

**FACULTY
OF MATHEMATICS
AND PHYSICS**
Charles University

DOCTORAL THESIS

Václav Březina

**Supramolecular complexes of oxoporphyrinogens
with organic molecules**

Department of Macromolecular Physics

Supervisor of the doctoral thesis: doc. RNDr. Lenka Hanyková, Dr.

Study programme: Physics

Study branch: 4F4: Biophysics, chemical
and macromolecular physics

Prague 2022

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In date

signature of the author

I would like to thank to my supervisor Lenka Hanyková for her support, patience, useful advice and help with NMR measurements. Many thanks belong to my consultant Jan Labuta for endless discussions about data processing, interpretation of the results and for arranging my internship at the National Institute for Materials Science (NIMS) in Tsukuba, Japan. Revisions of the text of this thesis by my supervisor and consultant are greatly appreciated. I am also grateful to Jonatan Hill from NIMS for providing the oxoporphyrinogen samples. Last but not least, I thank to my family and my girlfriend Katka for moral support.

Title: Supramolecular complexes of oxoporphyrinogens with organic molecules

Author: Václav Březina

Department: Department of Macromolecular Physics

Supervisor: doc. RNDr. Lenka Hanyková, Dr., Department of Macromolecular Physics

Abstract:

Oxoporphyrinogens are flat macrocyclic molecules possessing binding and protonation sites, and capable of light absorption in the visible region. These properties are prerequisites for a colorimetric molecular sensor, i.e. a specific detector of other molecules in the sample. In this work, we studied chromic properties of three oxoporphyrinogens, **OxP** and its partially (**Bz₂OxP**) and fully (**Bz₄OxP**) N-benzylated derivatives. Their colorimetric response to organic acids is caused by protonation and subsequent formation of supramolecular host-guest complex. We have shown that colorimetric sensitivity is highest for **OxP** and gradually weakens for **Bz₂OxP** and **Bz₄OxP** since the N-benzylation blocks the central binding sites, decreasing binding affinity of the oxoporphyrinogens. Furthermore, solvatochromic response of the oxoporphyrinogens to varying solvent polarity showed similar sensitivity decrease in **Bz₂OxP** and **Bz₄OxP**. The chromic and binding properties were studied by UV/vis and NMR spectroscopy, host-guest binding models were applied to describe the formation of oxoporphyrinogen-acid complexes.

In NMR spectra of protonated **OxP** and **Bz₂OxP**, we detected several dynamic processes, manifested as chemical exchange, such as prototropic tautomerization (i.e. change of the protonation site) or rotation of bulky side groups in **Bz₂OxP**. These processes were observed at different temperatures and acid concentrations, the corresponding transition rate coefficients were obtained by fitting of the NMR exchange lineshapes. In protonated **Bz₂OxP**, the effect of temperature obeys the Eyring equation, and the effect of acid concentration has been modeled using a kinetic model derived from competitive host-guest binding.

This thesis presents detailed mechanism of color changes in **OxP**, **Bz₂OxP** and **Bz₄OxP** and confirms potential application as colorimetric sensors in nonpolar media. Our NMR study of dynamic processes in the oxoporphyrinogens contributes to understanding the concentration dependence of transition rate coefficients, which is always system-specific. The presented models can be applied to a wide range of dynamic systems.

Keywords: oxoporphyrinogens, supramolecular complexes, colorimetric sensors, dynamic molecular processes

Contents

Preface	1
1 Introduction	2
1.1 Overview of tetrapyrroles	2
1.2 Properties of investigated oxoporphyrinogens	4
1.3 Aims and structure of this work	5
2 Quantitative description of supramolecular binding	7
2.1 Description of chemical reactions	7
2.1.1 Equilibrium constant	7
2.1.2 Meaning of activity coefficients	9
2.1.3 The rate law	10
2.1.4 Eyring equation	11
2.2 Host-guest binding	13
2.2.1 H:G 1:1 binding model	13
2.2.2 Aspects of multiple binding sites	14
2.2.3 Competitive host-ligand binding	17
2.2.4 Cooperativity	18
3 Nuclear magnetic resonance spectroscopy	22
3.1 Basic phenomena in NMR	22
3.1.1 Nuclei in magnetic field	22
3.1.2 Chemical shift	23
3.1.3 Relaxation	24
3.1.4 J-coupling	24
3.1.5 Bloch equations	24
3.1.6 Effect of radiofrequency field	25
3.2 Measurement of NMR	25
3.2.1 NMR spectrometer	25
3.2.2 Pulsed NMR experiment	26
3.2.3 Lorentzian spectral lineshape	27
3.3 Chemical exchange	28
3.3.1 Kinetics of spin states	30
3.3.2 Kinetics of two-state exchange	32
3.3.3 Kinetics of three-state exchange	33
3.3.4 Bloch-McConnell equations	34
3.3.5 Solution of Bloch-McConnell equations	35
3.3.6 Two-state exchange spectral lineshape	36
3.3.7 Three-state exchange spectral lineshape	37

3.3.8	Exchange schemes containing a fast process	38
3.4	NMR of host-guest systems	39
3.4.1	Chemical exchange in host-guest complexes	39
3.4.2	NMR titration binding isotherms	40
3.4.3	Determination of enantiomeric excess	43
4	UV/vis spectroscopy	48
4.1	Absorbance measurement	48
4.2	Beer-Lambert law of absorption	49
4.3	UV/vis titration experiment	49
4.4	Isosbestic points	50
4.5	Halochromism and solvatochromism	51
5	Singular value decomposition	53
5.1	Mathematical properties	53
5.1.1	Basics of SVD	53
5.1.2	Pseudoinverse and linear regression	55
5.2	Analysis of UV/vis spectral series via SVD	56
5.2.1	Linear combination of absorbing components	58
5.2.2	Determining the number of absorbing species	58
5.2.3	Transformation from abstract to real spectra	60
5.2.4	Combination coefficients	62
5.2.5	Decomposition to the absorbing components without SVD	63
5.2.6	SVD analysis – summary	64
5.3	Analogy with principal component analysis	64
5.3.1	Calculation of principal components	64
5.3.2	Comparison to SVD	65
5.3.3	Application	66
6	Chromic and binding properties of OxP derivatives	67
6.1	Halochromism of OxP	68
6.2	Halochromism of Bz₂OxP	70
6.3	Chromism of Bz₄OxP	71
6.4	Mechanism of halochromism	73
6.5	NMR analysis of host-guest binding	76
6.6	Solvatochromism of OxP	79
6.7	Solvatochromism of Bz₂OxP	81
6.8	Solvatochromism of Bz₄OxP	83
6.9	Chromism and binding – summary	84
7	Dynamic processes and chemical exchange in OxP derivatives	87
7.1	Dynamics of OxP	87
7.1.1	Prototropic tautomerization	87
7.1.2	Anion binding at the porphyrinogen center	90
7.1.3	Macrocyclic inversion	91
7.2	Dynamics of Bz₂OxP	94
7.2.1	Topomerization at hemiquinonoid groups	94
7.2.2	Anion binding at the porphyrinogen center	98
7.2.3	Prototropic tautomerization	99
7.3	Dynamic processes – summary	106

8 Conclusion	109
References	111
List of Abbreviations	122
A Overview of binding models	123
A.1 Binding model 1:1	123
A.2 Binding model 1:2	123
A.3 Binding model 1:3	124
A.4 Binding model 1:4	125
A.5 Competitive binding 1:1 host-ligand binding	125
A.5.1 First solution	126
A.5.2 Second solution	127
B Overview of NMR spectral lineshapes	128
B.1 Lineshape formulae and <i>Mathematica</i> codes	128
B.1.1 Lorentzian lineshape	128
B.1.2 Two-state exchange lineshape	129
B.1.3 Three-state exchange lineshape	131
B.1.4 Spectral lineshape of two J-coupled spins	133
B.2 Rapid derivation of <i>N</i> -state exchange lineshape in <i>Mathematica</i>	135
B.3 Rapid derivation of three-state populations	136
C Experimental details, models and data analysis	137
C.1 Properties of used chemicals	137
C.2 UV/vis spectroscopy measurements	137
C.3 RGB coordinates of UV/vis spectra	138
C.4 NMR spectroscopy measurements	138
C.5 NMR analysis of host-guest binding	139
C.6 Fitting of exchange lineshapes	141
C.7 Fitting of experimental data in Bz₂OxP + (<i>R</i>)-CSA system	144
C.8 Kinetic models of Bz₂OxP + (<i>R</i>)-CSA system	147
C.9 Prototropic tautomerization processes in Bz₂OxP	151
D SVD analysis	155
D.1 Bz₂OxP + (<i>R</i>)-CSA titration	155
D.2 OxP + TFA titration	156
D.3 Bz₂OxP + DFA titration	158
D.4 Bz₄OxP + DFA titration	159
D.5 OxP + DMF titration	160
D.6 Bz₂OxP + DMF at different vol.f.	162
D.7 Bz₄OxP + DMF at different vol.f.	163
E NMR spectra	164
E.1 Spectra of acids	164
E.2 OxP + DFA titration	165
E.3 OxP + (<i>R</i>)-CSA titration	166
E.4 Bz₂OxP + DFA titration	167
E.5 Bz₂OxP + (<i>R</i>)-CSA titration	168
E.6 Bz₂OxP + (<i>rac</i>)-CSA titration	169

E.7	Bz₄OxP + DFA titration	170
E.8	OxP + DFA 1.1 equiv. variable temperature	171
E.9	OxP + DFA 18 equiv. variable temperature	172
E.10	Bz₂OxP variable temperature	173
E.11	Bz₂OxP + DFA 35 equiv. variable temperature	174
E.12	Bz₂OxP + (<i>R</i>)-CSA 0.59 equiv. variable temperature	175
E.13	Bz₂OxP + (<i>R</i>)-CSA 14 equiv. variable temperature	176
E.14	Bz₂OxP + (<i>rac</i>)-CSA 0.5 equiv. variable temperature	177
E.15	Bz₂OxP + (<i>rac</i>)-CSA 17 equiv. variable temperature	178
E.16	OxP + DMF titration	179
E.17	Detection of OH signal in Bz₂OxP	180
E.18	Bz₄OxP assignment - HMBC	181

List of publications	182
-----------------------------	------------

Preface

Compounds capable of light absorption in the visible region as well as host-guest complexation with variety of ligands are perfect candidates for molecular sensors. These properties are present in porphyrins possessing rich conjugated system. Similar compounds, porphyrinogens (with sp^3 meso-carbon), lack the conjugated system and thus are colorless. However, they are capable of binding a broad spectrum of organic molecules, including metallic or non-metallic anions. The advantages of porphyrins and porphyrinogens are conveniently combined in *oxoporphyrinogens* (with sp^2 meso-carbon), which are subject of this work. Oxoporphyrinogen (**OxP**) and its derivatives have been previously studied from the perspective of colorimetric detection of various molecules, anions and acids in nonpolar media. Several dynamic molecular processes have also been observed in these compounds. Here, we provide detailed analysis of chromism and internal dynamics of **OxP** and its N-benzylated derivatives in the presence of organic acids. Their solvatochromic properties are also studied. This work is mainly based on our publications [1] and [2].

1 Introduction

1.1 Overview of tetrapyrroles

Compounds of our interest belong to a wide class called *tetrapyrroles*, containing four pyrrole rings typically joined by single-atom links at the pyrrole α -positions [3]. Scheme in Fig. 1.1 shows simplified classification of tetrapyrroles, presenting example molecules for each group (names of depicted compounds are highlighted by boldface in the text below). Basic division

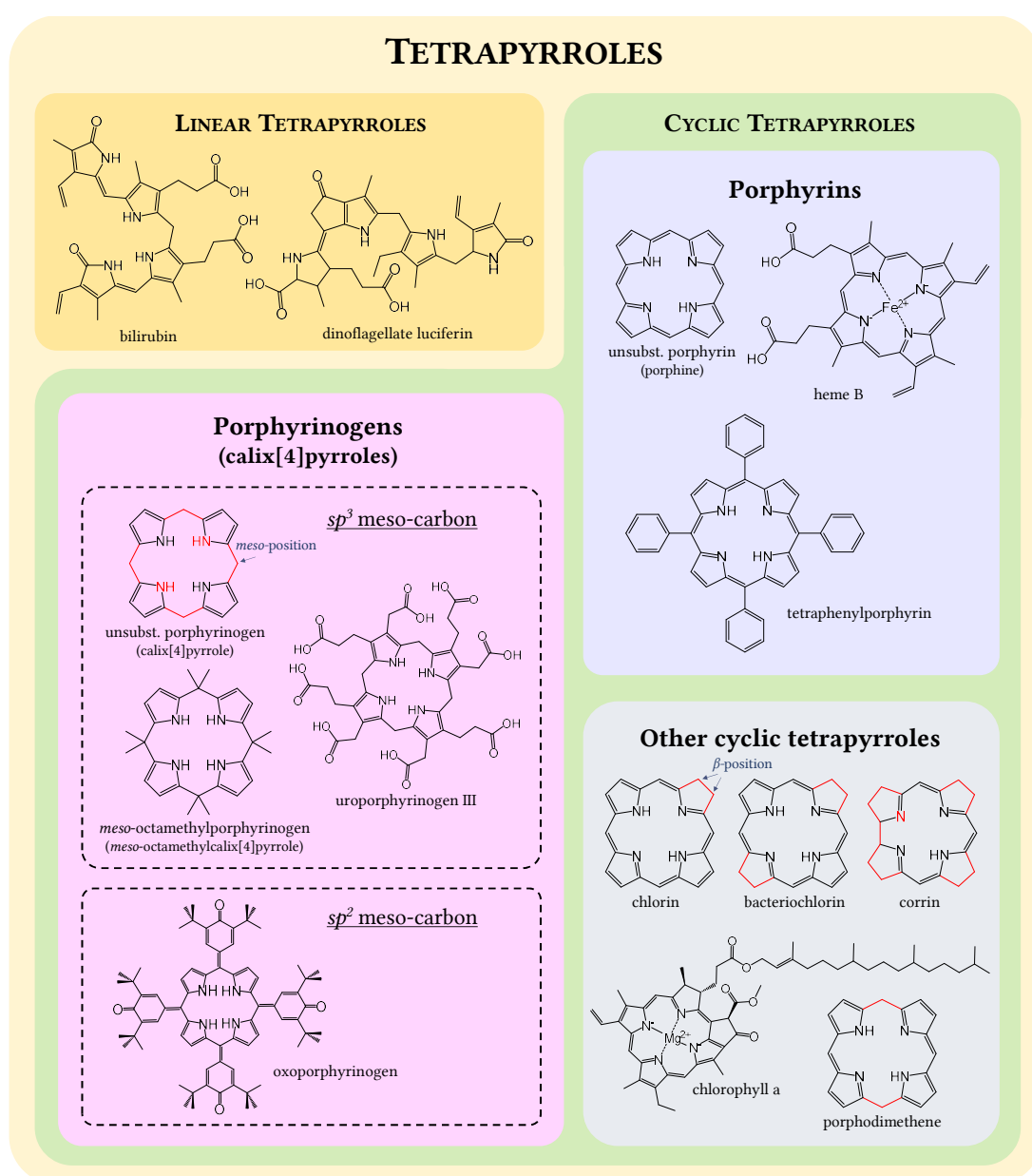


Figure 1.1: Simplified structure-based classification of tetrapyrroles. Several tetrapyrrole subclasses are omitted. Red color denotes differences of the unsubstituted rings from porphine.

comprises linear (open) tetrapyrroles and cyclic (closed) tetrapyrroles. Both groups represent essential natural dyes and metabolites [4, 5] as well as widely-studied artificially synthesized molecules [6].

Linear tetrapyrroles include bile pigments such as **bilirubin** or biliverdin, breakdown products of heme [5]. Another example is **dinoflagellate luciferin**, which produces bioluminescence ($\lambda_{\max} = 475$ nm) after oxidation catalyzed by luciferase enzyme [7, 8].

Cyclic tetrapyrroles are classified according to the type of macrocycle conjugation. Many of them are capable of coordinating a metal atom in their molecular center, which is denoted as ‘chelate’ or by a prefix ‘metallo-’. Prototypes of the tetrapyrrole macrocycle are fully conjugated *porphyrins*, whose unsubstituted form is called **porphine**. Their name is derived from a Greek word for ‘purple’. A well known natural metalloporphyrin is heme (most abundant form **heme B**), coordinating Fe^{2+} atom. Heme is located in red blood cells, bound to hemoglobin protein as a prosthetic group, and serves for transport of oxygen [9]. Man-made porphyrins have many applications in both chelated and free-base forms as discussed below. A common example is **tetraphenylporphyrin** (TPP), that has already appeared in thousands of scientific publications [10]. Interestingly, TPP is the first molecule, whose vibrational modes have been experimentally imaged [11].

Another group of cyclic tetrapyrroles are *porphyrinogens*, also called *calix[4]pyrroles* [12]. Their central nitrogens are bound by single bonds, and unsubstituted porphyrinogens contain four NH groups. While free-base (i.e. non-chelated) porphyrins such as TPP possess flat macrocycle, porphyrinogen center is more saddle-like. In this work we distinguish two types of porphyrinogens based on their *meso*-position (marked in Fig. 1.2): (i) “true” porphyrinogens with sp^3 hybridization on *meso*-carbons, (ii) *oxoporphyrinogens* with the same formal ring structure as (i) but with sp^2 *meso*-carbon hybridization. Structure of unsubstituted **porphyrinogen** shows the presence of two additional NH groups and interruption of conjugation in contrast to porphine (denoted by red color in Fig. 1.1). The incomplete conjugation renders the “true” porphyrinogens colorless. In the heme metabolism, porphyrinogens are precursors to porphyrins, which also gave them the name. Examples of metabolic porphyrinogens are **uroporphyrinogen III**, coproporphyrinogen III and protoporphyrinogen IX [5]. A frequently investigated synthetic porphyrinogen is **meso-octamethylcalix[4]pyrrole** [12]. This molecule enables various chemical modifications for different applications, further details are in the next section. The other group based on porphyrinogen structure and also subject of this work are *oxoporphyrinogens*, possessing sp^2 *meso*-carbon hybridization. **Oxoporphyrinogen (OxP)** has been first synthesized by Milgrom [13]. Oxoporphyrinogens are colored in contrast to “true” porphyrinogens due to the complete conjugation of the oxoporphyrinogen ring. On the other hand, both types of porphyrinogen share anion binding capabilities owing to the central NH groups.

There are other possible modifications of the porphyrin macrocycle illustrated in Fig. 1.1 (differences from the porphyrin ring are denoted by red color), with different degree of saturation at *meso*- or β -positions (marked in Fig. 1.1). Many of them constitute essential natural compounds. Reduced porphyrins (addition of hydrogens at the β -carbons) **chlorin** and **bacteriochlorin** constitute the core of chlorophylls (plants and algae) and bacteriochlorophylls¹ (photosynthetic bacteria) [14, 15]. These compounds are magnesium chelates and include many variants, e.g. **chlorophyll a** and chlorophyll b are the green plant pigments. Chlorophylls and bacteriochlorophylls serve for light harvesting, energy transfer and also electron transfer in natural photosystems [15]. Another macrocycle, **corrin**, is substantially saturated and lacks one of the methine link. A member of the corrins is cobalamin,² also known as vitamin B₁₂ [4]. As the name suggests, it is a chelate of cobalt. Interestingly, the already mentioned uroporphyrinogen

¹Some bacteriochlorophylls are based on the bacteriochlorin but others on the chlorin ring.

²There are several forms of cobalamin, e.g. hydroxycobalamin or cyanocobalamin [16].

III is common biosynthetic precursor to hemes, chlorophylls and also cobalamin [4, 15]. The last shown example of tetrapyrrole macrocycle is **porphodimethene**.

Applications of artificial cyclic tetrapyrroles utilize their binding properties (central NH binding sites), electronic structure or the ability to form π - π stacked structures [17]. Porphyrin-based compounds can serve as molecular probes for detection of cations or anions [18, 19]. Similarly, “true” porphyrinogens serve as probes³ of anions or neutral species [6, 12, 21]. TPP has wide applications in photochemistry, for example, as photosensitizer in light-induced oxidation⁴ [22]. Protonated TPP has the capability of forming helical J-aggregates for chiral recognition [23]. Further application include, for example, photoacoustic imaging probes [24, 25] or photodynamic therapy [26, 27].

1.2 Properties of investigated oxoporphyrinogens

Oxoporphyrinogen (**OxP**) and its derivatives⁵ N_{21}, N_{23} -bis(4-bromobenzyl)oxoporph. (**Bz₂OxP**) and also $N_{21}, N_{22}, N_{23}, N_{24}$ -tetrakis(4-bromobenzyl)oxoporph. (**Bz₄OxP**) are the subject of the current work. All three oxoporphyrinogens are shown in Fig. 1.2a–c, their X-ray crystal structures are given in Fig. 6.1 in Chapter 6. Atoms in the central macrocycle are numbered in the same way as porphine [3], see Fig. 1.2d. Figure 1.2 also shows several names of particular sites, which will be often used in the in chapters with experimental results. Methine links numbered 5, 10, 15 and 20 are denoted as *meso* positions. Labeling of α (1, 4, 6, 9, ...) and β (2, 3, 7, 8, ...) positions is derived from that of pyrrole in Fig. 1.2e. The bulky side groups are denoted as hemiquinonoid or cyclohexadienilidene, the corresponding hydrogens at *ortho*-position will be referred to as *ortho*-H. Carbonyl C=O sites can be protonated by an acid, the central NH groups serve as binding site for anions or neutral molecules. Next to the carbonyl, two *tert*-butyl groups are located, consisting of an sp^3 carbon with three methyls. Central nitrogen atoms are N-alkylated with bromobenzyl groups. Noteworthy, chlorin and porphyrin macrocycles show NH tautomerism, where the two central hydrogens “jump” among the four nitrogen atoms [28, 29]. Such scenario is, however, impossible in **Bz₂OxP** since the N-alkylation prevents any rearrangements.

³For UV/vis probing, a chromophore is attached to colorless porphyrinogens or displacement assay technique is used (weakly bound colored anion is displaced by more strongly binding analyte, which leads to color change of the colored anion [20]).

⁴High triplet-state energy of tetraphenylporphyrin (142.3 kJ/mol) enables excitation of triplet oxygen to its singlet state (94.2 kJ/mol) [22].

⁵The 4-bromobenzyl derivatives are more soluble in chloroform than the corresponding non-brominated derivatives.

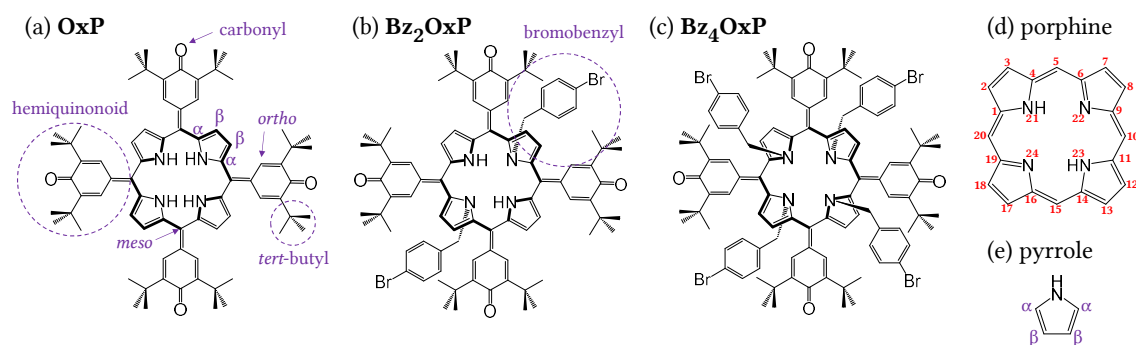


Figure 1.2: Oxoporphyrinogens, notation of important sites and groups. (a–c) Oxoporphyrinogens investigated in this work. (d) Labeling of pyrrole positions transferable to other cyclic tetrapyrroles, (e) labeling of pyrrole positions.

Various aspects of **OxP** and **Bz₂OxP** were studied by Jonathan Hill, Jan Labuta and coworkers [30–42]. Colorimetric detection of anions, solvents, acid or bases in solution [30–33] was observed as well as detection of trace water in tetrahydrofuran [34]. Nanocomposite film with **OxP** provides naked-eye discrimination of methanol from ethanol [35]. Another aspect is observation of dynamic processes present in oxoporphyrinogens [36, 37]. Finally, a novel approach was developed for NMR determination of enantiomeric excess of chiral analytes using oxoporphyrinogens (in particular **OxP** and **Bz₂OxP**) as chirality sensors [37–42]. This method is described in detail in Sec. 3.4.3 (determination of enantiomeric excess is also possible in TPP [39] or other compounds [43, 44]).

1.3 Aims and structure of this work

This work is based on our publications [1] and [2] including figures and parts of text. It is also a continuation of the master's thesis⁶ [45]. Here, we investigate oxoporphyrinogen (**OxP**) and two N-alkylated derivatives denoted as **Bz₂OxP** and **Bz₄OxP** and their interaction with organic acids. In particular, chiral camphorsulfonic acid (CSA), achiral difluoroacetic acid (DFA) and trifluoroacetic acid (TFA) are used. The effect of chirality on NMR spectra has been taken into account, especially in fast exchange processes. Furthermore, solvatochromic color changes of the oxoporphyrinogens are observed in the mixture of polar dimethylformamid (DMF) and nonpolar chloroform. Our aims are as follows:

- Determination of the effect of N-alkylation on color changes. Color changes of **OxP**, **Bz₂OxP** and **Bz₄OxP** induced by CSA or DFA will be studied in chloroform solvent.
- Analysis of UV/vis spectral changes by singular value decomposition. This type of analysis has not been used before in our systems, it allows determination of the exact number of absorbing species, their UV/vis spectra (and the corresponding degree of protonation) and the corresponding populations.
- Determination of binding constants. Binding isotherms for UV/vis and NMR titration experiments will be constructed and fitted by appropriate host-guest binding models.
- Investigation of solvatochromism. Color changes induced by solvent polarity change or hydrogen bonding with the oxoporphyrinogens will be studied in the DMF/chloroform mixture.
- Identification of dynamic molecular processes such as rotation of side groups or tautomerization. These processes will be observed in NMR spectra in a form of so-called chemical exchange.
- Derivation of analytical lineshape for three-state chemical exchange. The lineshape will be calculated from the Bloch-McConnell equations and will be used to retrieve the transition rate coefficients from the exchange spectra by fitting.
- Determination of transition rate coefficients and energy barriers as obtained from lineshape fitting of experimentally obtained exchange spectra. Concentration dependence of the transition rate coefficients will be described by host-guest binding models, and temperature dependence of the transition rate coefficients will be analyzed in terms of Eyring equation.

⁶Available at <https://dspace.cuni.cz/handle/20.500.11956/67480>.

This work consists of theoretical part (Chapter 1–5) and experimental part (Chapter 6–7). Extensive appendices provide an overview of binding models, NMR lineshapes, details about experimental measurements, data analyses and measured NMR spectra. Codes for implementation in *Mathematica* software are also included.

In particular, structure of this work is as follows: The current **Chapter 1** introduces the investigated oxoporphyrinogens and gives an overview of the wider class of tetrapyrroles. **Chapter 2** provides theory of supramolecular binding and chemical kinetics. **Chapter 3** introduces the NMR spectroscopy with focus on chemical exchange. Analytical three-state exchange lineshape, derived in our publication [2], is also shown in this chapter. Basics of UV/vis spectroscopy are given in **Chapter 4**. In **Chapter 5**, comprehensive description of singular value decomposition and its application to UV/vis spectra is presented. In **Chapter 6**, chromic and binding properties of the oxoporphyrinogens are studied experimentally with UV/vis and NMR spectroscopy. Results from our publication [1] are presented here in an extended form. In the next experimental chapter, **Chapter 7**, dynamic processes in **OxP** and **Bz₂OxP** are studied. Part of this research is published in our article [2], nevertheless, large portion of these results remains unpublished. Results of this work and further applications are concisely summarized in **Chapter 8**.

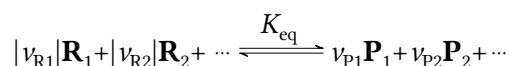
2 Quantitative description of supramolecular binding

Supramolecular structures encompass molecular assemblies without formation of covalent bonds. The assemblies are held together by hydrogen bonding or by electrostatic forces. These weak binding forces enable incessant formation and dissociation of the supramolecular complex, implying that the supramolecular systems are dynamic even after they achieve equilibrium. Formation and dissociation of supramolecular complexes follow the general laws of kinetics and equilibrium for chemical reactions, which are explained in this chapter. From the theoretical point of view, the matter is described in detail in textbooks of Atkins [46] or McQuarrie and Simon [47]. From the practical point of view, the textbook of Connors [48] is the basic source.

2.1 Description of chemical reactions

2.1.1 Equilibrium constant

A general chemical reaction converting reactants \mathbf{R}_j to products \mathbf{P}_j is described by the following scheme



Scheme 2.1: General scheme of a chemical reaction.

where v_j are stoichiometric numbers,¹ $v_j < 0$ for reactants (educts) and $v_j > 0$ for products [46]. The quantity K_{eq} is called *equilibrium constant* (also denoted as binding or stability constant) [48].

Chemical reactions usually take place in contact with thermal reservoir at constant pressure. Therefore, description using Gibbs free energy potential $dG = -SdT + Vdp + \mu_{\mathbf{R}1}dn_{\mathbf{R}1} + \mu_{\mathbf{R}2}dn_{\mathbf{R}2} + \mu_{\mathbf{P}2}dn_{\mathbf{P}2} + \mu_{\mathbf{P}1}dn_{\mathbf{P}1} \dots$ is convenient since the contact with the reservoir ensures $dT = dp = 0$. Thus, for constant temperature and pressure,

$$(dG)_{p,T} = \mu_{\mathbf{R}1}dn_{\mathbf{R}1} + \mu_{\mathbf{R}2}dn_{\mathbf{R}2} + \mu_{\mathbf{P}2}dn_{\mathbf{P}2} + \mu_{\mathbf{P}1}dn_{\mathbf{P}1} \dots, \quad (2.1)$$

where n_j is the amount of component j (in mol) and μ_j is the chemical potential of component j defined as $\mu_j = \frac{\partial G}{\partial n_j}$ (in J.mol⁻¹).

Number of molecules n_j in the system deviate from their initial values n_j^i while approaching to their equilibrium values. Thus, $n_j = n_j^i + v_j \xi$, where ξ is the extent of reaction (in moles). Hence, the Gibbs energy change can be expressed as $(dG)_{p,T} = \sum_j v_j \mu_j d\xi$. Also from Eq. (2.1)

$(dG)_{p,T} = \left(\frac{\partial G}{\partial \xi} \right)_{p,T} d\xi$, therefore by comparison

$$\left(\frac{\partial G}{\partial \xi} \right)_{p,T} = \sum_j v_j \mu_j. \quad (2.2)$$

In textbooks [47], the right-hand side of Eq. (2.2) is denoted as $\Delta_r G$, in particular,

$$\Delta_r G = \sum_j v_j \mu_j. \quad (2.3)$$

¹stoichiometric numbers <https://goldbook.iupac.org/terms/view/S06025>

Note that $\Delta_r G$ equals to a derivative of the Gibbs energy or for a difference of chemical potentials (multiplied by appropriate stoichiometric numbers).

Principle of minimum Gibbs potential states that any unconstrained internal parameter (in our case the reaction extent ξ) minimizes the Gibbs potential in equilibrium [49], yielding the condition for equilibrium

$$\left(\frac{\partial G}{\partial \xi}\right)_{p,T} = 0 \implies \Delta_r G = \sum_j \nu_j \mu_j = 0. \quad (2.4)$$

This is a general condition valid for liquids as well as gases and solids. If $\Delta_r G < 0$, the reaction spontaneously proceeds in forward direction and if $\Delta_r G > 0$, the reaction spontaneously proceeds in backward direction until equilibrium is achieved at $\Delta_r G = 0$. Hence, the crucial property of the system of interest is the exact form of chemical potential as function of concentration of all present chemical species.

The chemical potential of j -th component of multi-component ideal gas is expressed as $\mu_j = \mu_j^\circ(p, T) + RT \ln x_j$. This also holds for the ideal liquid solution, where the liquids do not interact [50]. For non-ideal liquid solution the chemical potential can be written in similar manner as

$$\mu_j = \mu_j^\circ(p, T) + RT \ln a_j, \quad (2.5)$$

where a_j is the *activity* of j -th component defined so that Eq. (2.5) is exactly valid for a non-ideal solution. Values of the *standard chemical potential* μ_j° are by definition related to a particular reference state in which $a_j = 1$. The choice of reference state includes values of pressure, temperature and concentrations of the experimental solution. As a common practice, the standard state is chosen at infinite dilution of the solutes (although other choices are possible, see page 30 in [48]). The activity can be directly related to concentration by

$$a_j = \gamma_j [j], \quad (2.6)$$

where $[j]$ is the concentration of j -th component. The activity coefficients γ_j correct for non-ideality of the solution. For more details, see the next section.

Using the equation for chemical potential, Eq. (2.5), the equilibrium condition in Eq. (2.4) is rewritten as

$$\ln \frac{a_{\mathbf{P}1}^{\nu_{\mathbf{P}1}} a_{\mathbf{P}2}^{\nu_{\mathbf{P}2}} \dots}{a_{\mathbf{R}1}^{\nu_{\mathbf{R}1}} a_{\mathbf{R}2}^{\nu_{\mathbf{R}2}} \dots} = -\frac{\sum_j \nu_j \mu_j^\circ}{RT}. \quad (2.7)$$

Although the activities depend on concentration, the right-hand side and consequently the left-hand side depend only on pressure and temperature. Therefore, we define the *equilibrium constant* as

$$K_{\text{eq}} = \exp\left(-\frac{\Delta_r G^\circ}{RT}\right), \quad (2.8)$$

where $\Delta_r G^\circ = \sum_j \nu_j \mu_j^\circ$ is the *standard Gibbs energy of reaction* (in $\text{J}\cdot\text{mol}^{-1}$). Hence, the standard Gibbs energy can be calculated from the equilibrium constant using

$$\Delta_r G^\circ = -RT \ln K_{\text{eq}}. \quad (2.9)$$

The equilibrium constant binds together activities of all components in equilibrium,

$$K_{\text{eq}} = \frac{a_{\mathbf{P}1}^{\nu_{\mathbf{P}1}} a_{\mathbf{P}2}^{\nu_{\mathbf{P}2}} \dots}{a_{\mathbf{R}1}^{\nu_{\mathbf{R}1}} a_{\mathbf{R}2}^{\nu_{\mathbf{R}2}} \dots}. \quad (2.10)$$

For dilute solutions ($\gamma_j = 1$) the equilibrium constant is equal to

$$K_{\text{eq}} = \frac{[\mathbf{P}1]^{\nu_{\mathbf{P}1}} [\mathbf{P}2]^{\nu_{\mathbf{P}2}} \dots}{[\mathbf{R}1]^{\nu_{\mathbf{R}1}} [\mathbf{R}2]^{\nu_{\mathbf{R}2}} \dots}. \quad (2.11)$$

To conclude, value of $\Delta_r G$ decides if the reaction is in equilibrium or proceeds in forward/backward direction. On the other hand, $\Delta_r G^\circ$ determines the value of the equilibrium constant K_{eq} and consequently the equilibrium concentration of all chemical species. Because $\Delta_r G^\circ = \Delta_r H^\circ - T\Delta_r S^\circ$, it is possible to determine the energetic contribution $\Delta_r H^\circ$ and entropic contribution $\Delta_r S^\circ$ from temperature dependence of K_{eq} . If $\Delta_r H^\circ$ and $\Delta_r S^\circ$ are independent of temperature, plot of $\ln K_{\text{eq}}$ vs. $1/T$ yields straight line with the slope $-\Delta_r H^\circ/R$ and intercept $\Delta_r S^\circ/R$ according to

$$\ln K_{\text{eq}} = -\frac{\Delta_r H^\circ}{R} \frac{1}{T} + \frac{\Delta_r S^\circ}{R}. \quad (2.12)$$

An equilibrium constant is defined not only for the overall reaction scheme according to Eq. (2.10) but also for every elementary reaction. Each elementary reaction has its own independent equilibrium constant and corresponding standard Gibbs energy of reaction.

2.1.2 Meaning of activity coefficients

Activity coefficients can be defined in multiple ways, depending on the use of a particular concentration scale (see page 33 in [48]). Besides the molar concentration scale ($[j] = n_j/V$ [mol/l]) implying $a_j = \gamma_j[j]$, also molar fraction scale ($x_j = n_j/\sum n_j$ [mol/mol]) implying $a_j = \gamma_j^x x_j$ can be used. The latter approach is useful for interpretation of activity coefficients. The Gibbs energy of multi-component solution $G = \sum_j n_j \mu_j$ can be reformulated as

$$\frac{G(T, p, n_j)}{n} = \underbrace{\sum_j x_j \mu_j^\circ(T, p, n_j)}_{\text{pure components}} + \underbrace{RT \sum_j x_j \ln x_j}_{\text{mixing entropy}} + \underbrace{RT \sum_j x_j \ln \gamma_j^x}_{\text{non-ideality}}, \quad (2.13)$$

where R is the gas constant. The first two terms in Eq. (2.13) correspond to an ideal solution including mixing entropy but not mixing enthalpy (non-interacting particles). The last term stands for enthalpic and entropic contributions from non-ideality, which explains the physical meaning of activity coefficients. This expression for Gibbs energy including activity coefficients is important not only for chemical reactions but also for other thermodynamic applications involving chemical potential, e.g. vapor pressure in vapor-liquid equilibrium or construction of phase diagrams [51].

Values of activity coefficients and exact meaning of μ_j° depend on the choice of reference state. If the reference state is defined at pure components, then the limit $x_j \rightarrow 1$ implies $\gamma_j^x \rightarrow 1$, and μ_j° is equal to molar Gibbs energy of molecules of the pure component j , which interact with other molecules of the same type. On the other hand, if the reference state is at infinitely diluted components, then $[j] \rightarrow 0$ implies $\gamma_j \rightarrow 1$, and μ_j° describes isolated molecule of the j -th component interacting with surrounding molecules of the medium (solvent). Illustration of actual measured activity coefficients γ_j^x in two-component solution is given in Fig. 2.1. Activity coefficients of solution mixtures can be obtained experimentally by vapor-liquid equilibria measurement (a standard technique) or other methods, such as IR spectroscopy [52–54].

The activity coefficients depend on temperature, pressure and $N_{\text{spc}} - 1$ concentration/molar fraction values, where N_{spc} is the number of present chemical species. Gibbs-Duhem relation [49, 50] at constant temperature and pressure, $\sum n_j d\mu_j = 0$, binds together chemical potentials of a multi-component mixture. After the substitution $\mu_j = \mu_j^\circ + RT \ln \gamma_j$, the activity coefficients are related by

$$\sum x_j d \ln \gamma_j = 0. \quad (2.14)$$

For example, for two-component system with γ as a function of x_2 in the form $\gamma_1 = \gamma_1(x_2)$, $\gamma_2 = \gamma_2(x_2)$ it follows (page 53 in [50]) that

$$\frac{d \ln \gamma_2}{dx_2} = -\frac{x_1}{x_2} \frac{d \ln \gamma_1}{dx_2}. \quad (2.15)$$

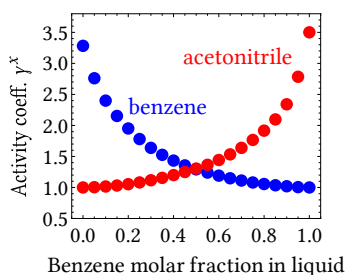


Figure 2.1: Activity coefficients for benzene-acetonitrile solution at 298 K obtained from vapor pressure measurements at vapor-liquid equilibrium. Values adopted from [52].

This condition must be obeyed for construction of models for activity coefficients. Simple Margules model employs polynomials in x_j , but for practical purposes, more realistic models are used, e.g. NRTL (non-random two-liquid), UNIQUAC (universal quasichemical) or UNIFAC (UNIQUAC Functional-group Activity Coefficients) models [50, 51, 55]. Dortmund Data Bank is a database of thermodynamic properties, where the activity coefficients of many chemical mixtures can be obtained [56].

2.1.3 The rate law

The central relation governing rates in chemical kinetics, *law of mass action*, was discovered in the second half of 19th century by contributions of Guldberg, Waage, van't Hoff and Horstmann [57, 58]. This law describes elementary reactions (not overall reactions) as collisions between the reacting molecules. The most common types of reactions are unimolecular reactions (e.g. decay of a complex or a conformational change) and bimolecular reactions (e.g. complex formation). Probability of the collisions is proportional to concentration of both reactants. If we consider a bimolecular reaction $\mathbf{R}_1 + \mathbf{R}_2 \rightleftharpoons \mathbf{P}_1 + \mathbf{P}_2$, then formally

$$\begin{aligned} \text{forward reaction rate} &= \kappa_{\text{fwd}}[\mathbf{R}_1][\mathbf{R}_2], \\ \text{backward reaction rate} &= \kappa_{\text{rev}}[\mathbf{P}_1][\mathbf{P}_2], \end{aligned} \quad (2.16)$$

where κ_{fwd} and κ_{rev} are the *reaction rate coefficients* for forward and backward (reverse) chemical reaction, respectively.² These coefficients depend on temperature and pressure but not on concentrations of the chemical species in the system. Difference of forward and backward reaction rates defines the *reaction rate* v [59] as

$$v = \kappa_{\text{fwd}}[\mathbf{R}_1][\mathbf{R}_2] - \kappa_{\text{rev}}[\mathbf{P}_1][\mathbf{P}_2]. \quad (2.17)$$

Time change of a particular chemical species concentration can be obtained from reaction rates of corresponding elementary reactions, for example in the case of one elementary reaction, $d[\mathbf{R}_1]/dt = -v$ or $d[\mathbf{P}_2]/dt = v$. Then, construction of the governing differential equations is

²Although reaction rate coefficients κ_{\bullet} are usually denoted in the literature as ' k_{\bullet} ', here, we apply the Greek letter ' κ ' to avoid confusion with transition rate coefficients defined in Chapter 3.

straightforward,

$$\begin{aligned}\frac{d[\mathbf{R}_1]}{dt} &= -\kappa_{\text{fwd}}[\mathbf{R}_1][\mathbf{R}_2] + \kappa_{\text{rev}}[\mathbf{P}_1][\mathbf{P}_2], \\ \frac{d[\mathbf{R}_2]}{dt} &= -\kappa_{\text{fwd}}[\mathbf{R}_1][\mathbf{R}_2] + \kappa_{\text{rev}}[\mathbf{P}_1][\mathbf{P}_2], \\ \frac{d[\mathbf{P}_1]}{dt} &= -\kappa_{\text{rev}}[\mathbf{P}_1][\mathbf{P}_2] + \kappa_{\text{fwd}}[\mathbf{R}_1][\mathbf{R}_2], \\ \frac{d[\mathbf{P}_2]}{dt} &= -\kappa_{\text{rev}}[\mathbf{P}_1][\mathbf{P}_2] + \kappa_{\text{fwd}}[\mathbf{R}_1][\mathbf{R}_2].\end{aligned}\quad (2.18)$$

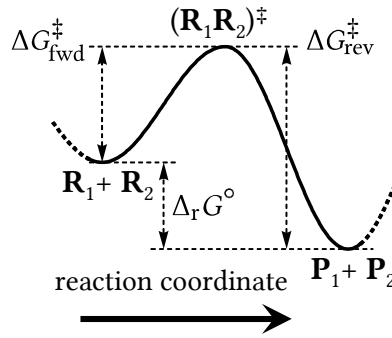
It is a set of first order nonlinear differential equations, which can be solved numerically. Notably, the equilibrium constant in dilute limit can be derived from the chemical kinetics because in equilibrium both reaction rates are equal, i.e. $v = 0$ (or equivalently, time derivatives in Eq. (2.18) are equal to zero), yielding

$$K_{\text{eq}} = \frac{\kappa_{\text{fwd}}}{\kappa_{\text{rev}}} = \frac{[\mathbf{P}_1][\mathbf{P}_2]}{[\mathbf{R}_1][\mathbf{R}_2]}.\quad (2.19)$$

All governing reactions must conserve the total mass, hence summing all subequations in Eq. (2.18) must yield $\frac{d}{dt}([\mathbf{R}_1] + [\mathbf{R}_2] + [\mathbf{P}_1] + [\mathbf{P}_2]) = 0$, i.e. all right-hand side terms of the subequations must sum to zero. The rate law is applicable to describe chemical kinetics in wide range of systems. For example, in enzyme-catalyzed reactions the chemical kinetics usually assumes short initial pre-steady-state mode (v dramatically changes in time) followed by approximate steady-state (v approximately constant) [2, 60].

2.1.4 Eyring equation

Temperature dependence of reaction rate coefficients is obtained from the transition state theory of chemical reactions [46, 47, 61, 62]. Energetic scheme for bimolecular reaction is given in the following scheme. The reaction coordinate is an abstract 1D representation of geometric



Scheme 2.2: Transition state in an elementary bimolecular reaction.

parameters of the two reacting molecules along the reaction pathway. The forward reaction progresses when reactants \mathbf{R}_1 and \mathbf{R}_2 collide with energy big enough to overcome the barrier $\Delta G_{\text{fwd}}^{\ddagger}$ and reach the transition state $(\mathbf{R}_1\mathbf{R}_2)^{\ddagger}$ and vice versa for the backward reaction. The transition state is a saddle point in the potential energy surface (depending on all atomic coordinates in the system), while states $\mathbf{R}_1 + \mathbf{R}_2$ or $\mathbf{P}_1 + \mathbf{P}_2$ are minima.

The reaction rate coefficients κ_{fwd} and κ_{rev} are directly related to barrier heights $\Delta G_{\text{fwd}}^{\ddagger}$ and $\Delta G_{\text{rev}}^{\ddagger}$, respectively according to the Eyring equation [63]

$$\kappa_j = \frac{\eta k_{\text{Boltz}} T}{h} \exp\left(-\frac{\Delta G_j^{\ddagger}}{RT}\right),\quad (2.20)$$

where ΔG_j^\ddagger is the Gibbs energy barrier (in $\text{J}\cdot\text{mol}^{-1}$) corresponding to κ_j , η is the transition probability, h is the Planck constant and R is the molar gas constant. The transition probability η describes how likely the system moves to the other minimum after it reaches the transition state, it is usually set to unity because it cannot be measured experimentally. The Eyring equation provides temperature dependence of reaction rate coefficients.

Relationship of the equilibrium constant and reaction rate coefficients $K_{\text{eq}} = \kappa_{\text{fwd}}/\kappa_{\text{rev}}$ (for dilute solutions, Eq. (2.19)) provides the relationship between the barrier heights and the standard reaction Gibbs energy. Using the definition of equilibrium constant (Eq. (2.8)) and the Eyring equation for κ_{fwd} and κ_{rev} , it follows that

$$\exp\left(-\frac{\Delta_r G^\circ}{RT}\right) = K_{\text{eq}} = \frac{\kappa_{\text{fwd}}}{\kappa_{\text{rev}}} = \exp\left(-\frac{\Delta G_{\text{fwd}}^\ddagger - \Delta G_{\text{rev}}^\ddagger}{RT}\right). \quad (2.21)$$

Comparison of the first and last expressions yields

$$\Delta_r G^\circ = \Delta G_{\text{fwd}}^\ddagger - \Delta G_{\text{rev}}^\ddagger \quad (2.22)$$

implying that $\Delta_r G^\circ$ is the difference between minima in the Gibbs energy profile corresponding to $\mathbf{R}_1 + \mathbf{R}_2$ and $\mathbf{P}_1 + \mathbf{P}_2$, see Scheme 2.2.

Variable temperature measurements are a standard means to obtain the value of energy barrier of a molecular process. Similarly to the standard reaction Gibbs energy, the energy barrier can be decomposed to its enthalpic and entropic contributions,

$$\Delta G_j^\ddagger = \Delta H_j^\ddagger - T\Delta S_j^\ddagger. \quad (2.23)$$

Using this decomposition in the Eyring equation, the measured dependence of $\kappa_j(T)$ can be visualized in Eyring plot of $\ln(\kappa_j/T)$ versus $1/T$. Assuming ΔH_j^\ddagger and ΔS_j^\ddagger are constant along the reaction coordinate, the Eyring plot is a straight line because

$$\ln \frac{\kappa_j}{T} = -\frac{\Delta H_j^\ddagger}{R} \frac{1}{T} + \ln \frac{k_{\text{Boltz}}}{h} + \frac{\Delta S_j^\ddagger}{R} \quad (2.24)$$

as illustrated in Fig. 2.2.

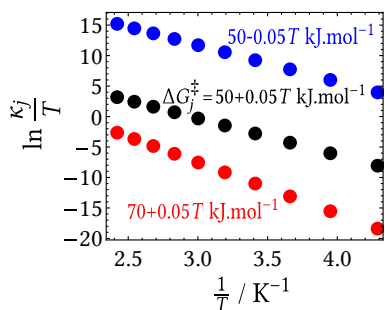


Figure 2.2: Illustration of Eyring plot for various values of $\Delta G_j^\ddagger = \Delta H_j^\ddagger - T\Delta S_j^\ddagger$. Temperatures range from -40 to 140 $^\circ\text{C}$ with 20 $^\circ\text{C}$ step.

Values of ΔH_j^\ddagger and ΔS_j^\ddagger can be obtained from slope and intercept, respectively, from the Eyring plot according to Eq. (2.24). Possible deviations from linearity in the Eyring plot can be caused by the presence of multiple processes with different energy barriers, which sum to the observed apparent reaction rate coefficient [2]. Another possible cause is dependence of ΔH_j^\ddagger and ΔS_j^\ddagger on temperature, which can be modeled by additional terms with molar heat capacity change ΔC_p^\ddagger in the Eyring equation [64, 65].

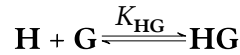
2.2 Host-guest binding

In many cases of supramolecular binding, the interacting molecules are denoted as host (usually a large molecule) and guest (usually a small molecule), nevertheless, the choice is rather arbitrary. Common binding mechanisms are, for example, hydrogen bonding, ion pairing or formation of coordination complexes (metal-ligand interaction) [66]. The host molecule often contains hydrogen bond donor sites or Lewis basic donor atoms (with lone electron pair). On the other hand, the guest molecule often contains hydrogen bond acceptor sites or Lewis acidic acceptor atoms [67]. Examples of host molecules include proteins, cage molecules, cyclic molecules with central cavity (e.g. crown ethers or cyclodextrin) or zeolites. Guest molecules include, for example, metal cations, halide anions, hormones, pheromones and other small organic molecules with functional group(s) [67].

During the construction of host-guest binding models in the following sections, infinite dilution will be assumed and consequently, all activity coefficients are set to unity (i.e. $\gamma_j = 1$). As a result, these models cannot describe the host-guest binding in the whole concentration range. The formalism of host-guest binding can be readily used for description of protonation with an acid in nonpolar medium, where the cation and anion originating from the acid are held closely together due to electrostatic forces. In polar medium (e.g. in water), the cation and anion strongly interact with the medium thus moving freely in the solution [68]. In addition, ionic species in polar solvent strongly behave non-ideally and the approximation $\gamma_j \approx 1$ is not valid even for dilute solutions.

2.2.1 H:G 1:1 binding model

In this work, host molecules are consistently denoted as **H** and guest molecules as **G**. Although host-guest complexes $\mathbf{H}_m\mathbf{G}_n$ with arbitrary stoichiometry can generally exist, in this work we assume the presence of only one host in the complex. The most elementary interaction is formation of **HG** complex with 1:1 **H:G** stoichiometry described by the following scheme



Scheme 2.3: Host-guest binding scheme (**H:G** 1:1 stoichiometry).

Binding strength of the complex is characterized by the equilibrium constant $K_{\mathbf{HG}}$ defined as $K_{\mathbf{HG}} = \exp\left(-\frac{\Delta_r G^\circ}{RT}\right)$, where $\Delta_r G^\circ = \mu_{\mathbf{HG}}^\circ - \mu_{\mathbf{H}}^\circ - \mu_{\mathbf{G}}^\circ$. According to Eq. (2.11), the equilibrium constant is related to concentrations of the chemical species. Together with the balance of total mass, the complete set of equations is obtained,

$$K_{\mathbf{HG}} = \frac{[\mathbf{HG}]}{[\mathbf{H}][\mathbf{G}]}, \quad (2.25a)$$

$$[\mathbf{H}]_t = [\mathbf{H}] + [\mathbf{HG}], \quad (2.25b)$$

$$[\mathbf{G}]_t = [\mathbf{G}] + [\mathbf{HG}], \quad (2.25c)$$

where $[\bullet]_t$ denotes total concentration. Then the concentrations can be calculated analytically as

$$[\mathbf{H}] = \frac{K_{\text{HG}}([\mathbf{H}]_t - [\mathbf{G}]_t) - 1 + \sqrt{[K_{\text{HG}}([\mathbf{H}]_t - [\mathbf{G}]_t) - 1]^2 + 4K_{\text{HG}}[\mathbf{H}]_t}}{2K_{\text{HG}}}, \quad (2.26a)$$

$$[\mathbf{HG}] = \frac{K_{\text{HG}}([\mathbf{H}]_t + [\mathbf{G}]_t) + 1 - \sqrt{[K_{\text{HG}}([\mathbf{H}]_t - [\mathbf{G}]_t) - 1]^2 + 4K_{\text{HG}}[\mathbf{H}]_t}}{2K_{\text{HG}}}, \quad (2.26b)$$

$$[\mathbf{G}] = \frac{K_{\text{HG}}([\mathbf{G}]_t - [\mathbf{H}]_t) - 1 + \sqrt{[K_{\text{HG}}([\mathbf{H}]_t - [\mathbf{G}]_t) - 1]^2 + 4K_{\text{HG}}[\mathbf{H}]_t}}{2K_{\text{HG}}}. \quad (2.26c)$$

The above equations enable calculation of concentrations of all species when the total concentrations and K_{HG} value are known. On the other hand, measurement of concentration of the chemical species can be used for determination of K_{HG} . The effect of different K_{HG} values is illustrated in Fig. 2.3. Implementation of Eq. (2.26) in *Mathematica* is given in Appendix A.1.

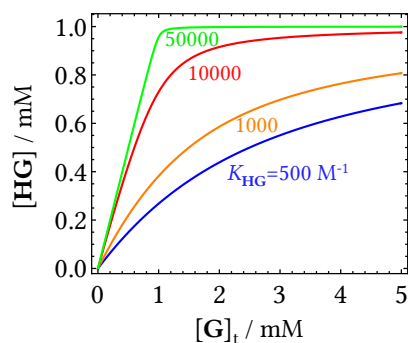
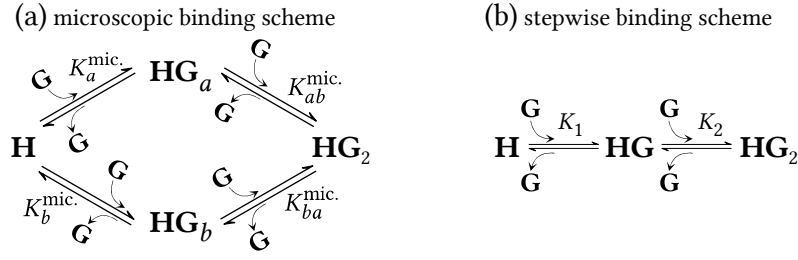


Figure 2.3: Concentration of complexed host in host-guest 1:1 binding model for different equilibrium constants. Value of $[\mathbf{H}]_t = 1$ mM is kept constant, $[\mathbf{G}]_t$ is changing.

The host-guest binding model does not account for interactions with the medium \mathbf{M} (i.e. solvent), which can be approximately described as $\mathbf{HM} + \mathbf{GM} \rightleftharpoons \mathbf{HG} + \mathbf{MM}$. Hence, the host-guest binding model is relevant only in situations, where the contributions of medium-medium and medium-solute (i.e. medium-host and medium-guest) interaction energies to the standard reaction Gibbs energy are negligible. Consequently, host-guest models cannot be applied in the case of weak binding corresponding to values of $K_{\text{HG}} \lesssim 1$ (see pages 6 and 13 in [48]). For example, hydrophobic association in polar solvents is governed mainly by strong medium-medium interactions.

2.2.2 Aspects of multiple binding sites

Host molecules can possess more than one binding site for the guest. Let us consider two binding sites denoted as a and b , which enables formation of four chemical species for the host, namely \mathbf{H} , \mathbf{HG}_a , \mathbf{HG}_b and \mathbf{HG}_2 as shown in Scheme 2.4a. The species \mathbf{HG}_a and \mathbf{HG}_b are denoted according to the occupied binding site.



Scheme 2.4: Host-guest 1:2 binding model. (a) microscopic binding scheme described by microscopic binding constants, (b) stepwise binding scheme described by stepwise binding constants.

Each binding step in Scheme 2.4a is described by a *microscopic binding constant* defined as

$$K_a^{\text{mic.}} = \frac{[\text{HG}_a]}{[\text{H}][\text{G}]}, \quad (2.27\text{a})$$

$$K_{ab}^{\text{mic.}} = \frac{[\text{HG}_2]}{[\text{HG}_a][\text{G}]}, \quad (2.27\text{b})$$

$$K_b^{\text{mic.}} = \frac{[\text{HG}_b]}{[\text{H}][\text{G}]}, \quad (2.27\text{c})$$

$$K_{ba}^{\text{mic.}} = \frac{[\text{HG}_2]}{[\text{HG}_b][\text{G}]}. \quad (2.27\text{d})$$

The total number of microscopic binding constants increases rapidly with the number of binding sites n , but not all of them are independent [69], see Table 2.1.

Table 2.1: Number of microscopic and stepwise binding constants for a host-guest system, where the host molecule contains n binding sites.

	total No. microscopic	No. independent microscopic	No. stepwise
$n = 2$	4	3	2
$n = 3$	12	7	3
$n = 4$	32	15	4
$n = 5$	80	31	5
$n = 6$	192	63	6
general n	$n2^{n-1}$	$2^n - 1$	n

It is often impossible to distinguish between binding at site a or site b experimentally. Therefore, it is convenient to introduce *stepwise binding constants* as illustrated in Scheme 2.4b. The binding constant K_1 accounts for binding at either of the two sites, and K_2 accounts for binding of a second guest molecule at either HG_a or HG_b , formally

$$K_1 = \frac{[\text{HG}_a] + [\text{HG}_b]}{[\text{H}][\text{G}]} = \frac{[\text{HG}]}{[\text{H}][\text{G}]}, \quad (2.28\text{a})$$

$$K_2 = \frac{[\text{HG}_2]}{([\text{HG}_a] + [\text{HG}_b])[\text{G}]} = \frac{[\text{HG}_2]}{[\text{HG}][\text{G}]}, \quad (2.28\text{b})$$

where $[\text{HG}] = [\text{HG}_a] + [\text{HG}_b]$. In a general binding scheme, a stepwise binding constant K_j describes binding of j -th guest to the complex HG_{j-1} according to the definition

$$K_j = \frac{[\text{HG}_j]}{[\text{HG}_{j-1}][\text{G}]}. \quad (2.29)$$

The stepwise binding constants are uniquely determined from the (independent) microscopic binding constants but not vice versa because the number of the latter is always higher than the former, see Table 2.1. Thus, microscopic binding constants cannot be calculated from the stepwise binding constants and are mostly experimentally inaccessible except for symmetric systems, see Sec. 2.2.4 (although they have been determined in an asymmetric two-site system from measurements with many different guests [70]). Stepwise binding constants are sufficient for construction of binding models describing experimental data. However, microscopic binding constants describe the interaction between binding sites during cooperative binding as explained later in Sec. 2.2.4.

For an overall reaction $\mathbf{H} + n\mathbf{G} \rightleftharpoons \mathbf{HG}_n$, the *overall binding constant* is defined as

$$\beta_j = \frac{[\mathbf{HG}_j]}{[\mathbf{H}][\mathbf{G}]^j} = \prod_{j=1}^n K_j, \quad (2.30)$$

where $j = 1 \dots n$ (cf. page 47 in [48]). The overall binding constant is a product of stepwise binding constants. In some cases, one or more complexes \mathbf{HG}_j are present at undetectable concentrations in the sample solution (e.g. for highly cooperative binding steps, see Sec. 2.2.4). Then, the corresponding stepwise binding constants cannot be determined, but the binding can still be described using the product of corresponding stepwise binding constants.

It is impossible to find an analytical solution for concentrations of all species in an overall reaction $\mathbf{H} + n\mathbf{G} \rightleftharpoons \mathbf{HG}_n$ for arbitrary n . Thus, following the approach given in the paper of Hargrove [71], the concentrations of host-related species (i.e. $[\mathbf{H}]$ and $[\mathbf{HG}_j]$) and $[\mathbf{G}]_t$ are expressed as functions of $[\mathbf{H}]_t$ and $[\mathbf{G}]$. Subsequently, the free guest concentration is calculated numerically from a polynomial equation in the form $f([\mathbf{G}]; K_j, [\mathbf{H}]_t, [\mathbf{G}]_t) = 0$. Here, we give an example of host-guest binding with 1:4 stoichiometry. The equations for stepwise binding constants and total mass balance are

$$K_j = \frac{[\mathbf{HG}_j]}{[\mathbf{HG}_{j-1}][\mathbf{G}]}, \text{ for } j = 1, \dots, 4, \quad (2.31a)$$

$$[\mathbf{H}]_t = [\mathbf{H}] + [\mathbf{HG}] + [\mathbf{HG}_2] + [\mathbf{HG}_3] + [\mathbf{HG}_4], \quad (2.31b)$$

$$[\mathbf{G}]_t = [\mathbf{G}] + [\mathbf{HG}] + 2[\mathbf{HG}_2] + 3[\mathbf{HG}_3] + 4[\mathbf{HG}_4]. \quad (2.31c)$$

The concentrations of host-related species have a simple functional dependence on the free guest concentration as calculated from Eq. (2.31),

$$[\mathbf{H}] = \frac{[\mathbf{H}]_t}{1 + K_1[\mathbf{G}] + K_1K_2[\mathbf{G}]^2 + K_1K_2K_3[\mathbf{G}]^3 + K_1K_2K_3K_4[\mathbf{G}]^4}, \quad (2.32a)$$

$$[\mathbf{HG}] = \frac{K_1[\mathbf{G}][\mathbf{H}]_t}{1 + K_1[\mathbf{G}] + K_1K_2[\mathbf{G}]^2 + K_1K_2K_3[\mathbf{G}]^3 + K_1K_2K_3K_4[\mathbf{G}]^4}, \quad (2.32b)$$

$$[\mathbf{HG}_2] = \frac{K_1K_2[\mathbf{G}]^2[\mathbf{H}]_t}{1 + K_1[\mathbf{G}] + K_1K_2[\mathbf{G}]^2 + K_1K_2K_3[\mathbf{G}]^3 + K_1K_2K_3K_4[\mathbf{G}]^4}, \quad (2.32c)$$

$$[\mathbf{HG}_3] = \frac{K_1K_2K_3[\mathbf{G}]^3[\mathbf{H}]_t}{1 + K_1[\mathbf{G}] + K_1K_2[\mathbf{G}]^2 + K_1K_2K_3[\mathbf{G}]^3 + K_1K_2K_3K_4[\mathbf{G}]^4}, \quad (2.32d)$$

$$[\mathbf{HG}_4] = \frac{K_1K_2K_3K_4[\mathbf{G}]^4[\mathbf{H}]_t}{1 + K_1[\mathbf{G}] + K_1K_2[\mathbf{G}]^2 + K_1K_2K_3[\mathbf{G}]^3 + K_1K_2K_3K_4[\mathbf{G}]^4}. \quad (2.32e)$$

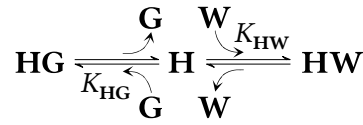
In the next step, Eq. (2.32a–e) are substituted into Eq. (2.31c) and rearranged, yielding the aforementioned equation $f([\mathbf{G}]; K_j, [\mathbf{H}]_t, [\mathbf{G}]_t) = 0$ in the form

$$\begin{aligned} &K_1K_2K_3K_4[\mathbf{G}]^5 + K_1K_2K_3(4K_4[\mathbf{H}]_t - K_4[\mathbf{G}]_t + 1)[\mathbf{G}]^4 + K_1K_2(3K_3[\mathbf{H}]_t - K_3[\mathbf{G}]_t + 1)[\mathbf{G}]^3 \\ &+ K_1(2K_2[\mathbf{H}]_t - K_2[\mathbf{G}]_t + 1)[\mathbf{G}]^2 + (K_1[\mathbf{H}]_t - K_1[\mathbf{G}]_t + 1)[\mathbf{G}] - [\mathbf{G}]_t = 0. \end{aligned} \quad (2.33)$$

Concentration $[G]$ is then calculated numerically from this polynomial equation in the interval $[G] \in [0, [G]_t]$, concentrations of other species are obtained from Eq. (2.32). Implementation and derivation of the binding models for up to four binding sites in *Mathematica* is given in Appendix A.2–A.4.

2.2.3 Competitive host-ligand binding

Competitive binding is a situation where two different ligands bind to a single binding site at the host molecule (and the binding of one ligand prevents binding of another ligand). Assuming the host molecule possesses one binding site (1:1 binding stoichiometry) and denoting the competitive ligands as **G** (guest) and **W** (for example, water), the binding can be described by two binding constants K_{HG} and K_{HW} according to Scheme 2.5 [2, 40].



Scheme 2.5: Competitive 1:1 host-ligand binding scheme with ligands **G** and **W**.

The corresponding equations for binding constants and mass balance are

$$K_{HG} = \frac{[\text{HG}]}{[\text{H}][\text{G}]}, \quad (2.34a)$$

$$K_{HW} = \frac{[\text{HW}]}{[\text{H}][\text{W}]}, \quad (2.34b)$$

$$[\text{H}]_t = [\text{H}] + [\text{HG}] + [\text{HW}], \quad (2.34c)$$

$$[\text{G}]_t = [\text{G}] + [\text{HG}], \quad (2.34d)$$

$$[\text{W}]_t = [\text{W}] + [\text{HW}]. \quad (2.34e)$$

Numeric solution, analogous to Eq. (2.32) and (2.33), requires solving a system of two second-order polynomial equations in $[G]$ and $[W]$, see Appendix A.5.1. Here, we show another solution, which leads to numeric treatment of third-order polynomial equation in $[H]$. First, the following equations

$$[\text{HG}] = \frac{K_{HG}[\text{H}][\text{G}]_t}{1 + K_{HG}[\text{H}]}, \quad (2.35a)$$

$$[\text{HW}] = \frac{K_{HW}[\text{H}][\text{W}]_t}{1 + K_{HW}[\text{H}]}, \quad (2.35b)$$

$$[\text{G}] = \frac{[\text{G}]_t}{1 + K_{HG}[\text{H}]}, \quad (2.35c)$$

$$[\text{W}] = \frac{[\text{W}]_t}{1 + K_{HW}[\text{H}]} \quad (2.35d)$$

are obtained by combination of Eq. (2.34a,b) and Eq. (2.34d,e). Then, substitution of Eq. (2.35a,b) into Eq. (2.34c) yields the aforementioned third-order polynomial equation in $[H]$,

$$[\text{H}]^3 K_{HG} K_{HW} + [\text{H}]^2 (K_{HG} + K_{HW} + [\text{G}]_t K_{HG} K_{HW} - [\text{H}]_t K_{HG} K_{HW} + K_{HG} K_{HW} [\text{W}]_t) + [\text{H}] (1 + [\text{G}]_t K_{HG} - [\text{H}]_t K_{HG} - [\text{H}]_t K_{HW} + K_{HW} [\text{W}]_t) - [\text{H}]_t = 0. \quad (2.36)$$

This equation is solved numerically in the interval $[H] \in [0, [H]_t]$. The resulting value of $[H]$ is then substituted into Eq. (2.35), yielding the concentrations of other chemical species. Implementation of this solution and derivation of the corresponding equations in *Mathematica* is given in Appendix A.5.2.

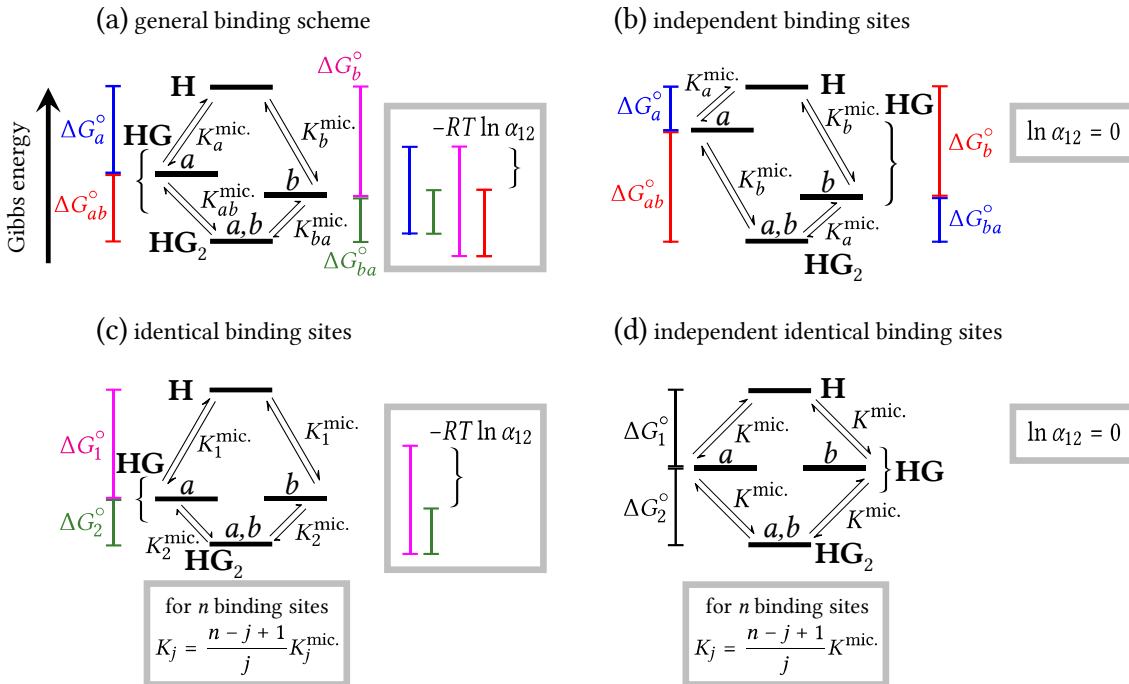
2.2.4 Cooperativity

Cooperative binding relates to host molecules with multiple binding sites. It arises, when the presence of a bound guest increases or decreases affinity to bind another guest (cf. [72] and page 78 in [48]). Let us illustrate this phenomenon with the host-guest binding model, where the host molecule possesses two binding sites a and b . Here, binding is described by four microscopic binding constants (Scheme 2.4a). For example, cooperative binding at the site a arises when binding steps $\mathbf{H} \rightarrow \mathbf{HG}_a$ (in the absence of another bound guest) and $\mathbf{HG}_b \rightarrow \mathbf{HG}_2$ (in the presence of another bound guest) have different affinities (i.e. $K_a^{\text{mic.}} \neq K_{ba}^{\text{mic.}}$). From the point of view of Gibbs energy profile, $\Delta G_a^\circ \neq \Delta G_{ba}^\circ$ (subscript ‘r’ in Δ_r omitted in the following text) as shown in Scheme 2.6a. On the other hand, independence of binding sites implies non-cooperative binding (for binding sites located far apart, e.g. at different ends of a large molecule) and requires $\Delta G_a^\circ = \Delta G_{ba}^\circ$ and $\Delta G_b^\circ = \Delta G_{ab}^\circ$, see Scheme 2.6b.

The four microscopic binding constants are not independent, which can be inferred from the Gibbs energy profile. The energy difference between the \mathbf{HG}_2 complex and the free host and guest species should be equal for both branches $a \rightarrow ab$ and $b \rightarrow ba$, i.e. $\mu_{\mathbf{HG}_2}^\circ - \mu_{\mathbf{H}}^\circ - 2\mu_{\mathbf{G}}^\circ = \Delta G_a^\circ + \Delta G_{ab}^\circ = \Delta G_b^\circ + \Delta G_{ba}^\circ$. Hence, using the formula $\Delta G_j^\circ = -RT \ln K_j$ (Eq. (2.9)) implies $-RT \ln (K_a^{\text{mic.}} K_{ab}^{\text{mic.}}) = -RT \ln (K_b^{\text{mic.}} K_{ba}^{\text{mic.}})$, and thus $K_a^{\text{mic.}} K_{ab}^{\text{mic.}} = K_b^{\text{mic.}} K_{ba}^{\text{mic.}}$. Subsequently, an interaction parameter α_{12} , describing interaction between both binding sites (cf. page 23 in [48]), can be defined as

$$\alpha_{12} = \frac{K_{ba}^{\text{mic.}}}{K_a^{\text{mic.}}} = \frac{K_{ab}^{\text{mic.}}}{K_b^{\text{mic.}}} \quad (2.37)$$

Thus, α_{12} describes cooperativity of binding at both a and b sites, see 2.6a. It is also a measure of standard reaction Gibbs energy difference since $\Delta G_{ab}^\circ - \Delta G_b^\circ = \Delta G_{ba}^\circ - \Delta G_a^\circ = -RT \ln \alpha_{12}$. Interacting sites imply $\alpha_{12} \neq 1$. If $\alpha_{12} > 1$, then binding at one site increases binding affinity for the other site (i.e. $K_{ba}^{\text{mic.}} > K_a^{\text{mic.}}$ and $K_{ab}^{\text{mic.}} > K_b^{\text{mic.}}$). If $\alpha_{12} < 1$, the binding affinity decreases upon



Scheme 2.6: Gibbs energy profile for host-guest 1:2 binding model. (a) General case, (b) independent (i.e. non-interacting) binding sites, (c) identical binding sites, (d) identical and independent binding sites. In (c) and (d), $\Delta G_1^\circ = \Delta G_a^\circ = \Delta G_b^\circ$ and $\Delta G_2^\circ = \Delta G_{ab}^\circ = \Delta G_{ba}^\circ$.

binding of single guest molecule (i.e. $K_{ba}^{\text{mic.}} < K_a^{\text{mic.}}$ and $K_{ab}^{\text{mic.}} < K_b^{\text{mic.}}$). Hence, the cooperativity assumes three qualitatively different modes according to the value of the interaction parameter,

$$\begin{aligned}\alpha_{12} &< 1 \text{ (negative cooperativity),} \\ \alpha_{12} &= 1 \text{ (no cooperativity),} \\ \alpha_{12} &> 1 \text{ (positive cooperativity).}\end{aligned}$$

The case $\alpha_{12} = 1$ must be taken cautiously when assessing independence of the binding sites because a fortuitous compensation of different effects in interacting sites can be manifested as $\alpha_{12} \approx 1$ (cf. [70] and page 86 in [48]).

The general relationship between stepwise and microscopic binding constants derived in [69] is

$$K_j = \frac{\sum_{l_1 < l_2 < \dots < l_j} K_{l_1}^{\text{mic.}} K_{l_1 l_2}^{\text{mic.}} \dots K_{l_1 l_2 \dots l_j}^{\text{mic.}}}{\sum_{l_1 < l_2 < \dots < l_{j-1}} K_{l_1}^{\text{mic.}} K_{l_1 l_2}^{\text{mic.}} \dots K_{l_1 l_2 \dots l_{j-1}}^{\text{mic.}}}, \quad (2.38)$$

where n is the number of binding sites and $l_j = a, b, c, \dots, n$. The binding site labels are ordered as $a < b < c < \dots$, thus, for example, $K_{acb}^{\text{mic.}}$ is not present in the formula as it is dependent on other constants. For $n = 2$, the relationships are $K_1 = K_a^{\text{mic.}} + K_b^{\text{mic.}}$ and $K_2 = \frac{K_a^{\text{mic.}} K_{ab}^{\text{mic.}}}{K_a^{\text{mic.}} + K_b^{\text{mic.}}}$. Note that the dependent $K_{ba}^{\text{mic.}}$ does not appear in these formulae as expected. There are three important special cases for a host with n binding sites.

Case 1 – Independent binding sites as shown in Scheme 2.6b. Here, each binding site has its own binding constant $K_{l_j}^{\text{mic.}}$ independent of occupancy of other binding sites, i.e. $K_{l_1 l_2 \dots l_j}^{\text{mic.}} = K_{l_j}^{\text{mic.}}$ where $l_j = a, b, c, \dots$, and all interaction parameters $\alpha_{j-1, j}$ are equal to unity. Substitution into Eq. (2.38) yields

$$K_j = \frac{\sum_{l_1 < l_2 < \dots < l_j} K_{l_1}^{\text{mic.}} K_{l_2}^{\text{mic.}} \dots K_{l_j}^{\text{mic.}}}{\sum_{l_1 < l_2 < \dots < l_{j-1}} K_{l_1}^{\text{mic.}} K_{l_2}^{\text{mic.}} \dots K_{l_{j-1}}^{\text{mic.}}}, \quad (2.39)$$

where $l_j = a, b, c, \dots$. For $n = 2$, the relationships are $K_1 = K_a^{\text{mic.}} + K_b^{\text{mic.}}$ and $K_2 = \frac{K_a^{\text{mic.}} K_b^{\text{mic.}}}{K_a^{\text{mic.}} + K_b^{\text{mic.}}}$. Systems with independent binding sites are always non-cooperative.

Case 2 – Identical binding sites as shown in Scheme 2.6c. Here, symmetry of the host molecule implies equal energy levels for a particular complex HG_j . Therefore, the microscopic binding constants only depend on the number of bound guests, not on the particular site, hence $K_{l_1 l_2 \dots l_j}^{\text{mic.}} = K_j^{\text{mic.}}$, where $j = 1, 2, \dots, n$. Their relationship with stepwise binding constants can be obtained either by substitution into Eq. (2.38) or by counting the occupied and unoccupied binding sites as follows

$$K_j = \frac{\text{No. unoccupied sites on } \text{HG}_{j-1}}{\text{No. occupied sites on } \text{HG}_j} K_j^{\text{mic.}} = \frac{n - (j - 1)}{j} K_j^{\text{mic.}} = \frac{n - j + 1}{j} K_j^{\text{mic.}}. \quad (2.40)$$

Overview of these relationships for different n is given in Table 2.2.

Systems with identical binding sites might be either cooperative or non-cooperative. Noteworthy, the microscopic binding constants can be directly calculated from the stepwise constants.

Case 3 – Independent identical binding sites as shown in Scheme 2.6d. Here, all binding steps are fully described by only one microscopic binding constant $K^{\text{mic.}}$ (cf. page 52 in [48]). The simplification $K_j^{\text{mic.}} = K^{\text{mic.}}$ in Eq. (2.40) yields

$$K_j = \frac{n - j + 1}{j} K^{\text{mic.}}. \quad (2.41)$$

Table 2.2: Relationship between stepwise and microscopic binding constants for a host-guest system, where the host molecule contains n identical binding sites.

	stepwise and microscopic binding constants					
$n = 2$	$K_1 = 2K_1^{\text{mic.}}$	$K_2 = \frac{1}{2}K_2^{\text{mic.}}$				
$n = 3$	$K_1 = 3K_1^{\text{mic.}}$	$K_2 = K_2^{\text{mic.}}$	$K_3 = \frac{1}{3}K_3^{\text{mic.}}$			
$n = 4$	$K_1 = 4K_1^{\text{mic.}}$	$K_2 = \frac{3}{2}K_2^{\text{mic.}}$	$K_3 = \frac{2}{3}K_3^{\text{mic.}}$	$K_4 = \frac{1}{4}K_4^{\text{mic.}}$		
$n = 5$	$K_1 = 5K_1^{\text{mic.}}$	$K_2 = 2K_2^{\text{mic.}}$	$K_3 = K_3^{\text{mic.}}$	$K_4 = \frac{1}{2}K_4^{\text{mic.}}$	$K_5 = \frac{1}{5}K_5^{\text{mic.}}$	
$n = 6$	$K_1 = 6K_1^{\text{mic.}}$	$K_2 = \frac{5}{2}K_2^{\text{mic.}}$	$K_3 = \frac{4}{3}K_3^{\text{mic.}}$	$K_4 = \frac{3}{4}K_4^{\text{mic.}}$	$K_5 = \frac{2}{5}K_5^{\text{mic.}}$	$K_6 = \frac{1}{6}K_6^{\text{mic.}}$

For $n = 2$, the relationships are $K_1 = 2K^{\text{mic.}}$ and $K_2 = 1/2K^{\text{mic.}}$. These systems are always non-cooperative and the microscopic constant can be determined from the stepwise constants.

The interaction parameter has been defined for the case $n = 2$ as ratio of microscopic binding constants, which can be generalized to arbitrary n provided the binding sites are identical (see Case 2 above). Thus, for binding step $\mathbf{HG}_{j-1} \rightarrow \mathbf{HG}_j$ it holds

$$\alpha_{j-1,j} = \frac{K_j^{\text{mic.}}}{K_{j-1}^{\text{mic.}}} \quad (2.42)$$

Using this definition, $-RT \ln \alpha_{j-1,j}$ can be interpreted as a difference of standard Gibbs energies of reaction (i.e. $-RT \ln \alpha_{j-1,j} = \Delta G_j^\circ - \Delta G_{j-1}^\circ$, where ΔG_j° corresponds to binding step $\mathbf{HG}_{j-1} \rightarrow \mathbf{HG}_j$), see Scheme 2.6c. Using Eq. (2.40) and (2.42), the interaction parameters can be calculated from the stepwise binding constants as (cf. equation (3) in [73])

$$\alpha_{j-1,j} = \frac{(n-j+2)j}{(n-j+1)(j-1)} \frac{K_j}{K_{j-1}} \quad (2.43)$$

Expressions for interaction parameters for different n are listed in Table 2.3. When the binding sites are not identical, binding step $\mathbf{HG}_{j-1} \rightarrow \mathbf{HG}_j$ cannot be described by one interaction parameter only. However, in these cases the microscopic binding constants are usually experimentally inaccessible.

Table 2.3: Interaction coefficients for host-guest binding, where the host molecule contains n identical binding sites.

	interaction parameter					
$n = 2$	$\alpha_{12} = \frac{4K_2}{K_1}$					
$n = 3$	$\alpha_{12} = \frac{3K_2}{K_1}$	$\alpha_{23} = \frac{3K_3}{K_2}$				
$n = 4$	$\alpha_{12} = \frac{8K_2}{3K_1}$	$\alpha_{23} = \frac{9K_3}{4K_2}$	$\alpha_{34} = \frac{8K_4}{3K_3}$			
$n = 5$	$\alpha_{12} = \frac{5K_2}{2K_1}$	$\alpha_{23} = \frac{2K_3}{K_2}$	$\alpha_{34} = \frac{2K_4}{K_3}$	$\alpha_{45} = \frac{5K_5}{2K_4}$		
$n = 6$	$\alpha_{12} = \frac{12K_2}{5K_1}$	$\alpha_{23} = \frac{15K_3}{8K_2}$	$\alpha_{34} = \frac{16K_4}{9K_3}$	$\alpha_{45} = \frac{15K_5}{8K_4}$	$\alpha_{56} = \frac{12K_6}{5K_5}$	

The type of cooperativity described above is called ‘allosteric cooperativity’, which can be either ‘homotropic’ or ‘heterotropic’. The former refers to cooperative binding of one type of guest as in our case, the latter refers to cooperative binding of different types of guest to the same host molecule. There are also other types of cooperative binding for guests, which attach to the host at multiple sites at once or for hydrogen bonding in DNA double helix (cf. pages 175 and 215 in [74]). A well-known example of allosteric cooperative binding is oxygen binding to

haemoglobin at its four binding sites. This system is described by various models in the literature, one of them is based on the formalism described here, which is referred to as Adair-Klotz model [75].

3 Nuclear magnetic resonance spectroscopy

Nuclear magnetic resonance (NMR) is a phenomenon when electromagnetic waves in the MHz range are absorbed in atomic nuclei immersed in external magnetic field and then emitted back at a particular resonant frequency. NMR spectroscopy exploits the NMR related phenomena to obtain useful information about matter at atomic level. In synthetic chemistry, it is applied for structure determination of newly synthesized substances as well as for sample purity checks. In chemical physics, binding and dynamic properties of various systems can be investigated. There are many textbooks dealing with theory of NMR, e.g. Levitt [76], Cavanagh et al. [77], Macomber [78] or Slichter [79]. In this chapter, short introduction to NMR spectroscopy is given as well as detailed treatment of chemical exchange in NMR as described in our paper [2], parts of the text are taken from that publication.

3.1 Basic phenomena in NMR

3.1.1 Nuclei in magnetic field

Atomic nuclei with nonzero spin I also possess the magnetic moment μ , whose magnitude is described by element-specific gyromagnetic ratio γ (ratio of magnetic dipole moment and angular momentum),

$$\mu = \gamma \hbar I. \quad (3.1)$$

In the presence of external magnetic field, energy levels of nuclear spins split into $2I + 1$ levels. This phenomenon is called the *Zeeman effect*. In the following text, let us consider only nuclei with spin $I = 1/2$ resulting in Zeeman splitting into two energy levels. Energy of a magnetic dipole in a magnetic field \mathbf{B}_0 is

$$E_{\text{mag}} = -\mu \cdot \mathbf{B}_0, \quad (3.2)$$

thus, the Zeeman splitting for $I = 1/2$ induces two energy levels separated by

$$\Delta E = |\gamma| B_0 \hbar = |\omega_0| \hbar \quad (3.3)$$

as illustrated in Fig. 3.1. The quantity

$$\omega_0 = -\gamma B_0 \quad (3.4)$$

is called the *Larmor frequency*. For $\gamma > 0$, the energy level of spin $|\uparrow\rangle$ is lower than spin $|\downarrow\rangle$, hence,

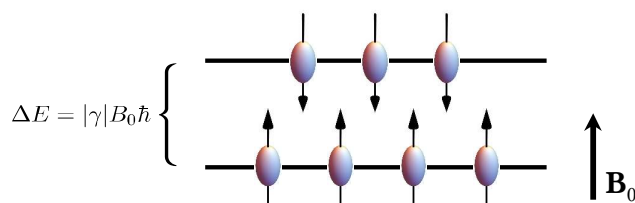


Figure 3.1: Illustration of Zeeman splitting for spin-1/2 nucleus with $\gamma > 0$ (spin and magnetic moments are oriented in the same direction). Populations of energy levels are not to scale.

in equilibrium, state $|\uparrow\rangle$ is more populated according to Boltzmann distribution,

$$\frac{p_{\uparrow}}{p_{\downarrow}} = \exp\left(-\frac{E_{\uparrow} - E_{\downarrow}}{RT}\right). \quad (3.5)$$

At room temperature, the population difference $p_{\uparrow} - p_{\downarrow}$ for hydrogen nucleus ($\gamma = 42.58 \text{ MHz/T}$) is very small, about 10^{-5} . The observed NMR signal is weak since it is proportional to the population difference.

In reality, most nuclear spins are not in their eigenstates $|\uparrow\rangle$ or $|\downarrow\rangle$ but in superposition of these two states (see section 11.1 in [76]). Thus, proper description of the whole ensemble of nuclear spins requires the use of spin density matrix $\hat{\rho}$ with full quantum mechanical approach. However, a classical macroscopic model can be applied for description of the basics of NMR measurement, while some phenomena in NMR require quantum mechanical approach. From macroscopic point of view, the ensemble of nuclear spins can be described by *magnetization*¹ \mathbf{M} , i.e. density of magnetic moments in a volume V ,

$$\mathbf{M} = \frac{1}{V} \sum \boldsymbol{\mu}. \quad (3.6)$$

Let us assume that the external magnetic field points in the z -direction, i.e. $\mathbf{B}_0 = B_0 \mathbf{e}_z$. Then equilibrium implies that the magnetization is also aligned with z -axis (in order to minimize the energy $E_{\text{mag}} = -\mathbf{M} \cdot \mathbf{B}_0$). When initially tilted away from the z -axis, the magnetization starts to precess around the z -axis with the Larmor frequency ω_0 and relaxes back to equilibrium.

3.1.2 Chemical shift

Local magnetic field B_{loc} at the position of reference nucleus is the sum of the external field B_0 and additional field B_{ind} induced by the presence of electrons in the surroundings. The induced field is proportional to the external field, $B_{\text{ind}} = -\sigma B_0$, where σ is the shielding constant. In general, the shielding constant is a tensor, which is averaged to a scalar due to rapid motion in liquid samples. Hence, $B_{\text{loc}} = B_0 + \sigma B_0$ and the nuclear magnetic spins precess with modified Larmor frequency

$$\omega_0 = -\gamma B_{\text{loc}} = -\gamma(1 - \sigma)B_0. \quad (3.7)$$

This influence of surrounding electrons is called *chemical shift* (see section 3.7 in [76]). It enables differentiation of symmetrically nonequivalent nuclear spins in NMR measurement. Thus, NMR provides a useful tool to identify chemical substances, as every molecule has its unique NMR fingerprint.

It is common practice to measure the chemical shift relative to a chemical standard present in the sample with Larmor frequency ω_{std} and shielding constant σ_{std} . Conversion from angular frequency ω to chemical shift δ is defined as [76, 77, 81]

$$\delta = \frac{\omega - \omega_{\text{std}}}{2\pi\nu_0} \times 10^6 = (\sigma - \sigma_{\text{std}}) \times 10^6, \quad (3.8)$$

where ν_0 is the spectrometer frequency and the frequency $2\pi\nu_0$ corresponds to $\sigma = 1$. The unit of chemical shift δ is 'ppm' (parts per million) since the frequency difference $(\omega - \omega_{\text{std}})/(2\pi)$ is in orders of Hz or kHz while ν_0 is in orders of MHz. Chemical shift only depends on shielding constants and does not depend on the external field B_0 because it is calculated as a relative

¹Nuclear magnetization of water in external field $B_0 = 11.7 \text{ T}$ is $M(\text{H}_2\text{O}) = 2c(\text{H}_2\text{O}) N_A \mu (p_{\uparrow} - p_{\downarrow}) = 9.6 \times 10^{-4} \text{ A} \cdot \text{m}^{-1}$, where $c(\text{H}_2\text{O}) = 55.5 \text{ M}$ is the molar water concentration, N_A the Avogadro constant, $\mu = 2.2 \times 10^{-27} \text{ A} \cdot \text{m}^2$ the magnetic moment of hydrogen nucleus and $p_{\uparrow} - p_{\downarrow} = \frac{\hbar\gamma B_0}{2k_B T} = 6.4 \times 10^{-6}$ the population difference at room temperature. This value is much smaller than the magnetization of a typical bar magnet $M(\text{magnet}) = 3.7 \times 10^5 \text{ A} \cdot \text{m}^{-1}$ [80].

frequency difference. Hence, measured chemical shifts of a particular molecule are identical on spectrometers with different B_0 . Note that due to historical reasons, the chemical shift axis in ppm is plotted in increasing order from right to left. Equation (3.8) provides a simple formula for the conversion from ppm to $\text{rad}\cdot\text{s}^{-1}$ valid for frequency difference,

$$\Delta\omega [\text{rad}\cdot\text{s}^{-1}] = 2\pi \times \nu_0 [\text{MHz}] \times \Delta\delta [\text{ppm}]. \quad (3.9)$$

Tetramethylsilan (TMS) is commonly used in nonpolar solvents as a NMR standard because it is inert to common chemicals and its nuclei have rather stable Larmor frequency value (temperature change by 100 °C leads to change of δ_{TMS} by 0.1 ppm [82]). It has twelve hydrogen atoms with identical chemical shift which provides strong NMR signal even at low TMS concentrations.

3.1.3 Relaxation

There are several types of interaction of the nuclear spins with each other and with their surroundings. These interactions are mostly averaged out by molecular motion in liquid or gaseous samples and thus not directly detectable in NMR measurements. Yet they are the cause of relaxation of the magnetization \mathbf{M} from any starting position to the equilibrium $\mathbf{M} \parallel \mathbf{B}_0$. There are two qualitatively different relaxation mechanisms (see sections 2.6 and 2.7 in [76]).

The first mechanism, spin-lattice relaxation, is described by the relaxation time T_1 . Here, energy is transferred from the observed spins to the surroundings (called lattice) to achieve thermodynamic equilibrium with minimum energy $E_{\text{mag}} = -\mathbf{M}\cdot\mathbf{B}_0$ [78]. This type of relaxation involves only the M_z component of the magnetization, hence, it is also called *longitudinal relaxation*.

The other relaxation mechanism with the relaxation time T_2 is called spin-spin relaxation. It is the result of direct dipole-dipole interaction between nuclear magnetic moments. Their interaction causes fluctuations in their local magnetic field B_{loc} and consequently their Larmor frequency, which leads to gradual dephasing during their precession. In this case no energy is dissipated, but entropy of the system increases.

Longitudinal relaxation rate $R_1 = 1/T_1$ and transverse relaxation rate $R_2 = 1/T_2$ are also used instead of relaxation times. For small molecules with low viscosity, such as pure water, at room temperature, typical values of T_1 and T_2 are about 3 – 4 s. However, in solutions of macromolecules or proteins the relaxation times are an order of magnitude or more shorter. In addition, T_2 is always smaller than T_1 .

3.1.4 J-coupling

An interaction, which is not averaged due to fast molecular motion, is indirect spin-spin coupling mediated by electrons in molecular bonds. It is also called J-coupling and its strength is described by the coupling constant J . J-coupling is manifested in NMR by splitting of a single NMR resonance at Larmor frequency ω_0 into two or more resonances (see section 3.8 in [76]). For example, if the observed nuclear spin I_1 is J-coupled to another nuclear spin $I_2 = 1/2$, then two resonances for I_1 are observed. In particular, observed frequency of I_1 is $\omega_0 + J/2$ if I_2 is in state $|\uparrow\rangle$, and $\omega_0 - J/2$ if I_2 is in state $|\downarrow\rangle$. Generally, if spin $I_2 > 1/2$ or the spin I_1 is coupled to several other nuclear spins, the observed Larmor frequency splits into more than two frequencies. J-coupling is a quantum phenomenon and it cannot be described by the classical model (see section 17.8 in [76]).

3.1.5 Bloch equations

Classical derivation of magnetization dynamics and consequently the spectral lineshapes for NMR is based on Bloch equations [83, 84]. Here, single nuclear spin can be represented by the

classical angular momentum \mathbf{L} of some rotating object related to the magnetic moment as $\boldsymbol{\mu} = \gamma\mathbf{L}$. Time change of \mathbf{L} can be induced by torque $\boldsymbol{\tau}$, i.e. $\frac{d\mathbf{L}}{dt} = \boldsymbol{\tau}$. The presence of external field causes torque on the magnetic moment, $\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}$ [85]. Combination of the expressions for torque yields $\frac{d\boldsymbol{\mu}}{dt} = \gamma\boldsymbol{\mu} \times \mathbf{B}$, which is valid for a single magnetic moment. Assuming homogeneous magnetic field in the whole sample, the same expression holds for the magnetization (i.e. volume density of magnetic moments). However, apart from the torque, both longitudinal (T_1) and transverse (T_2) relaxation mechanisms have to be taken into account, hence the resulting *Bloch equations* describing the evolution of magnetization are

$$\begin{aligned}\frac{dM_x}{dt} &= \gamma(\mathbf{M} \times \mathbf{B})_x - \frac{M_x}{T_2}, \\ \frac{dM_y}{dt} &= \gamma(\mathbf{M} \times \mathbf{B})_y - \frac{M_y}{T_2}, \\ \frac{dM_z}{dt} &= \gamma(\mathbf{M} \times \mathbf{B})_z + \frac{M_0 - M_z}{T_1}.\end{aligned}\quad (3.10)$$

Note that the torque $\gamma\mathbf{M} \times \mathbf{B}$ is perpendicular to both the magnetization \mathbf{M} and the magnetic field \mathbf{B} , which causes precession of \mathbf{M} around \mathbf{B} .

3.1.6 Effect of radiofrequency field

The macroscopic magnetization can be manipulated by application of radiofrequency (RF) electromagnetic field of amplitude B_{RF} perpendicular to the external field \mathbf{B}_0 , whose frequency ω_{RF} is close to the Larmor frequency of the nuclei. Because of this resonance, even small B_{RF} has substantial effect on the spin states. It can be shown (see section 8.4.2 in [76]) that during the RF irradiation, the magnetization undergoes nutation movement, which tilts the magnetization away from the equilibrium position ($\mathbf{M} \parallel \mathbf{B}_0$). Solution of the Bloch equations (Eq. 3.10) shows that the nutation frequency is

$$\omega_{\text{nut}} = \left| \frac{1}{2} \gamma B_{\text{RF}} \right|. \quad (3.11)$$

For example, if the RF field is applied during a pulse of duration τ_p , the magnetization rotates by the angle $\beta_p = \omega_{\text{RF}} \tau_p$. In NMR spectroscopy, this is used to manipulate the direction of magnetization using short pulses. For example, in order to tilt \mathbf{M} into the x - y plane, the pulse amplitude B_{RF} and τ_p are chosen such that $\omega_{\text{RF}} \tau_p = \frac{\pi}{2}$. This type of pulse is called ‘ $(\frac{\pi}{2})$ pulse’ (see section 10.8 in [76]), its typical duration is $\tau_p \approx 10 \mu\text{s}$.

3.2 Measurement of NMR

3.2.1 NMR spectrometer

NMR is measured with a NMR spectrometer as shown in Fig. 3.2a. The sample is located in a probe, which is inside a superconducting magnet (cooled with liquid helium and liquid nitrogen) generating the external magnetic field \mathbf{B}_0 , see Fig. 3.2b. RF field is generated by transverse coil wound perpendicular to the sample tube, see Fig. 3.2c. The coil serves as transmitter, inducing the NMR signal, and receiver, detecting the NMR signal. NMR spectroscopy requires high magnetic fields in order to obtain good spectral resolution, for example, magnetic field $B_0 = 11.7 \text{ T}$ induces Larmor frequency of 500 MHz for ^1H nuclei. Note that frequency of the RF field produced by NMR (in orders of 100 MHz) is too high for the current analog-digital converters. Hence, a reference wave with frequency ω_{ref} is subtracted from the NMR signal in so-called ‘quadrature detection’ (see sections 4.5.2 and A.5 in [76]; [87]), and the NMR signal with frequency ω_0 is digitized with the offset frequency $\Omega_0 = \omega_0 - \omega_{\text{ref}}$. Moreover, signal in complex form is obtained from quadrature

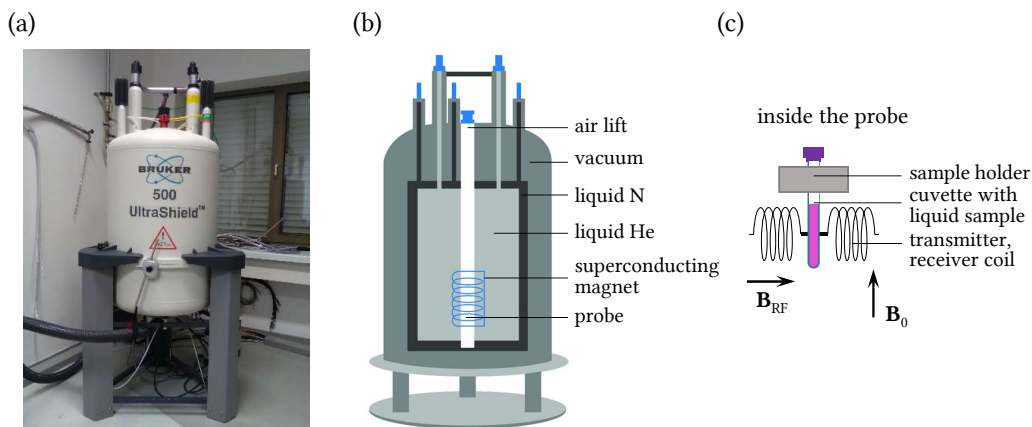


Figure 3.2: NMR spectrometer. (a) Photo of Bruker 500 MHz commercial spectrometer (500 MHz is the Larmor frequency for ^1H nuclei), (b) scheme of the spectrometer parts and (c) detailed view on the sample in the probe. The spectrometer scheme was adopted and edited from [86].

detection. Importantly, ω_{ref} is equal to the frequency of the RF pulse ω_{RF} (see Sec. 3.1.6) since they are produced by the same RF wave synthesizer (see section 4.6 and figure 4.14 in [76]). This fact is conveniently used for theoretical assessment, when a frame of reference rotating with frequency ω_{ref} is introduced (see section 10.6 in [76]).

3.2.2 Pulsed NMR experiment

The production and detection of NMR signal is illustrated in Fig. 3.3. The procedure starts from equilibrium (1), with the magnetization along the z -axis. Then, the magnetization \mathbf{M} is tilted to the x - y plane perpendicular to \mathbf{B}_0 after a ($\frac{\pi}{2}$) RF pulse from the transmitter coil (2). Thereafter, during so-called free induction decay (FID) period, it undergoes free precession with Larmor frequency ω_0 (3) until it relaxes back to equilibrium (4) due to the longitudinal and transverse relaxation mechanisms. Note that $|\mathbf{M}(t)|$ changes during the evolution due to the presence of relaxation.² Also, according to the definition $\omega_0 = -\gamma B_0$, nuclei with $\gamma > 0$ precess clockwise and nuclei with $\gamma < 0$ precess counterclockwise. During the relaxation period (3), the NMR signal is produced in the sample and detected as induced electric current in the receiver coil, which is proportional to the precessing magnetization.

Since the magnetization precesses around the z -axis, only the transverse magnetization component (transverse to the external magnetic field \mathbf{B}_0) is detected. The NMR signal is acquired with quadrature detection (see the previous section) producing complex signal in the form of $M_x(t) + iM_y(t)$, where the original Larmor frequency ω_0 is detected as an offset frequency $\Omega_0 = \omega_0 - \omega_{\text{ref}}$ (where $\omega_{\text{ref}} = \omega_{\text{RF}}$). Usually, frequency domain is preferred over the time domain to analyze NMR signal, for which Fourier transform (FT) is used.³

To compensate for low sensitivity of the NMR technique (mainly due to low $p_{\uparrow} - p_{\downarrow}$), high sample concentrations about 0.1–1 mM are typically used to obtain hydrogen spectrum. Also, the data acquisition described in Fig. 3.3 is repeated many times with subsequent data averaging, typically the number of scans $N_{\text{scan}} < 100$ for hydrogen nuclei and $N_{\text{scan}} = 2000$ – 20000 for carbon nuclei.

²During FID, Bloch equations (Eq. 3.10) have the form $\frac{d}{dt}\mathbf{M} = \omega_0\mathbf{M} \times \mathbf{B} + \mathbf{f}(T_1, T_2)$. Therefore, the magnitude of magnetization $|\mathbf{M}|$ changes in time due to the relaxation terms $\mathbf{f}(T_1, T_2)$ since $\frac{d}{dt}(\mathbf{M} \cdot \mathbf{M}) = 2\mathbf{M} \cdot \frac{d\mathbf{M}}{dt} = \mathbf{M} \cdot \{\omega_0\mathbf{M} \times \mathbf{B} + \mathbf{f}(T_1, T_2)\} = \mathbf{M} \cdot \mathbf{f}(T_1, T_2) \neq 0$.

³Pulsed FT NMR spectroscopy is the state-of-the-art technique and superseded the original continuous-wave (CW) NMR spectroscopy [88].

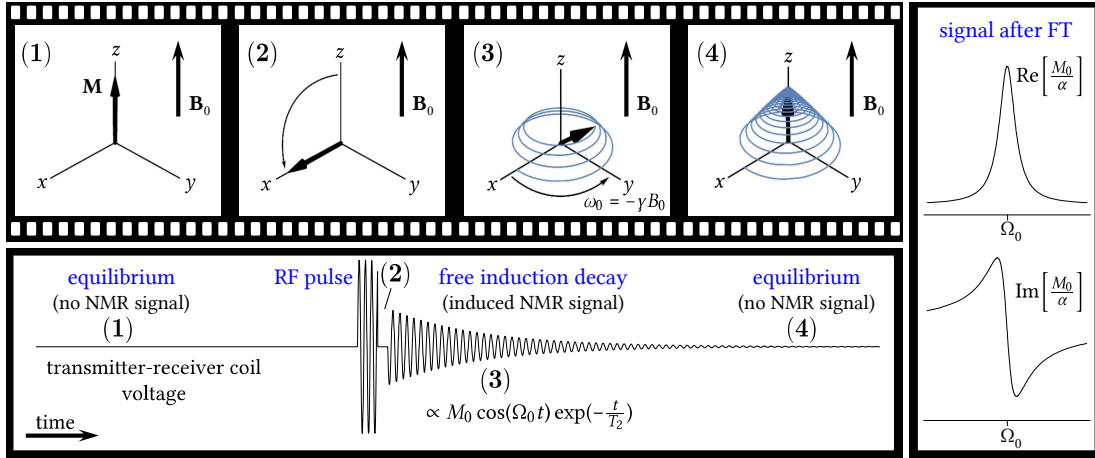


Figure 3.3: Illustration of pulsed NMR experiment. Snapshots (1–4) in the upper left panel correspond to points (1–4) in the bottom left panel showing the voltage in transmitter-receiver coil. The right panel shows Fourier transform of the NMR signal.

3.2.3 Lorentzian spectral lineshape

The NMR signal is detected during the free induction decay (FID) period ((3) in Fig. 3.3), hence it is highly relevant for calculation of the observed spectral lineshapes. During FID, total magnetic field is equal to the constant external field $\mathbf{B} = (0, 0, B_0)$, thus defining the complex transverse magnetization

$$M_{xy} = M_x + iM_y \quad (3.12)$$

Bloch equations (Eq. 3.10) reduce to

(laboratory frame)

$$\frac{dM_{xy}}{dt} = i\omega_0 M_{xy} - R_2 M_{xy}, \quad (3.13a)$$

$$\frac{dM_z}{dt} = R_1(M_0 - M_z). \quad (3.13b)$$

The complex transverse magnetization M_{xy} is proportional to so-called -1 quantum coherence (nondiagonal element of density matrix $\hat{\rho}$) in quantum description of NMR, which is directly proportional to the acquired NMR signal during FID.

Theory of NMR spectroscopy is often constructed in the frame of reference rotating with ω_{ref} . In this frame, magnetization due to spins with Larmor frequency ω_0 rotates with $\Omega_0 = \omega_0 - \omega_{\text{ref}}$. Quadrature detection yields digitized NMR signal with this offset frequency Ω_0 , therefore, the acquired signal is described by Bloch equations in rotating frame,

(rotating frame)

$$\frac{dM_{xy}}{dt} = i\Omega_0 M_{xy} - R_2 M_{xy}, \quad (3.14a)$$

$$\frac{dM_z}{dt} = R_1(M_0 - M_z). \quad (3.14b)$$

In the following text, we will consider spin dynamics only in the rotating frame of reference. Without the loss of generality, let us assume the initial position of magnetization after the $(\frac{\pi}{2})$

pulse along the x -axis, i.e. $M_{xy}(0) = M_x(0) = M_0$ (and $M_y = M_z = 0$), where M_0 is the initial equilibrium magnetization. Using this initial condition, the solution of Eq.(3.14) during FID is

$$M_{xy}(t) = M_0 \exp(i\Omega_0 t - R_2 t), \quad (3.15a)$$

$$M_z(t) = M_0 [1 - \exp(-R_1 t)]. \quad (3.15b)$$

Note that $M_{xy} \rightarrow 0$ and $M_z \rightarrow M_0$ as $t \rightarrow \infty$.

In order to calculate the spectral lineshape in frequency domain $S(\Omega)$, we apply Fourier transform (FT)⁴ on the above solutions for $M_{xy}(t)$. As the signal is collected from $t = 0$, the corresponding magnetization is multiplied by the Heaviside step function $\theta(t)$ or equivalently the integration proceeds in time interval $(0, +\infty)$ instead of $(-\infty, +\infty)$,

$$S(\Omega) = \text{FT}[M_{xy}\theta(t)] = \int_{-\infty}^{+\infty} M_{xy}\theta(t) \exp(-i\Omega t) dt = \int_0^{+\infty} M_{xy} \exp(-i\Omega t) dt. \quad (3.16)$$

For compact notation, we also introduce a quantity

$$\alpha = R_2 - i(\Omega - \Omega_0). \quad (3.17)$$

Then, the solution of Bloch equations for transverse magnetization during FID, i.e. the spectral lineshape, is

$$S(\Omega) = \frac{M_0}{R_2 - i(\Omega - \Omega_0)} = \frac{M_0}{\alpha}. \quad (3.18)$$

Real and imaginary parts of $S(\Omega)$ are called ‘absorption’ and ‘dispersion’ lineshapes, respectively,⁵

$$\text{Re}[S(\Omega)] = \text{Re} \left[\frac{M_0}{\alpha} \right] = \frac{M_0 R_2}{R_2^2 + (\Omega - \Omega_0)^2}, \quad (3.19a)$$

$$\text{Im}[S(\Omega)] = \text{Im} \left[\frac{M_0}{\alpha} \right] = \frac{M_0(\Omega - \Omega_0)}{R_2^2 + (\Omega - \Omega_0)^2}. \quad (3.19b)$$

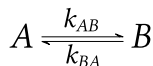
These lineshapes are shown in Fig. 3.3 (right panel). Full width at half maximum of the absorption lineshape is proportional to R_2 , i.e. transverse relaxation causes so-called ‘homogeneous broadening’ of the NMR lineshapes. However, further ‘inhomogeneous broadening’ is caused by inhomogeneity of the external magnetic field B_0 . The sign of the dispersion lineshape depends on the convention in FT. More details about the Lorentzian lineshape are given in Appendix B.1.1.

3.3 Chemical exchange

NMR spectroscopy is sensitive to molecular processes, which change chemical environments, i.e. states A, B, \dots , of a particular reference nuclear spin. Such processes, when they are reversible, are called *chemical exchange* in NMR terminology and lead to non-Lorentzian spectral lineshapes. We denote these transitions between states as spin kinetics. Chemical exchange is characterized by transition rate coefficients k_{ij} (in units of s^{-1}), where $i, j = A, B, \dots; i \neq j$. These coefficients are often denoted as ‘rate constants’, although, due to their possible dependence on concentration (for example, in a bimolecular reaction, see discussion in Sec. 3.4.1), we apply the term ‘transition rate coefficients’. The simplest and most common case is a two-state chemical exchange between A and B spin states (i.e. between states with different chemical environments and thus different offset Larmor frequencies Ω_A and Ω_B) described by two transition rate coefficients k_{AB}, k_{BA} as illustrated in Scheme 3.1.

⁴The convention for FT used in this work is $\hat{f}(\Omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\Omega t} dt$. The corresponding inverse FT is then $f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\Omega)e^{i\Omega t} d\Omega$.

⁵The terms ‘absorption’ and ‘dispersion’ have no physical meaning and are purely historical [76].



Scheme 3.1: Spin kinetics of two-state exchange.

Basic principles of chemical exchange are illustrated in Fig. 3.4 with the example of symmetric two-state exchange. In that case, transition rate coefficients k are identical in both directions (i.e. $k_{AB} = k_{BA} = k$), see Fig. 3.4a. Lifetime τ in both states is exponentially distributed [89] and $k = 1/\langle\tau\rangle$ plays a role of mean frequency of transitions (inverse of the mean lifetime), see Fig. 3.4b. Figure 3.4c shows that the reference nuclear spin precesses with offset frequency Ω_A while in state A and with Ω_B while in state B (decay of the FID signal due to T_2 relaxation is not visible during such short time period). Chemical exchange causes “mixing” of these Larmor frequencies resulting to nontrivial alteration of the resulting NMR spectrum, see Fig. 3.4d (bottom spectrum). There are three different regimes depending on the magnitude of k compared to the Larmor frequency difference $\Delta\Omega_{AB} = |\Omega_A - \Omega_B|$ as shown in Fig. 3.4e:

1. **Slow exchange regime** occurs when $k \ll \Delta\Omega_{AB}$. The spectral lineshape is very close to two Lorentzians placed at Ω_A and Ω_B . Mixing of the two Larmor frequencies is too slow and does not induce any significant spectral change.
2. **Intermediate exchange regime** occurs when $k \approx \Delta\Omega_{AB}$. The spectrum is broadened and significantly differs from the Lorentzian lineshape. The two peaks merge at a so-called ‘coalescence point’ k_c (when k increases, the minimum between two peaks disappears exactly at k_c).
3. **Fast exchange regime** occurs when $k \gg \Delta\Omega_{AB}$. Fast mixing of Larmor frequencies averages the signal into a single Lorentzian-like peak located at $p_A\Omega_A + p_B\Omega_B$.

Note that the intermediate regime is sometimes called ‘slow intermediate’ or ‘fast intermediate regime’ depending from which side the coalescence point is approached [76]. Furthermore, the area below the spectral lineshape (i.e. peak intensity) is not altered when k changes, see Appendix B.

Transition rate coefficient at the coalescence point k_c for symmetric two-state exchange can be calculated from the analytical lineshape [2, 76]. Assuming $R_2^A = R_2^B = R_2$, k_c can be obtained by the following approximative formula

$$k_c \approx \frac{\Delta\Omega_{AB}}{2\sqrt{2}} \left[1 - 1.138 \frac{R_2}{\Delta\Omega_{AB}} \right]. \quad (3.20)$$

This equation can be used for rapid estimation of k_c from a spectrum measured exactly at the coalescence point (no need for lineshape fitting) [90, 91]. Furthermore, the concept of coalescence can also be generalized for asymmetric two-state exchange (i.e. $k_{AB} \neq k_{BA}$). In contrast to symmetric two-state exchange, where the coalescence transition rate coefficient is a single value k_c , asymmetric exchange exhibits infinitely many pairs of transition rate coefficients $\{k_{AB,c}, k_{BA,c}\}$ at which the spectrum has coalescence lineshapes (provided constant values of Ω_A , Ω_B , and R_2). This generalization as well as the approximative formula for non-zero R_2 in Eq. (3.20) were published in our paper [2].

An actual example of a molecular process inducing symmetric two-state exchange is shown in Fig. 3.4f,g. Rotation of a bulky side group (Fig. 3.4f) causes jumps of a *tert*-butyl reference spin (all nine hydrogen nuclei denoted by blue circle have identical Larmor frequency due to the motional averaging – fast rotation around $\text{CH}_3\text{-C}$ bond and *tert*-butyl-C bond) between state A in the vicinity of bromobenzyl and state B further from the bromobenzyl. The corresponding NMR

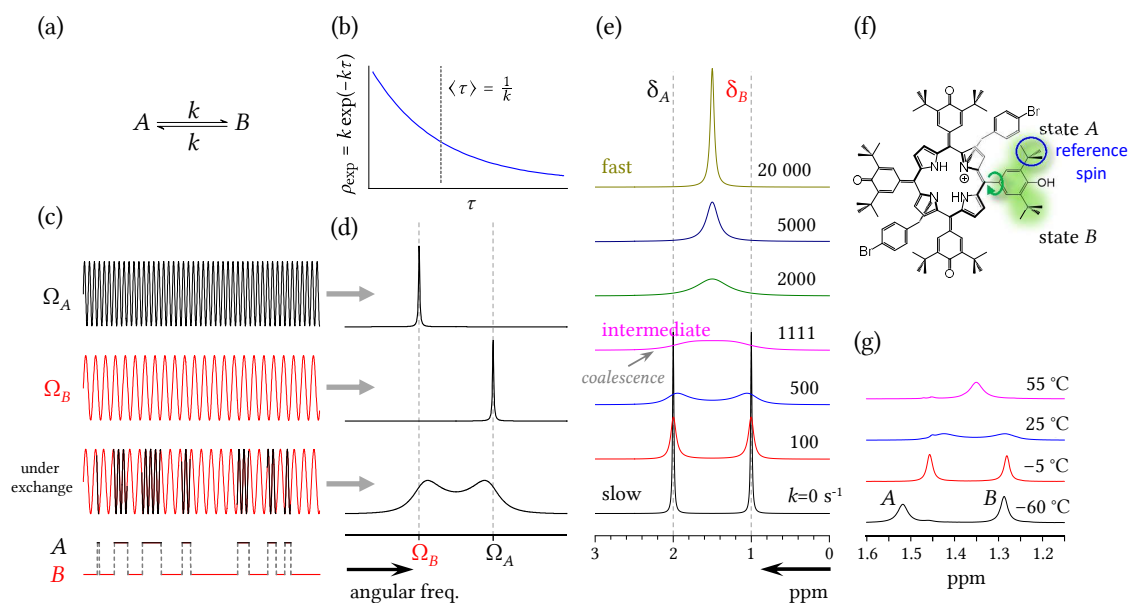


Figure 3.4: Illustration of symmetric two-state exchange with $R_2^A = R_2^B$. (a) Transition in both directions is characterized by the same transition rate coefficient k . (b) Lifetime between transitions has exponential distribution ρ_{exp} , k is the inverse of mean lifetime $\langle \tau \rangle$. (c) Transition between states A and B with different Larmor frequencies causes (d) corresponding spectra. (e) Exchange spectra in slow, intermediate (including the coalescence point at $k = 1111 \text{ s}^{-1}$) and fast regimes. (f) Symmetric two-state exchange in a real system, rotation of a bulky side group provides two different environments for the reference spin (denoted by blue circle). Transition rate coefficient increases with temperature. Note that the Larmor frequencies Ω_A and Ω_B change with temperature.

spectra (Fig. 3.4g) show increase of k with temperature. Note that the offset Larmor frequencies Ω_A and Ω_B change with temperature.

NMR lineshape fitting is a straightforward method to obtain transition rate coefficients and often requires only the acquisition of simple 1D NMR spectra [92–94] as described in Sec. 3.2.2. For more complicated molecules (e.g. proteins), 2D experiments are required. For example, HSQC spectra (see section 7.1 in [77]) can be analyzed, where either a 1D crosssection [95–97] or the full 2D spectrum [98–100] lineshape is fitted. NMR lineshape fitting procedure is most suitable for the analysis of exchange processes with transition rate coefficients approximately in the range of $10\text{--}10^5 \text{ s}^{-1}$.

Methods other than lineshape fitting exist for the determination of transition rate coefficients. However, those methods usually require elaborate NMR sequences and extended experimental time. They are also aimed at the analysis of different ranges of transition rate coefficients (both higher and lower). Reviews of these methods (e.g. ZZ-exchange, EXSY, R_2 relaxation dispersion) are available from Bain [101], Kleckner [102] or Furukawa [103].

3.3.1 Kinetics of spin states

Chemical exchange of a reference spin (i.e. the spin of interest within the structure of molecule) can be modeled with discrete states A, B, C, ... assuming that the transition time between any two states is short compared to the average lifetime in any state. Relative populations of states are denoted as p_A, p_B, p_C, \dots , where $\sum p_j = 1$. For illustration, let us consider two-state spin kinetics, where $[A]$ and $[B]$ are concentrations of spins in states A and B, respectively. Then the time change of $[A]$ is

$$\frac{d[A]}{dt} = -k_{AB}[A] + k_{BA}[B]. \quad (3.21)$$

The first term on the right-hand side accounts for decrease of $[A]$ due to the $A \rightarrow B$ transitions, the second term accounts for its increase due to the $B \rightarrow A$ transitions. The equation for time change of $[B]$ is constructed analogously. Using the definitions $p_A = \frac{[A]}{[A]+[B]}$ and $p_B = \frac{[B]}{[A]+[B]}$, the kinetic equations can be formulated as (assuming that the total concentration $[A] + [B]$ is constant)

$$\frac{d}{dt} \begin{pmatrix} p_A \\ p_B \end{pmatrix} = \mathbf{K} \begin{pmatrix} p_A \\ p_B \end{pmatrix}, \quad (3.22)$$

where

$$\mathbf{K} = \begin{pmatrix} -k_{AB} & k_{BA} \\ k_{AB} & -k_{BA} \end{pmatrix} \quad (3.23)$$

is the kinetic matrix for two states. Note that the sum of left-hand sides of Eq. (3.22) yields zero since $\frac{d}{dt}(p_A + p_B) = \frac{d}{dt}1 = 0$. Hence, the sum of right-hand sides of Eq. (3.22) also yields zero, which implies that all elements of j -th column of the kinetic matrix \mathbf{K} must sum to zero for arbitrary j , i.e. $\sum_i \mathbf{K}_{ij} = 0$. This finding holds for any number of states and can be used to test if \mathbf{K} is correctly constructed.

For a general number of states N , let us define the vector of populations $\mathbf{p} = (p_A, p_B, \dots, p_N)^T$, then the time dependence of populations is governed by

$$\frac{d\mathbf{p}}{dt} = \mathbf{K}\mathbf{p}. \quad (3.24)$$

Note that these first-order differential equations are linear, unlike the governing equations for chemical kinetics (Eq. 2.18).

We are interested in the equilibrium state, which requires the time derivative at the left-hand side of Eq. (3.24) to be zero and converts this differential equation into an algebraic equation

$$\mathbf{K}\mathbf{p} = \mathbf{0}. \quad (3.25)$$

Solution of this equation can be readily obtained by a diagrammatic method introduced in King and Altman [104] or Hill [105].

Apart from equilibrium, this solution also describes the steady-state. To understand the difference between equilibrium and steady-state, let us define a *population flux* from state i to state j as

$$J_{ij} = p_i k_{ij} - p_j k_{ji}. \quad (3.26)$$

It follows then from this definition that $J_{ij} = -J_{ji}$.

The condition for *equilibrium* is that population flux between each state i and j is equal to zero (i.e. $J_{ij} = 0$). In contrast, in *steady-state* mode non-zero fluxes (constant in time) are present, but the net flux \mathcal{J}_j to any state j is zero, so that the populations do not change over time,

$$\mathcal{J}_j = \sum_{\substack{i=A,B,\dots \\ i \neq j}} J_{ij} = 0, \quad (3.27)$$

for all $j = A, B, \dots$. Note that $\sum_{j=A,B,\dots} \mathcal{J}_j = 0$ even outside equilibrium or steady-state because the spin in states A, B, \dots is an isolated system. The condition for zero population flux at equilibrium can be reformulated using populations and transition rate coefficients setting $J_{ij} = 0$ in Eq. (3.26). Hence, at equilibrium (but not at steady-state),

$$\frac{p_j^{\text{eq}}}{p_i^{\text{eq}}} = \frac{k_{ij}}{k_{ji}}. \quad (3.28)$$

As follows from these relationships, the equilibrium conditions reduce the number of independent transition rate coefficients compared to any out-of-equilibrium state. This must be considered during the analysis of actual experimental data. Note that a steady-state spin kinetics can be achieved only in schemes containing a closed cycle [105]. Closed cycle exists in the scheme if it is possible to return to the same state through a different process (i.e. through a different transition state).

3.3.2 Kinetics of two-state exchange

As already mentioned above, two-state spin kinetics is characterized by the kinetic matrix $\mathbf{K} = \begin{pmatrix} -k_{AB} & k_{BA} \\ k_{AB} & -k_{BA} \end{pmatrix}$. Equilibrium populations calculated from $\mathbf{K}\mathbf{p} = 0$ are given in Eq. (3.29) in Table 3.1a. Since closed cycles are not contained in the simple two-state exchange (Table 3.1a), a steady-state mode is not feasible. It also follows from Eq. (3.23) and (3.25) that $J_{AB} = J_{BA} = 0$, which is equivalent to the equilibrium condition in Eq. (3.28). However, even in two-state kinetics, a closed cycle can arise if the transition between states is accomplished by more than one process.

Table 3.1: Two-state spin kinetics with one or two processes.

type	steady-state populations	equilibrium populations
(a) one process		
$A \xrightleftharpoons[k_{BA}]{k_{AB}} B$	no closed cycle, steady-state impossible	$p_A^{\text{eq}} = \frac{k_{BA}}{k_{AB} + k_{BA}} \quad (3.29a)$
		$p_B^{\text{eq}} = \frac{k_{AB}}{k_{AB} + k_{BA}} \quad (3.29b)$
(b) two processes		
	$p_A^{\text{ss}} = \frac{k_{BA}^I + k_{BA}^{II}}{k_{AB}^I + k_{AB}^{II} + k_{BA}^I + k_{BA}^{II}} \quad (3.30)$	$p_A^{\text{eq}} = \frac{k_{BA}^I}{k_{AB}^I + k_{BA}^I} = \frac{k_{BA}^{II}}{k_{AB}^{II} + k_{BA}^{II}} \quad (3.32a)$
	$p_B^{\text{ss}} = \frac{k_{AB}^I + k_{AB}^{II}}{k_{AB}^I + k_{AB}^{II} + k_{BA}^I + k_{BA}^{II}} \quad (3.31)$	$p_B^{\text{eq}} = \frac{k_{AB}^I}{k_{AB}^I + k_{BA}^I} = \frac{k_{AB}^{II}}{k_{AB}^{II} + k_{BA}^{II}} \quad (3.32b)$

To illustrate this case, see the scheme of two-state kinetics with two reversible processes in Table 3.1b. There are two independent reversible processes I (k_{AB}^I and k_{BA}^I) and II (k_{AB}^{II} and k_{BA}^{II}) both of which cause interconversion of A and B states. Since a closed cycle is present, in theory, there can be non-zero population flux in clockwise ($J_{AB}^I = J_{BA}^{II} > 0$), or counterclockwise ($J_{AB}^I = J_{BA}^{II} < 0$) directions. However, the NMR spectral lineshape for a two-process two-state exchange (Table 3.1b) is in principle indistinguishable from one-process two-state exchange (Table 3.1a). Consequently, steady-state and equilibrium modes in the two-process case also cannot be distinguished (both can be mapped on an equilibrium one-process spectral lineshape). There are many examples of two-state exchange with more than one process, some of which enable steady-state, including the conversion of a substrate to a product in reversible enzyme-catalyzed reactions according to the Michaelis-Menten scheme [60], conformational variation in dimesityl systems, where both mesityl rings can flip about the connecting single bonds [91, 106, 107] or other compounds with correlated internal motions [108, 109]. These examples are described in detail in our paper [2].

3.3.3 Kinetics of three-state exchange

Transitions among three states are described by six transition rate coefficients according to the scheme in Table 3.2a with the corresponding kinetic matrix

$$\mathbf{K} = \begin{pmatrix} -k_{AB} - k_{AC} & k_{BA} & k_{CA} \\ k_{AB} & -k_{BA} - k_{BC} & k_{CB} \\ k_{AC} & k_{BC} & -k_{CA} - k_{CB} \end{pmatrix}. \quad (3.33)$$

The equilibrium/steady-state populations can be determined using the equation $\mathbf{K}\mathbf{p} = \mathbf{0}$ with populations $\mathbf{p} = (p_A, p_B, p_C)^\top$, which can also be reformulated using population fluxes (Eq. (3.26)) as

$$J_{BA} + J_{CA} = 0, \quad (3.34a)$$

$$J_{AB} + J_{CB} = 0, \quad (3.34b)$$

$$J_{AC} + J_{BC} = 0. \quad (3.34c)$$

The expressions for populations of the general three-state kinetics and two special cases are summarized in Table 3.2. In contrast to simple two-state kinetics (Table 3.1a), general three-state kinetics (Table 3.2a) contains a cycle and thus enabling the existence of steady-state. However, imposing symmetry (Table 3.2b) or complexity reduction (Table 3.2c) to the general three-state scheme leads only to equilibrium solution and any steady-state solution is absent.

In contrast to simple two-state kinetics (Table 3.1a), general three-state kinetics (Table 3.2a) contains a cycle and thus enabling the existence of steady-state. The expressions for steady-state and equilibrium populations of the general three-state kinetics and two special cases are summarized in Table 3.2. For the special cases, it can be seen that imposing symmetry (Table 3.2b) or complexity reduction (Table 3.2c) to the general three-state scheme leads only to solutions in equilibrium and a steady-state solution is absent (for more details, see text below).

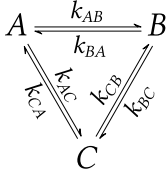
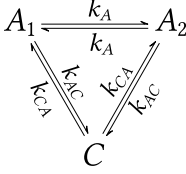
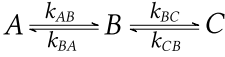
In the case of general three-state kinetics, the populations for steady-state p_j^{ss} are given in Eq. (3.35) in Table 3.2a. At equilibrium, this solution further simplifies when it is combined with the equilibrium conditions ($\frac{p_j^{\text{eq}}}{p_i^{\text{eq}}} = \frac{k_{ij}}{k_{ji}}$, Eq. (3.28)). Since the transition rate coefficients are not mutually independent at equilibrium, we obtain

$$1 = \frac{p_C^{\text{eq}} p_B^{\text{eq}} p_A^{\text{eq}}}{p_B^{\text{eq}} p_A^{\text{eq}} p_C^{\text{eq}}} = \frac{k_{BC} k_{AB} k_{CA}}{k_{CB} k_{BA} k_{AC}}. \quad (3.39)$$

If we choose, for example, k_{BA} as the dependent rate coefficient, then the equilibrium populations can be obtained according to Eq. (3.36abc) in Table 3.2a. The dependent transition rate coefficient k_{BA} is calculated according to Eq. (3.36d). See Appendix B.3 for derivation of the populations in *Mathematica*. It is possible to select any other transition rate coefficient as dependent. It follows from these relationships that the equilibrium condition reduces the number of independent transition rate coefficients, and this fact must be taken into consideration during analysis of any actual experimental data. Prior to use of the three-state lineshape formula derived later in Sec. 3.3.7, the populations and transition rate coefficients for steady-state should be related using Eq. (3.35) and for equilibrium using Eq. (3.36).

The first special case of three-state kinetics is denoted *half-symmetric* and is shown in Table 3.2b. In this case, the states *A* and *B* are denoted as A_1 and A_2 , respectively, due to the symmetry in the corresponding chemical species. Hence, the states A_1 and A_2 are equally populated, and the system is described by three transition rate coefficients: k_{AC} , k_{CA} and k_A . Populations given by Eq. (3.37) already imply equilibrium because they obey the equilibrium conditions in Eq. (3.28). Steady-state with non-zero net fluxes is not possible for this kinetic scheme. The second special case is a *consecutive* kinetic scheme, which is shown in Table 3.2c. As this scheme lacks any closed cycle, only equilibrium is possible with populations expressed in Eq. (3.38).

Table 3.2: Three-state spin kinetics and its special cases.

type	steady-state populations	equilibrium populations
(a) general		
	$p_j^{ss} = \frac{\pi_j}{\pi_A + \pi_B + \pi_C} \quad (3.35)$ $j = A, B, C$ $\pi_A = k_{BA}k_{CA} + k_{BC}k_{CA} + k_{BA}k_{CB}$ $\pi_B = k_{AB}k_{CA} + k_{AB}k_{CB} + k_{AC}k_{CB}$ $\pi_C = k_{AC}k_{BA} + k_{AB}k_{BC} + k_{AC}k_{BC}$	$p_A^{eq} = \frac{k_{BC}k_{CA}}{k_{AC}k_{BC} + k_{AC}k_{CB} + k_{BC}k_{CA}} \quad (3.36a)$ $p_B^{eq} = \frac{k_{AC}k_{CB}}{k_{AC}k_{BC} + k_{AC}k_{CB} + k_{BC}k_{CA}} \quad (3.36b)$ $p_C^{eq} = \frac{k_{AC}k_{BC}}{k_{AC}k_{BC} + k_{AC}k_{CB} + k_{BC}k_{CA}} \quad (3.36c)$ $k_{AB} = \frac{k_{AC}k_{BA}k_{CB}}{k_{BC}k_{CA}} \quad (3.36d)$
(b) half-symmetric		
	<p style="text-align: center;">symmetry prevents steady-state</p>	$p_{A_1}^{eq} = p_{A_2}^{eq} = \frac{k_{CA}}{2k_{CA} + k_{AC}} \quad (3.37a)$ $p_C^{eq} = \frac{k_{AC}}{2k_{CA} + k_{AC}} \quad (3.37b)$
(c) consecutive		
	<p style="text-align: center;">no closed cycle, steady-state impossible</p>	$p_A^{eq} = \frac{k_{BA}k_{CB}}{k_{BA}k_{CB} + k_{AB}k_{BC} + k_{AB}k_{CB}} \quad (3.38a)$ $p_B^{eq} = \frac{k_{AB}k_{CB}}{k_{BA}k_{CB} + k_{AB}k_{BC} + k_{AB}k_{CB}} \quad (3.38b)$ $p_C^{eq} = \frac{k_{AB}k_{BC}}{k_{BA}k_{CB} + k_{AB}k_{BC} + k_{AB}k_{CB}} \quad (3.38c)$

3.3.4 Bloch-McConnell equations

Exchange spin dynamics is described by *Bloch-McConnell equations* [77, 110], which originate from Bloch equations (Eq. (3.10)) and possess additional terms accounting for transfer of magnetization through the exchange process. For two-state chemical exchange the Bloch-McConnell equations are as follows

(laboratory frame)

$$\frac{dM_x^A}{dt} = \gamma(\mathbf{M}^A \times \mathbf{B}^A)_x - R_2^A M_x^A - k_{AB}M_x^A + k_{BA}M_x^B, \quad (3.40a)$$

$$\frac{dM_y^A}{dt} = \gamma(\mathbf{M}^A \times \mathbf{B}^A)_y - R_2^A M_y^A - k_{AB}M_y^A + k_{BA}M_y^B, \quad (3.40b)$$

$$\frac{dM_z^A}{dt} = \gamma(\mathbf{M}^A \times \mathbf{B}^A)_z + R_1^A(M_0^A - M_z^A) - k_{AB}M_z^A + k_{BA}M_z^B, \quad (3.40c)$$

$$\frac{dM_x^B}{dt} = \gamma(\mathbf{M}^B \times \mathbf{B}^B)_x - R_2^B M_x^B - k_{BA}M_x^B + k_{AB}M_x^A, \quad (3.40d)$$

$$\frac{dM_y^B}{dt} = \gamma(\mathbf{M}^B \times \mathbf{B}^B)_y - R_2^B M_y^B - k_{BA}M_y^B + k_{AB}M_y^A, \quad (3.40e)$$

$$\frac{dM_z^B}{dt} = \gamma(\mathbf{M}^B \times \mathbf{B}^B)_z + R_1^B(M_0^B - M_z^B) - k_{BA}M_z^B + k_{AB}M_z^A. \quad (3.40f)$$

Let us use these equations to describe FID in constant external field \mathbf{B}_0 . Using the definition for complex transverse magnetization of nuclei in state $j = A, B$, $M_{xy}^j = M_x^j + iM_y^j$, we obtain

(rotating frame)

$$\frac{dM_{xy}^A}{dt} = i\Omega_A M_{xy}^A - R_2^A M_{xy}^A - k_{AB} M_{xy}^A + k_{BA} M_{xy}^B, \quad (3.41a)$$

$$\frac{dM_{xy}^B}{dt} = i\Omega_B M_{xy}^B - R_2^B M_{xy}^B - k_{BA} M_{xy}^B + k_{AB} M_{xy}^A. \quad (3.41b)$$

Equations (3.41) can be conveniently expressed in matrix form for arbitrary number of states N . Let us define a vector of complex transverse magnetizations $\mathbf{M}_{xy} = (M_{xy}^A, M_{xy}^B, \dots, M_{xy}^N)^\top$. Then the McConnell equations of general N -state exchange (of non- J -coupled spins) for FID signal are

(rotating frame)

$$\frac{d}{dt} \mathbf{M}_{xy}(t) = (i\mathbf{L} - \mathbf{R}_2 + \mathbf{K}) \mathbf{M}_{xy}(t), \quad (3.42)$$

where the diagonal⁶ matrix $\mathbf{L} = \text{diag}(\Omega_A, \Omega_B, \dots, \Omega_N)$ describes evolution due to the external magnetic field, the diagonal matrix $\mathbf{R}_2 = \text{diag}(R_2^A, R_2^B, \dots, R_2^N)$ accounts for transverse relaxation, and the kinetic matrix \mathbf{K} characterizes flux of magnetization from one state to another due to the chemical exchange as discussed in detail in previous sections. Bloch-McConnell equations can be solved analytically or numerically in the frequency domain, which will be explained in the next section. Note that this classical description is not applicable in the presence of J -coupling, which requires quantum mechanical approach [36, 111].

3.3.5 Solution of Bloch-McConnell equations

In literature reports [76, 77, 92, 101] the differential Bloch-McConnell equations (Eq. (3.42)) are usually solved using matrix exponential involving matrix diagonalization. Subsequently, the time-dependent vector of complex transverse magnetizations $\mathbf{M}_{xy}(t)$ is converted to the frequency domain using FT. This method is suitable for symmetric two-state exchange, however, it already becomes quite inconvenient for the case of an asymmetric two-state exchange. Therefore, other studies [112–114] use Bloch-McConnell equations in the frequency domain for numerical simulations of exchange spectra. FT converts the differential Bloch-McConnell equations into algebraic equations, which can be solved either numerically or analytically using standard methods.

The following derivation holds for an arbitrary number of states and arbitrary kinetic schemes. First, FT is applied to Eq. (3.42). Note that the signal from the spectrometer is not actually infinite, but it is equal to zero for $t < 0$. Therefore, the functions $f^j(t) = \theta(t)M_{xy}^j(t)$ should be transformed instead of $M_{xy}^j(t)$, $j = A, B, \dots, N$. The function $\theta(t)$ is the Heaviside step function, equal to zero for $t < 0$ and unity otherwise. The left-hand side of Eq. (3.42) multiplied by the Heaviside function can be rewritten using the Dirac delta function $\delta(t)$,

$$\theta(t) \frac{dM_{xy}^j(t)}{dt} = \frac{d}{dt} [\theta(t)M_{xy}^j(t)] - M_{xy}^j(t) \frac{d\theta(t)}{dt} = \frac{df(t)}{dt} - M_{xy}^j(t)\delta(t). \quad (3.43)$$

FT of the first term on the right-hand side is⁷

$$\widehat{\frac{df(t)}{dt}} = i\Omega \hat{f}(\Omega), \quad (3.44)$$

⁶'diag(\bullet)' represents a diagonal matrix with diagonal elements listed in the parentheses.

⁷The convention for FT used in this work is $\hat{f}(\Omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\Omega t} dt$. The corresponding inverse FT is then $f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\Omega)e^{i\Omega t} d\Omega$.

and FT of the second term is

$$\int_{-\infty}^{+\infty} M_{xy}^j(t) \delta(t) e^{-i\Omega t} dt = M_{xy}^j(t=0). \quad (3.45)$$

Without loss of generality, let us assume that the initial transverse magnetization is real and positive, $M_{xy}^j(t=0) = M_0^j$. Using Eq. (3.43), (3.44) and (3.45), FT of Eq. (3.42) has the following form

$$[i(\Omega \mathbf{1} - \mathbf{L}) + \mathbf{R} - \mathbf{K}] \begin{pmatrix} \hat{f}_A(\Omega) \\ \hat{f}_B(\Omega) \\ \vdots \\ \hat{f}_N(\Omega) \end{pmatrix} = \begin{pmatrix} M_0^A \\ M_0^B \\ \vdots \\ M_0^N \end{pmatrix}, \quad (3.46)$$

where $\mathbf{1}$ is the unit matrix, $\mathbf{L} = \text{diag}(\Omega_A, \Omega_B, \dots, \Omega_N)$, $\mathbf{R} = \text{diag}(R_2^A, R_2^B, \dots, R_2^N)$ and \mathbf{K} is the appropriate kinetic matrix. Using the definition $\alpha_j = R_j + i(\Omega - \Omega_j)$ we can rewrite Eq. (3.46) as

$$[\mathbf{A} - \mathbf{K}] \begin{pmatrix} \hat{f}_A(\Omega) \\ \hat{f}_B(\Omega) \\ \vdots \\ \hat{f}_N(\Omega) \end{pmatrix} = M_0 \begin{pmatrix} p_A \\ p_B \\ \vdots \\ p_N \end{pmatrix}, \quad (3.47)$$

where $\mathbf{A} = \text{diag}(\alpha_A, \alpha_B, \dots, \alpha_N)$ and $M_0^j = M_0 p_j$.

Solution of this set of linear equations for $\hat{f}_j(\Omega)$ is straightforward. Resulting complex NMR signal is detected as the sum of all components,

$$S(\Omega) = \hat{f}_A(\Omega) + \hat{f}_B(\Omega) + \dots + \hat{f}_N(\Omega). \quad (3.48)$$

Solution for two- and three-state exchange is given below. Implementation of this analytical procedure in *Mathematica* code to obtain four- and five-state exchange spectral lineshapes is shown in Appendix B.2. The derivation of the analytical solution also provides a scheme for the numerical simulation of the spectrum [112, 113]. Multiplying Eq. (3.47) by $[\mathbf{A} - \mathbf{K}]^{-1}$ from left yields $\hat{f}_j(\Omega)$ values for a fixed Ω . Hence, for N states, inversion of an $N \times N$ matrix is required for each frequency value. Implementation of the numerical solution in *MATLAB* code is shown in our paper [2].

So-called steady-state approximation (in the time domain) for continuous-wave NMR spectroscopy [115–117] provides an equation fully analogous to Eq. (3.47) without a requirement to apply FT. However, the angular frequency Ω in those derivations stands for the frequency of applied continuous radiofrequency field, not for the frequency of the detected signal (as in pulse NMR approach).

3.3.6 Two-state exchange spectral lineshape

The solution of the system of Bloch-McConnell equations (Eq. (3.42)) for two-state exchange is a standard part of literature reports in the symmetric [76, 77, 101] and also asymmetric case [92, 118, 119]. The formula in complex form reads

$$S_{\text{two-state exch.}}(\Omega) = M_0 \frac{p_A \alpha_B + p_B \alpha_A + k_{AB} + k_{BA}}{\alpha_A \alpha_B + k_{AB} \alpha_B + k_{BA} \alpha_A}, \quad (3.49a)$$

where

$$\alpha_j = R_2^j + i(\Omega - \Omega_j), \quad (3.49b)$$

for $j = A, B$. Total transverse magnetization is denoted as $M_0 = M_0^A + M_0^B$. At equilibrium, the populations and transition rate coefficients in Eq. (3.49) are connected by the algebraic Eq. (3.29).

The expression α_j^{-1} describes the complex Lorentzian lineshape with Larmor frequency Ω_j and transverse relaxation R_2^j , see Sec. 3.2.3. This is useful for calculation of two limit cases explaining the spectral behavior in Fig. 3.4e: (i) The infinitely slow exchange limit gives

$$\lim_{k_{AB}, k_{BA} \rightarrow 0} S_{\text{two-state exch.}}(\Omega) = M_0 \left(\frac{p_A}{\alpha_A} + \frac{p_B}{\alpha_B} \right), \quad (3.50)$$

which is, as expected, sum of two Lorentzian peaks. (ii) The infinitely fast exchange limit should be done in equilibrium regime, i.e. keeping the ratio of transition rate coefficients constant and equal to the ratio of populations, which yields

$$\lim_{k_{AB}, k_{BA} \rightarrow \infty} S_{\text{two-state exch.}}(\Omega) \left[\frac{k_{AB}}{k_{BA}} = \frac{p_B}{p_A} \right] = \frac{M_0}{p_A \alpha_A + p_B \alpha_B}. \quad (3.51)$$

This expression is a single Lorentzian peak located at averaged frequency $p_A \Omega_A + p_B \Omega_B$ with averaged relaxation rate $p_A R_2^A + p_B R_2^B$. Lineshapes with larger number of states behave analogously in slow and fast limit.

For lineshape fitting, it is convenient to express the real (absorption) part of the spectrum from Eq. (3.49). The result and its implementation in *Mathematica* can be found in Appendix B.1.2. The analytical lineshape can also be used to calculate the position of coalescence point in Eq. (3.20), detailed treatment is given in our paper [2].

3.3.7 Three-state exchange spectral lineshape

The analytical solution of the system of Bloch-McConnell equations (Eq. (3.42)) for three-state exchange was, to the best of our knowledge, first published in our paper [2] (although special cases were presented during the early years of NMR spectroscopy [118, 120]). It should be noted that Kovrigin [112] analyzed numerically several properties of three-state (and four-state) chemical exchange. There also exist NMR spectral simulation programs (e.g. *NmrLineGuru* [113] and *TITAN* [98]), which enable the numerical fitting of multi-state exchange lineshapes.

The spectral lineshape for three-state chemical exchange, derived using the procedure in Sec. 3.3.5, is as follows

$$S_{\text{three-state exch.}}(\Omega) = M_0 \frac{\mathcal{P}}{\mathcal{Q}}, \quad (3.52)$$

where

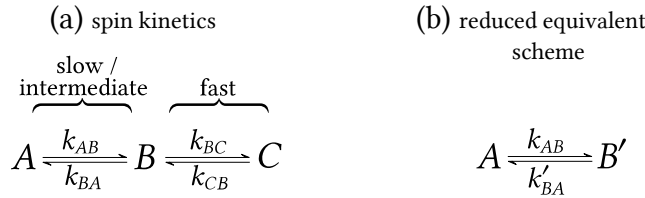
$$\begin{aligned} M_0 &= M_0^A + M_0^B + M_0^C, \\ \mathcal{P} &= p_A[\alpha_B \alpha_C + \alpha_B(k_{CA} + k_{CB} + k_{AC}) + \alpha_C(k_{BA} + k_{BC} + k_{AB})] \\ &\quad + p_B[\alpha_A \alpha_C + \alpha_A(k_{CA} + k_{CB} + k_{BC}) + \alpha_C(k_{AB} + k_{AC} + k_{BA})] \\ &\quad + p_C[\alpha_A \alpha_B + \alpha_A(k_{BA} + k_{BC} + k_{CB}) + \alpha_B(k_{AB} + k_{AC} + k_{CA})] \\ &\quad + \pi_A + \pi_B + \pi_C, \\ \mathcal{Q} &= \alpha_A \alpha_B \alpha_C + \alpha_A \alpha_B(k_{CA} + k_{CB}) + \alpha_A \alpha_C(k_{BA} + k_{BC}) + \alpha_B \alpha_C(k_{AB} + k_{AC}) \\ &\quad + \alpha_A \pi_A + \alpha_B \pi_B + \alpha_C \pi_C, \\ \pi_A &= k_{BA} k_{CA} + k_{BC} k_{CA} + k_{BA} k_{CB}, \\ \pi_B &= k_{AB} k_{CA} + k_{AB} k_{CB} + k_{AC} k_{CB}, \\ \pi_C &= k_{AC} k_{BA} + k_{AB} k_{BC} + k_{AC} k_{BC}. \end{aligned}$$

Quantities π_j are proportional to the steady-state population flux as can be deduced from Eq. (3.35).

Details about lineshape fitting in different modes (i.e. equilibrium, steady-state and out-of-steady-state mode with time-dependent populations) of the spin kinetics and discussion about the interdependency of the fitted parameters during lineshape analysis can be found in our paper [2]. For lineshape fitting, it is convenient to express the real (absorption) part of the spectrum from Eq. (3.52). The result and its implementation in *Mathematica* can be found in Appendix B.1.3. It is also worth mentioning that a steady-state in the general three-state exchange scheme cannot be distinguished from an equilibrium by analyzing a single NMR spectrum since the equilibrium lineshape (Eq. (3.52) and (3.36)) can also be fitted to the steady-state lineshape (Eq. (3.52) and (3.35)), although they are not mathematically identical [2].

3.3.8 Exchange schemes containing a fast process

Let us consider a consecutive exchange with slow or intermediate exchange between states A and B and fast exchange between B and C , see Scheme 3.2a.



Scheme 3.2: Consecutive three-state exchange containing a fast exchange process. (a) Spin kinetics scheme and (b) the corresponding reduced equivalent scheme.

The two fast-exchanging peaks merge, so that only two peaks can be observed. This situation can be modeled as limit of $k_{BC}, k_{CB} \rightarrow \infty$ in Eq. (3.52), where the ratio $k_{BC}/k_{CB} = p_C/p_B$ is kept constant. The resulting equation is

$$S_{\text{three-state exch. limit}}(\Omega) = M_0 \frac{\mathcal{P}}{\mathcal{Q}}, \quad (3.53)$$

where

$$\begin{aligned}
 M_0 &= M_0^A + M_0^B + M_0^C, \\
 \mathcal{P} &= p_A \frac{p_B \alpha_B + p_C \alpha_C}{p_B + p_C} + (p_B + p_C) \alpha_A + \frac{k_{BA} p_B}{p_B + p_C} + k_{AB}, \\
 \mathcal{Q} &= \alpha_A \frac{p_B \alpha_B + p_C \alpha_C}{p_B + p_C} + \frac{k_{BA} p_B}{p_B + p_C} \alpha_A + k_{AB} \frac{p_B \alpha_B + p_C \alpha_C}{p_B + p_C}.
 \end{aligned}$$

In comparison with Eq. (3.49), it becomes obvious that three-state exchange containing a fast process can be modeled as a two-state exchange between states A and B' with transition rate coefficients k_{AB} and $k_{B'A}$, where $k_{B'A}$ has the following relationship to the transition rate coefficient k_{BA} with physical meaning

$$k_{B'A} = k_{BA} \frac{p_B}{p_B + p_C}. \quad (3.54)$$

The offset resonant frequency of the state B' is the population weighted average of the frequencies Ω_B and Ω_C , in particular $\Omega_{B'} = (p_B \Omega_B + p_C \Omega_C)/(p_B + p_C)$ (by analogy with fast asymmetric two-state exchange). Similarly, it holds that $R_2^{B'} = (p_B R_2^B + p_C R_2^C)/(p_B + p_C)$ for the relaxation rate. We can view the B' replacement state as being merged B and C since $M_{B'} = M_B + M_C$ (and consequently $p_{B'} = p_B + p_C$). The full kinetic scheme is simplified to a *reduced equivalent scheme* as seen in Scheme 3.2b. To maintain the transition rate from B' to A (one-way population flux, i.e.

$k_{B'A}p_{B'}$) the same as the transition rate from B to A ($k_{BA}p_B$) in the original scheme, the equality $k_{B'A}p_{B'} = k_{BA}p_B$ must hold. Because $p_{B'}$ is larger than p_B , the modified constant $k_{B'A}$ should be smaller than the original constant k_{BA} . Note that only the transition rate coefficient *out* of the replacement state B' is modified.

Modified transition rate coefficient $k_{B'A}$ can also be obtained directly from the differential equations $\frac{dp}{dt} = \mathbf{Kp}$ without the use of analytical lineshapes. In this case, the equations are as follows

$$\frac{dp_A}{dt} = -k_{AB}p_A + k_{BA}p_B, \quad (3.55a)$$

$$\frac{dp_B}{dt} = -(k_{BA} + k_{BC})p_B + k_{AB}p_A + k_{CB}p_C, \quad (3.55b)$$

$$\frac{dp_C}{dt} = -k_{CB}p_C + k_{BC}p_B. \quad (3.55c)$$

After adding Eq. (3.55b) and Eq. (3.55c), expanding $p_B = \frac{p_B}{p_B + p_C}(p_B + p_C)$ and substituting $p_B + p_C = p_{B'}$ we obtain

$$\frac{dp_A}{dt} = -k_{AB}p_A + k_{BA} \frac{p_B}{p_B + p_C} p_{B'}, \quad (3.56a)$$

$$\frac{dp_{B'}}{dt} = -k_{BA} \frac{p_B}{p_B + p_C} p_{B'} + k_{AB}p_A. \quad (3.56b)$$

Prefactors of $p_{B'}$ in the above equations yield the correct form of $k_{B'A}$ (cf. Eq. (3.54)).

3.4 NMR of host-guest systems

3.4.1 Chemical exchange in host-guest complexes

In this section, we provide the relationship between *reaction rate coefficients* κ_{ij} (introduced in Sec. 2.1.3), which describe the kinetics of host-guest binding, and *transition rate coefficients* k_{ij} which describe transitions of a particular spin between corresponding states. Reaction rate coefficients are independent of concentration and their temperature dependence is governed by the Eyring equation (Eq. 2.20). In general, transition rate coefficients may or may not be dependent on concentration. The relationship between transition and reaction rate coefficients also automatically reveals the concentration dependence of transition rate coefficients, cf. section 5.6.1 in [77].



Scheme 3.3: Two-state exchange with 1:1 host-guest binding. Schemes for (a) chemical kinetics and (b) corresponding spin kinetics of a nucleus located at the host molecule.

Let us consider the 1:1 binding of a host (\mathbf{H}) with a guest (\mathbf{G}) in Scheme 3.3a with an equilibrium constant

$$K_{\mathbf{HG}} = \frac{[\mathbf{HG}]}{[\mathbf{H}][\mathbf{G}]} = \frac{\kappa_{AB}}{\kappa_{BA}}. \quad (3.57)$$

Guldberg-Waage's law of mass action (Sec. 2.1.3) assumes that the forward and backward rates of a chemical reaction are proportional to the concentrations of reacting molecules,

$$\text{forward reaction rate} = \kappa_{AB}[\mathbf{G}][\mathbf{H}], \quad (3.58a)$$

$$\text{backward reaction rate} = \kappa_{BA}[\mathbf{HG}]. \quad (3.58b)$$

The free host **H** can be assigned to state *A* and the complex **HG** to state *B* ($[A] = [\mathbf{H}]$ and $[B] = [\mathbf{HG}]$). The variation of concentration with time of the spin state *B* associated with the **HG** complex, i.e. the transition rate, is given by Scheme 3.3b and Eq. (3.24) (multiplied by $[\mathbf{H}]_t$ to transform populations to concentrations, subscript 't' denotes total concentration), yielding $\frac{d[B]}{dt} = k_{AB}[A] - k_{BA}[B]$. In the latter equation, the two right-hand side terms are identified as forward and backward transition rates, that is

$$\text{forward transition rate} = k_{AB}[A], \quad (3.59a)$$

$$\text{backward transition rate} = k_{BA}[B]. \quad (3.59b)$$

A comparison of Eq. (3.58) with Eq. (3.59), using $[A] = [\mathbf{H}]$ and $[B] = [\mathbf{HG}]$, implies that k_{AB} depends on free guest concentration unlike k_{BA} ,

$$k_{AB} = \kappa_{AB}[\mathbf{G}], \quad (3.60a)$$

$$k_{BA} = \kappa_{BA}. \quad (3.60b)$$

The free guest concentration $[\mathbf{G}]$ and consequently the value of k_{AB} increases with $[\mathbf{G}]_t$ and decreases with $[\mathbf{H}]_t$. For bimolecular elementary reactions the transition rate coefficients usually depend linearly on concentration, while for unimolecular reaction (i.e. decay of the complex) the transition rate coefficients are independent of concentration.

The above formulae (Eq. (3.60)) imply that simple 1:1 host-guest binding does not allow (in general) for symmetric two-state exchange because during the titration experiments (i.e. measurement of spectra at different guest concentrations when the guest is gradually added into the host solution), the p_B/p_A ratio always increases upon addition of guest molecules. Only two out of the three parameters κ_{AB} , κ_{BA} and $K_{\mathbf{HG}}$ are independent. The dependence of $[\mathbf{H}]$, $[\mathbf{HG}]$ and $[\mathbf{G}]$ on the total host and guest concentrations ($[\mathbf{H}]_t$ and $[\mathbf{G}]_t$) can be expressed in an analytical form for 1:1 host-guest binding, see Sec. 2.2.1.

Figure 3.5 illustrates a simulated titration of a host with a guest, which form a complex according to Scheme 3.3a. Simulated spectra shown in Fig. 3.5a contain two resonances in an intermediate exchange regime. The resonances belong to the host in states *A* (corresponding to **H**) and *B* (corresponding to **HG**). As the total amount of guest ($[\mathbf{G}]_t$) increases from zero to one equivalent (from bottom to the top in Fig. 3.5a), the population of state *B* increases (Fig. 3.5b) as well as the concentration of free and bound guest ($[\mathbf{G}]$ and $[\mathbf{HG}]$ in Fig. 3.5c). Because the exchange is in an intermediate regime, the apparent peak maxima shift towards each other and eventually coalesce (Fig. 3.5d, details about coalescence in asymmetric two-state exchange can be found in our paper [2]). In accordance with Eq. (3.60), the transition rate coefficient k_{AB} depends on guest concentration, while k_{BA} does not. Note that although the function $k_{AB}([\mathbf{G}])$ is linear, the function $k_{AB}([\mathbf{G}]_t)$ is nonlinear and can be calculated using Eq. (3.60a) and Eq. (2.26c).

Another example of a three-state exchange in the competitive host-ligand binding model (two types of ligands), including the lineshapes during a simulated titration experiment can be found in our paper [2].

3.4.2 NMR titration binding isotherms

Titration of the host with the guest is a standard experiment to investigate host-guest supramolecular binding. At the beginning, the sample solution contains only the host molecule and the guest is gradually added from a stock solution (the stock solution may contain the same concentration of host as in the sample to prevent dilution). If the host and guest molecules do not interact, the measured spectrum is just sum of the pure **H** and pure **G** spectra. Their interaction is manifested by appearance of new peaks due to the **HG** complex (slow exchange), by shift of

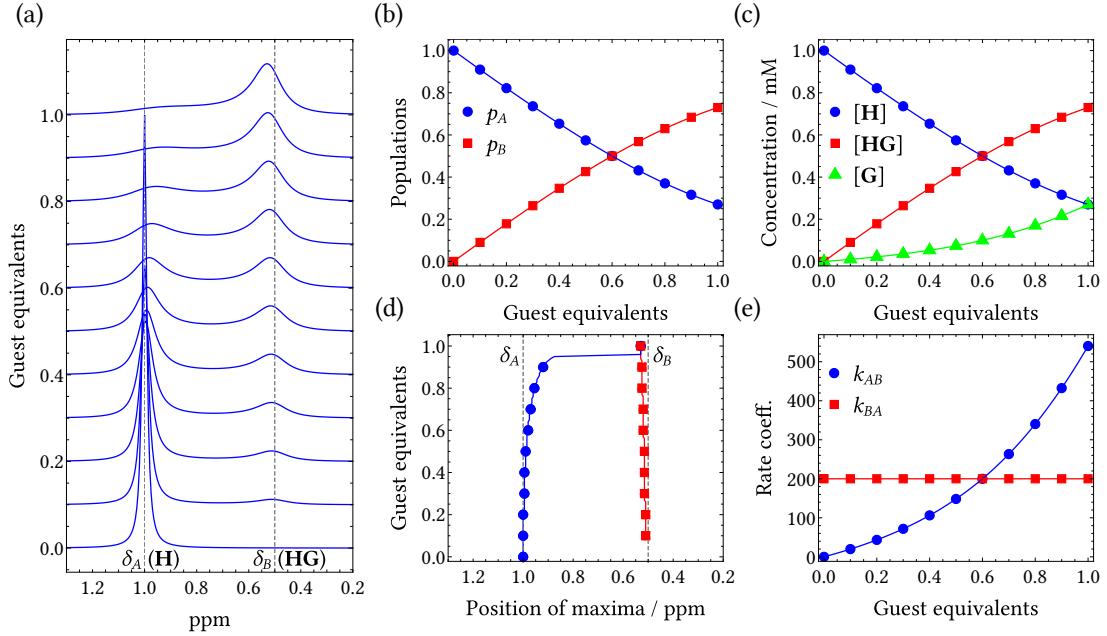


Figure 3.5: Illustration of two-state exchange combined with 1:1 host-guest binding in a simulated titration of host with guest (at constant $[\mathbf{H}]_t$). (a) Simulated spectral lineshapes (y-scaling of spectra is adjusted for clarity). Concentration of guest corresponds to the value where the simulated spectrum meets the y-axis. The spectrum of the free host is cropped for clarity. Concentration dependence of (b) populations, (c) concentration of all species, (d) apparent positions of peak maxima and (e) transition rate coefficients. Discrete points in plots of various quantities correspond to the spectra shown on the left. Simulation parameters: $R_2^A = R_2^B = 30 \text{ s}^{-1}$, $\kappa_{AB} = 2 \times 10^6 \text{ s}^{-1} \text{ M}^{-2}$, $K_{\text{HG}} = 1 \times 10^4 \text{ M}^{-1}$, $[\mathbf{H}]_t = 1 \text{ mM}$, spectrometer frequency $\nu_0 = 500.13 \text{ MHz}$ (which is used for conversion from ppm to $\text{rad} \cdot \text{s}^{-1}$). Chemical shifts of pure A (corresponding to \mathbf{H}) and pure B (\mathbf{HG}) states are $\delta_A = 1 \text{ ppm}$ and $\delta_B = 0.5 \text{ ppm}$, respectively. Equivalents of guest are defined as $[\mathbf{G}]_t/[\mathbf{H}]_t$. The spectral lineshapes were calculated using Eq. (3.49) and concentrations determined using Eq. (2.26).

resonances (fast exchange) or by more complex change of the spectral lineshapes (intermediate exchange regime).

A *binding isotherm* describes dependence of a spectral feature (e.g. chemical shift or integrated peak intensity) on the total concentration of added guest in a titration at constant temperature. The binding isotherm can also be calculated from a binding model, which enables fitting of the experimental data and determination of binding constants. Figure 3.6 illustrates construction of binding isotherms using the example of 1:1 host-guest binding from the previous section. Slow exchange between free and complexed host enables integration of isolated peaks (numerically or using lineshape fitting) due to free and complexed host as shown in Fig. 3.6a. The binding isotherm corresponding to observed complexed fraction of host can be constructed as $p_{\text{HG,obs}} = \frac{\text{area}(A)}{\text{area}(A)+\text{area}(B)}$, see Fig. 3.6b. Complexed fraction can also be calculated from the host-guest binding model using Eq. (2.26b), $p_{\text{HG,calc}} = \frac{[\text{HG}]}{[\mathbf{H}]_t}$. Hence, $p_{\text{HG,obs}}$ obtained from the experiment can be fitted with $p_{\text{HG,calc}}$ with K_{HG} as the fitting parameter. To conclude, slow exchange binding isotherm can be constructed as

$$\text{complexed fraction} = \frac{\sum \text{area}(\text{peaks due to the complex})}{\sum \text{area}(\text{all peaks})}. \quad (3.61)$$

Fast exchange regime causes averaging of observed signals and only one peak is observed (cf. Eq. (3.51)), whose observed chemical shift δ_{obs} is the population-weighted average of chemical

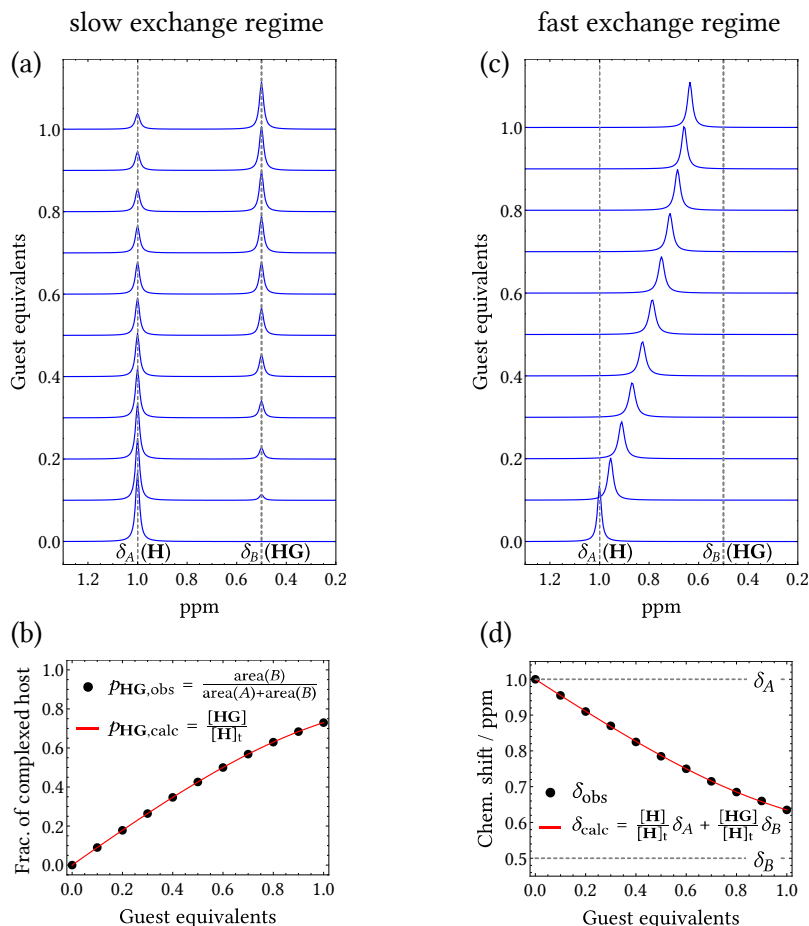


Figure 3.6: Illustration of NMR titration isotherms for two-state exchange combined with 1:1 host-guest binding in simulated titrations of host with guest (at constant $[\mathbf{H}]_t$). (a) Simulated spectral lineshapes in slow exchange regime. (b) Fraction of complexed host and the slow exchange binding isotherm. (c) Simulated spectral lineshapes in fast exchange regime. (d) Position of the observed averaged peak (δ_{obs}) and the fast exchange binding isotherm. Concentration of guest corresponds to the value where the simulated spectrum meets the y -axis in (a,c), y -scaling of the spectra is adjusted for clarity. Discrete points in (b,d) correspond to the spectra shown in (a,c), respectively. Simulation parameters are $\kappa_{AB} = 2 \times 10^4 \text{ s}^{-1}\text{M}^{-2}$ in (a) and $\kappa_{AB} = 2 \times 10^8 \text{ s}^{-1}\text{M}^{-2}$ in (b), the other parameters are identical as in Fig. 3.5. Equivalents of guest are defined as $[\mathbf{G}]_t/[\mathbf{H}]_t$. The spectral lineshapes were calculated using Eq. (3.49) and concentrations determined using Eq. (2.26).

shifts δ_j of all states j , which are under mutual fast exchange. Mathematically

$$\delta_{\text{obs}} = \sum_{j=A,B,\dots} p_j \delta_j. \quad (3.62)$$

Considering our example with two states shown in Fig. 3.6c, the binding isotherm is $\delta_{\text{obs}} = p_A \delta_A + p_B \delta_B$ (see page 194 in [48]). The observed chemical shift can be fitted with $\delta_{\text{calc}} = \frac{[\mathbf{H}]}{[\mathbf{H}]_t} \delta_A + \frac{[\mathbf{HG}]}{[\mathbf{H}]_t} \delta_B$ with K_{HG} as the fitting parameter, where $[\mathbf{H}]$ and $[\mathbf{HG}]$ are calculated using Eq. (2.26), see Fig. 3.6d.

If the NMR resonances are in intermediate regime as in Fig. 3.5a, then the spectral pattern must be fitted with the appropriate exchange lineshape and the aforementioned slow and fast binding isotherms are not applicable.

3.4.3 Determination of enantiomeric excess

Mirror image of a *chiral molecule* is non-superposable with the original (i.e. it lacks symmetry elements of mirror plane, inversion center or improper rotation axis [121]). The two mirror forms are called *enantiomers* and are often denoted as (*R*) and (*S*).⁸ Synthesis of chiral molecules mostly produces a mixture of both enantiomers, which is characterized by *enantiomeric excess* e.e. defined as

$$\text{e.e.} = \frac{[R] - [S]}{[R] + [S]} \quad (3.63)$$

Thus, e.e. = 1 and e.e. = -1 imply enantiopure (*R*)- and (*S*)-form, respectively; e.e. = 0 implies racemate (i.e. 1:1 mixture of (*R*)- and (*S*)-form).

NMR spectra of both enantiomers are identical and thus indistinguishable. Their distinction is enabled by the presence of additional chiral agents in the sample (alternatively, the solvent itself can serve as chiral agent [122]), which form covalent (chiral derivatizing agents) or non-covalent (chiral solvating agents, chiral lanthanide shift reagents) bonds with the chiral analyte [122–124]. Binding of the chiral agent with two enantiomeric forms of chiral analyte produces so-called *diastereomers*, i.e. stereoisomers which are not mutual mirror images⁹ (possible types of diastereomers in the case of 1:1 binding are (*R,R*) and (*R,S*) or alternatively (*S,R*) and (*S,S*)). Diastereomers have different NMR spectra thus enabling determination of e.e. The use of chiral derivatizing agents is limited by kinetic resolution (i.e. different reaction rate for (*R*)- and (*S*)-enantiomers) or by racemization (e.g. by heat) during the analyte-agent reaction [125]. These disadvantages are not present, when non-covalently bound agents are used.

Let us consider non-covalent binding of a chiral solvating agent denoted as host **H** (enantiopure, can be either (*R*)- or (*S*)-form) with enantiomeric mixture of guest analytes (*R*)-**G** and (*S*)-**G** according to the scheme in Fig. 3.7a. Fast exchange between free molecules and their complex is required to produce sharp signals for peak integration. The reference spin of (*R*)-**G** undergoes exchange between states R_{free} and R_{bound} forming a replacement state R' (see Sec. 3.3.8), similarly, reference spin of (*S*)-**G** undergoes exchange between states S_{free} and S_{bound} forming a replacement state S' , see Fig. 3.7b. According to Eq. (3.62) signals of states R' and S' are located at

$$\delta_{R'} = p_{R,\text{free}}\delta_{R,\text{free}} + p_{R,\text{bound}}\delta_{R,\text{bound}} \quad (3.64a)$$

$$\delta_{S'} = p_{S,\text{free}}\delta_{S,\text{free}} + p_{S,\text{bound}}\delta_{S,\text{bound}} \quad (3.64b)$$

Despite fast exchange, two signals are present because they originate from different guest molecules. The splitting $\Delta\delta = \delta_{R'} - \delta_{S'}$ is equal to [123]

$$\Delta\delta = p_{R,\text{free}}\delta_{R,\text{free}} + p_{R,\text{bound}}\delta_{R,\text{bound}} - p_{S,\text{free}}\delta_{S,\text{free}} - p_{S,\text{bound}}\delta_{S,\text{bound}} \quad (3.65)$$

Most importantly, intensity of the peaks is proportional to the concentration of each analyte enantiomer, which enables determination of e.e. by peak integration. Different binding affinity of the agent **H** to (*R*)-**G** or (*S*)-**G** has no effect on resulting areas of the split peaks because of the undergoing exchange between free and bound forms of guest.

Determination of e.e. is illustrated with 2-methylbutyric acid analyte (guest) and 1,2-diphenylethane-1,2-diamine chiral solvating agent (host) [124], see Fig. 3.7c. Signal due to the reference guest methyl group (denoted by blue circle in Fig. 3.7c) forms a doublet due to J-coupling as shown in Fig. 3.7d. Additional splitting by $\Delta\delta$ is attributed to the formation of diastereomers. Then the integrated intensities of peaks due to (*R*)- and (*S*)-form reflect the enantiomeric excess.

⁸Descriptors (*R*)/(*S*) are used, when chirality is caused by a particular asymmetric center in the molecule, mostly an sp^3 carbon atom. Other descriptors exist, see page 304 in [74].

⁹diastereomers <https://goldbook.iupac.org/terms/view/D01679>

Note that $\Delta\delta$ must be large enough so that the two signals are separated, its value differs for each chiral solvating agent [122]. Furthermore, $\Delta\delta$ is the biggest if the chiral solvating agent is enantiopure. This method also works when 1:2 host-guest complexes are formed, however, value of $\Delta\delta$ differs for each stoichiometry [122].

A different method of e.e. determination was introduced recently. It is based on complexation of chiral analytes with so-called prochiral solvating agents [37–40, 42], in particular with oxoporphyrinogens introduced in Sec. 1.2. To understand this method, we will explain a few terms related to so-called ‘topicity’, which is illustrated in Fig. 3.8. Topicity [126] is a property of a pair of sites, in our case the sites are hydrogen atoms. Figure 3.8a illustrates *homotopic*¹⁰ pair of sites, where substitution of the red hydrogen atom for chlorine atom leads to the same result as substitution of the green hydrogen atom. Sites are homotopic when they can interchange through proper rotation C_n . Figure 3.8b illustrates *enantiotopic*¹¹ pair of sites, where substitution of the red or green hydrogen atoms results into two different enantiomers (replacement with a chiral group leads to diastereomers). Enantiotopic molecular sites are related through improper rotation¹² S_n , but not through proper rotation C_n . Note that sites located exactly in the mirror symmetry plane are not enantiotopic. Finally, Fig. 3.8c illustrates *diastereotopic*¹³ pair of sites, where substitution of the red or green hydrogen atoms results into two different diastereomers. Diastereotopic sites are constitutionally equivalent (e.g. bound to the same C atom) but not related by any symmetry operation [126]. Enantiotopic or diastereotopic nuclei are also called ‘heterotopic’. Pair of atomic nuclei located at homotopic or enantiotopic sites possess identical NMR chemical shift unlike diastereotopic sites. Topicity enables us to define *prochiral* molecules as those having at least one pair of enantiotopic sites [126, 127]. Generally, prochirality is “the geometric property of an achiral object (or spatial arrangement of points or atoms) which is capable of becoming chiral in a single desymmetrization step”.¹⁴ The mirror symmetry plane between enantiotopic sites is denoted as *prochiral plane*.

When a prochiral solvating agent (host) forms a complex with a chiral analyte (guest) according to the scheme in Fig. 3.7a, the complexes $\mathbf{H}(R)\text{-G}$ and $\mathbf{H}(S)\text{-G}$ are mutual mirror images as follows from the definition of prochirality. Hence, there is chirality transfer from guest to host [40]. Supramolecular complexes are bound weakly, therefore the guest molecule undergoes fast movement at the host binding site forming so-called ‘chiral field’ [40]. Symmetry of the chiral field depends on symmetry of the host and guest molecules, see Chapter 7 for details. Pair of enantiotopic nuclear spins of host are chosen as reference spins and their NMR spectrum is observed. As shown in Fig. 3.7e, both reference spins are in state C in free host. In the $\mathbf{H}(R)\text{-G}$ complex, the reference spin 1 is in state A and the reference spin 2 is in state B . Since the reference spins are enantiotopic, complex $(S)\text{-G}$ induces state B the reference spin 1 and state A in the reference spin 2. In other words, the presence of prochiral plane in the host shapes the chiral field so that the exchange $(R) \leftrightarrow (S)$ of enantiomers in the complex causes swapping of states $A \rightarrow B$ and $B \rightarrow A$ of the two reference enantiotopic spins. Host-guest complex formation must be fast on the NMR timescale to permit measurement of e.e., thus, only one averaged peak can be observed for each reference spin. According to Eq. (3.62), chemical shifts of the observed averaged peaks are

$$\delta(1) = p_{\mathbf{HG}}(p_{A(1)}\delta_A + p_{B(1)}\delta_B) + p_{\mathbf{H}}\delta_C, \quad (3.66a)$$

$$\delta(2) = p_{\mathbf{HG}}(p_{B(2)}\delta_B + p_{A(2)}\delta_A) + p_{\mathbf{H}}\delta_C, \quad (3.66b)$$

¹⁰homotopic <https://goldbook.iupac.org/terms/view/H02857>

¹¹enantiotopic <https://goldbook.iupac.org/terms/view/E02083>

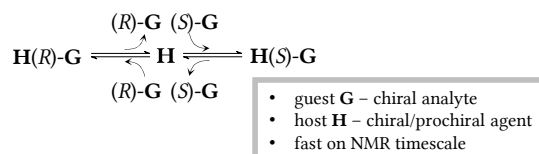
¹²Improper rotation includes inversion i ($S_1 = C_1\sigma_h = i$) and mirror reflection σ_h ($S_2 = C_2\sigma_h = \sigma_h$) [121].

¹³diastereotopic <https://goldbook.iupac.org/terms/view/D01685>

¹⁴prochirality <https://goldbook.iupac.org/terms/view/P04859>

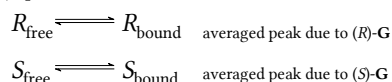
NMR DETERMINATION OF E.E.

(a) general chemical kinetics scheme

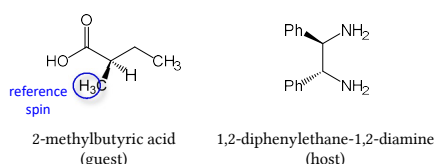


Chiral solvating agent

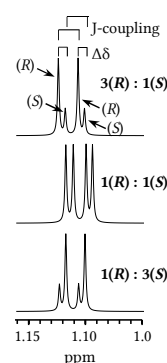
(b) spin kinetics



(c) analyte and chiral solvating agent

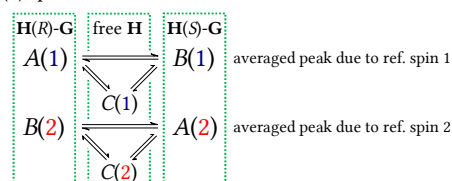


(d) spectral manifestation

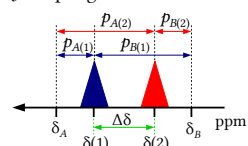


Prochiral solvating agent

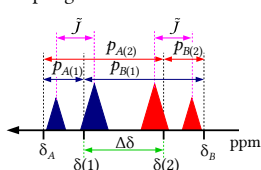
(e) spin kinetics



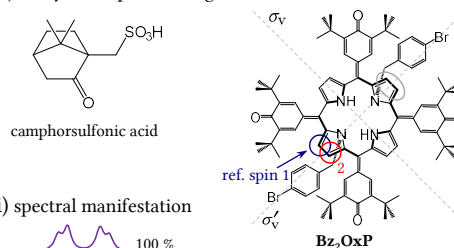
(f) without J-coupling



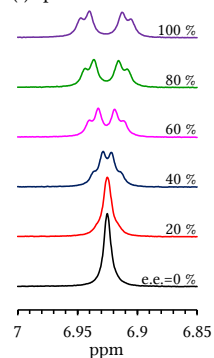
(g) with J-coupling



(h) analyte and prochiral agent



(i) spectral manifestation



(j) linear e.e. splitting

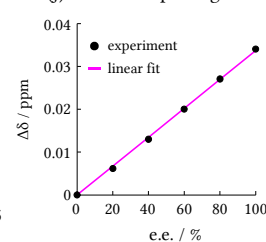


Figure 3.7: Determination of enantiomeric excess (e.e.) using chiral or prochiral solvating agents. (a) Chemical kinetics scheme usable in both methods. The processes must be fast on NMR timescale. **Chiral solvating agent:** (b) Spin kinetics of two reference spins located at two different molecules (R)-G and (S)-G. (c–d) Example of a chiral analyte and a chiral solvating agent adopted from [124]. The reference methyl peak due to G (denoted by blue circle in (c)) forms a doublet due to J-coupling, and further splits by $\Delta\delta$ as the diastereomers are formed. Intensity ratio of peaks due to (R)-G and (S)-G is p_R/p_S . **Prochiral solvating agent:** (e) Spin kinetics of two reference spins located at the same host molecule in HG complex. (f) Determination of $\Delta\delta$ from spectrum without J-coupling and (g) with J-coupling. The former spectral shape consists of two Lorentzians, the latter can be found in [40] or in Appendix B.1.4. Splitting due to J-coupling \bar{J} is related to the coupling constant J by \bar{J} [ppm] = J [Hz]/ ν_0 [MHz], where ν_0 is the spectrometer frequency). Schemes in (e,f) are valid for $p_H = 0$. (h–j) Example of chiral analyte and prochiral solvating agent **Bz₂OxP** adopted from [45]. The enantiotopic reporter groups of prochiral **Bz₂OxP** are its β -protons denoted by color circles in (h) (reference spin 1 and 2). Another pair of identical reporter groups denoted by gray ellipse lies symmetrically relative to the mirror plane σ_v . Spectral splitting in (i) is linear with e.e. according to Eq. (3.67) as shown in (j).

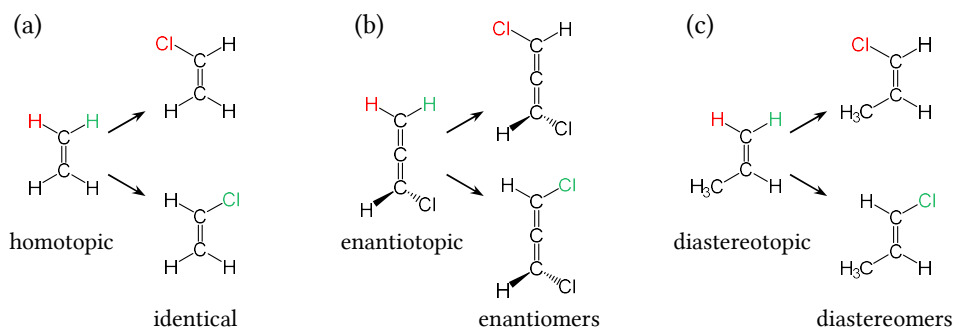


Figure 3.8: Illustration of topicity. The colored pair of hydrogen atoms are (a) homotopic, when replacement of the red hydrogen results to identical structure as replacement of the green hydrogen, (b) enantiotopic, when the replacement leads to enantiomers and (c) diastereotopic, when replacement produces diastereomers.

where $p_{A(1)}$, $p_{B(1)}$, $p_{A(2)}$ and $p_{B(2)}$ are relative fractions of the reference spins 1 and 2 in the corresponding states A or B in the host-guest complex, it holds $p_{A(1)} + p_{B(1)} = 1$ and $p_{A(2)} + p_{B(2)} = 1$. Furthermore, $p_{\text{H}} + p_{\text{HG}} = 1$ holds for populations of free (p_{H}) and complexed (p_{HG}) host. Let us denote relative fractions of (R)- and (S)-form of guest as p_{R} and p_{S} , respectively, where $p_{\text{R}} + p_{\text{S}} = 1$. Using Eq. (3.63) in the form of e.e. = $\frac{p_{\text{R}} - p_{\text{S}}}{p_{\text{R}} + p_{\text{S}}}$ and using the relationships $p_{A(1)} = p_{\text{R}}$, $p_{B(1)} = p_{\text{S}}$, $p_{A(2)} = p_{\text{S}}$ and $p_{B(2)} = p_{\text{R}}$, Eq. (3.66) imply that splitting $\Delta\delta = \delta(1) - \delta(2)$ between signals of reference spins 1 and 2 equals to [40]

$$\Delta\delta = \Delta\delta_{\text{max}} \times \text{e.e.} \times p_{\text{HG}}, \quad (3.67)$$

where $\Delta\delta_{\text{max}} = \delta_A - \delta_B$. Hence, the observed splitting $\Delta\delta$ is directly proportional to e.e. Figures 3.7f,g show how $\Delta\delta$ is determined without or with the presence of J-coupling. These schemes are valid for $p_{\text{HG}} = 0$, otherwise the observed peak positions are both shifted towards δ_C , however, $\Delta\delta$ would not be affected. Note that both averaged peaks have the same intensity. Determination of $\Delta\delta$ for J-coupled nuclei in Fig. 3.7g requires lineshape fitting, the relevant formula for the spectral lineshape is given in Appendix B.1.4.

An actual example is e.e. determination of camphorsulfonic acid analyte using **Bz₂OxP** prochiral solvating agent [45], see Fig. 3.67h. The reporter spins (reference spins 1 and 2) belong to enantiotopic β -hydrogens denoted by blue and red circles (another pair of identical reporter groups denoted by gray ellipse lie symmetrically relative to the mirror plane σ_v , their signals overlap with the signal of the reference spins). These atoms are located symmetrically with respect to prochiral mirror plane σ'_v and are mutually J-coupled, hence the resulting spectral shapes shown in Fig. 3.67i follow the scheme in Fig. 3.67g. Spectral lineshape fitting revealed linear dependence of $\Delta\delta$ on e.e. shown in Fig. 3.67j as expected from Eq. (3.67). Note that C_{2v} symmetry of **Bz₂OxP** causes that each nuclear spin of the molecule has its enantiotopic counterpart reflected through σ_v or σ'_v symmetry plane, which, at the same time, are not related by the C_2 rotation axis. Only the N atoms and the hydrogens of NH groups are an exception as they lie exactly on the symmetry plane.

To conclude, contrary to the use of chiral solvating agents, where e.e. is determined by integration, prochiral solvating agents determine e.e. from the extent of peak splitting $\Delta\delta$. Importantly, sign of e.e. cannot be determined from one measurement since e.e. and $-e.e.$ produce identical peak splitting. The sign can be determined when, for example, small amount of enantiopure (R)-form of analyte is added into the sample, then $\Delta\delta$ increases upon addition if e.e. > 0 or decreases if e.e. < 0 . Also, value of $\Delta\delta_{\text{max}}$ must be obtained independently using enantiopure analyte. Oxoporphyrinogens or TPP (see Sec. 1.2) are well-described prochiral solvating agents, furthermore, a sp^3 meso-carbon porphyrinogen [43] or a coordination complex of Zn with salen-

like ligand [44], acting as prochiral solvating agents, have been published. Determination of e.e. using prochiral solvating agents is also possible in host-guest complexes with 1:2 stoichiometry [37]. Other methods of chiral discrimination in NMR use chiral liquid crystals [128, 129].

4 UV/vis spectroscopy

Ultraviolet/visible (UV/vis) spectroscopy measures interaction of light with electronic structure of analytes. This structure is affected by protonation and supramolecular host-guest binding, thus, UV/vis is a good instrument to investigate these processes typically in titration experiments. Analysis of UV/vis titrations is conveniently performed using singular value decomposition as described in the next chapter.

4.1 Absorbance measurement

The experimental setup inside a UV/vis spectrometer is shown in Fig. 4.1 (see section 13 in [130] for details). The source (incandescent lamp in combination with diffraction grating) produces essentially monochromatic light beam passing through the sample in a cuvette and further to the photodetector. Light interacts with the analyte in three ways: (i) transmission, when the light passes through the sample without changes, (ii) scattering, when the photons change direction and (iii) absorption, when the light induces dissipative processes in the analyte and its energy is transformed to heat. The measurement is performed for all wavelengths in the accessible range, typically 200–900 nm. Incident intensity $I_0(\lambda)$ and transmitted intensity $I(\lambda)$ (Fig. 4.1) are measured by the photodetector (the former is measured without the presence of analyte) for all wavelengths in the accessible range, typically 200–900 nm. Then, the UV/vis spectrum is constructed as transmittance [131]

$$T(\lambda) = \frac{I(\lambda)}{I_0(\lambda)} \quad (4.1)$$

or absorbance

$$A(\lambda) = \log \frac{I_0(\lambda)}{I(\lambda)} = -\log T, \quad (4.2)$$

where ‘log’ denotes decadic logarithm. In this work, we use absorbance for visualization of measured UV/vis spectra. Note that according to Fig. 4.1, the measured absorbance accounts not only for absorption but also for scattering of light in the sample.

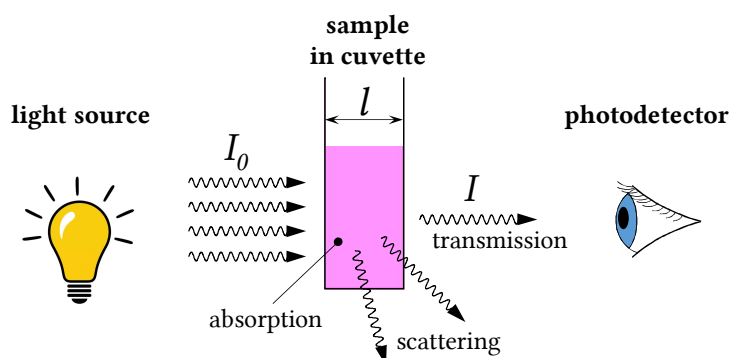


Figure 4.1: Experimental setup of UV/vis measurement.

Absorption of light is directly related to transitions between electronic energy levels of the analyte molecules. In this regard, HOMO-LUMO¹ transitions play the most significant role [132].

¹Highest occupied molecular orbitals and lowest unoccupied molecular orbitals, respectively.

UV/vis spectra in gas state reflect splitting of electronic states due to vibrational and rotational molecular motions [133]. However, in liquid samples due to collisions with solvent, rotation of the analyte is restricted and energies of vibrations are modified in a nonuniform way, thus, rotational and vibrational lines disappear (see chapter 14b in [130] and chapter 7 in [134]). Typical UV/vis absorption bands of liquid samples have approximately log-normal shape [135, 136]. Scattering is caused by the presence of large molecules or particles (≈ 100 nm diameter) [137].

4.2 Beer-Lambert law of absorption

Intensity of light transmitted through the sample decreases exponentially as formulated by the Beer-Lambert law of absorption

$$I = I_0 10^{-\epsilon l[\mathbf{X}]}, \quad (4.3)$$

where l is the cuvette length, $[\mathbf{X}]$ is concentration of the absorbing molecules and ϵ is the molar decadic absorption coefficient (also called extinction coefficient [131]).² The law yields a simple expression for absorbance,

$$A = \epsilon l[\mathbf{X}]. \quad (4.4)$$

This expression is readily generalized for a mixture of N absorbing components \mathbf{X}_i ,

$$\frac{A(\lambda)}{l} = \epsilon_1(\lambda)[\mathbf{X}_1] + \epsilon_2(\lambda)[\mathbf{X}_2] + \dots + \epsilon_N(\lambda)[\mathbf{X}_N]. \quad (4.5)$$

Deviations from linearity in Eq. (4.4) might occur at high analyte concentrations due to several reasons (see pages 306–315 in [130]). For instance, analyte molecules at high concentrations can interact with each other, and thus change their optical properties compared to diluted analyte. Also, absorptivity depends on the refractive index, which might change at high analyte concentration. Another factors, such as the presence of scattering introduce nonlinearity [138]. Furthermore, instrumental deviations such as stray radiation or too wide monochromator window can corrupt the linearity in Eq. (4.4). Absorbance values of $A > 2.5$ or $A < 0.05$ possess high uncertainty due to limited accuracy of the photodetector (table 13-4 in [130]), hence the typical analyte concentrations are 10^{-5} M and typical cuvette width is 1 cm or less.

4.3 UV/vis titration experiment

Titration experiments (i.e. measurement of spectra at different guest concentrations when the guest is gradually added into the host solution, discussed for NMR in Sec. 3.4.2) in UV/vis spectroscopy are a common method to investigate molecular interactions. A model example is titration of **Bz₂OxP** with (*R*)-CSA shown in Fig. 4.2a. Here, host-guest complex with 1:1 stoichiometry is formed, thus, increase of the absorption band of **HG** at 745 nm as well as decrease of the absorption band of **H** at 500 nm are observed. In order to quantify the host-guest binding, binding isotherm is constructed from the absorbance values, which change the most during the titration. In this case it is the absorption maximum at 745 nm. The Beer-Lambert law for mixture of absorbing species in Eq. (4.5) implies the analytical expression for fitting of the binding isotherm in the form (page 148 in [48])

$$A_{\lambda_{\max}} = A_{\mathbf{H}} \frac{[\mathbf{H}]}{[\mathbf{H}]_{\text{t}}} + A_{\mathbf{HG}} \frac{[\mathbf{HG}]}{[\mathbf{H}]_{\text{t}}}, \quad (4.6)$$

²Instead of A or ϵ , absorption cross section $\sigma = \frac{1}{N_{\text{A}}[\mathbf{X}]} \ln \frac{I_0}{I} = \frac{1}{N_{\text{A}}[\mathbf{X}]} \frac{A}{\log e}$ (in m^2) is used, where e is the Euler number and N_{A} is the Avogadro number. It can be expressed using the molar decadic absorption coefficient $\sigma = \frac{\epsilon}{N_{\text{A}} \log e} \approx 3.82 \times 10^{-24} \epsilon \text{ m}^2$ [131].

where $A_j = l\epsilon_j[\mathbf{H}]_t$. Fitting to the experimental data with fitting parameters A_H , A_{HG} and K_1 provides excellent match as shown in Fig. 4.2b. Such analysis exploits only very small part of the total measured data. In the next chapter, we show another method based on singular value decomposition in which the whole measured spectra are analyzed at once.

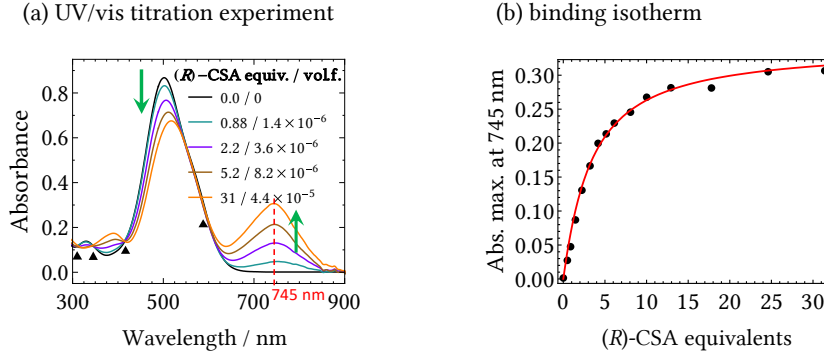


Figure 4.2: Illustration of a titration experiment. (a) Titration of **Bz₂OxP** with (R)-CSA. (b) Binding isotherm constructed from absorbance at 745 nm (red dashed line in (a)), fitted with Eq. (4.6).

4.4 Isosbestic points

When the spectral change involves only two absorbing species, i.e. one converts into the other, *isosbestic points* (points of same extinction) can arise [131]. These are points $\{\lambda_{i.p.}, A(\lambda_{i.p.})\}$, where the absorbance does not change, for example, during a titration experiment, see black triangles in Fig. 4.2a. The necessary condition is that the total concentration of all absorbing species $[\mathbf{X}]_t$ remains constant (or the spectra are rescaled to constant concentration, see footnote 7 on page 58). Let us denote all N_{spc} absorbing species present in the sample as \mathbf{X}_i with total concentration $[\mathbf{X}]_t = [\mathbf{X}_1] + [\mathbf{X}_2] + \dots + [\mathbf{X}_N]$ and suppose interconversion only between \mathbf{X}_1 and \mathbf{X}_2 . Then

$$\begin{aligned} \frac{A(\lambda_{i.p.})}{l} &= \epsilon_1(\lambda_{i.p.})[\mathbf{X}_1] + \epsilon_2(\lambda_{i.p.})[\mathbf{X}_2] + \dots + \epsilon_{N_{\text{spc}}}(\lambda_{i.p.})[\mathbf{X}_{N_{\text{spc}}}] \\ &= \epsilon_1(\lambda_{i.p.}) \{ [\mathbf{X}]_t - [\mathbf{X}_2] - \dots - [\mathbf{X}_{N_{\text{spc}}}] \} + \epsilon_2(\lambda_{i.p.})[\mathbf{X}_2] + \dots + \epsilon_{N_{\text{spc}}}(\lambda_{i.p.})[\mathbf{X}_{N_{\text{spc}}}] \\ &= \underbrace{\epsilon_1(\lambda_{i.p.})[\mathbf{X}]_t}_{\text{const.}} + \underbrace{\{ \epsilon_2(\lambda_{i.p.}) - \epsilon_1(\lambda_{i.p.}) \}}_0 [\mathbf{X}_2] + \dots + \underbrace{\{ \epsilon_{N_{\text{spc}}}(\lambda_{i.p.}) - \epsilon_1(\lambda_{i.p.}) \}}_{\text{const.}} [\mathbf{X}_{N_{\text{spc}}}], \quad (4.7) \end{aligned}$$

and all the right-hand side terms are constant or equal to zero. Therefore, isosbestic points can be formed where the individual spectra of the interconverting absorbing species cross (marked by black triangles in Fig. D.1g in Appendix D). Identification of isosbestic points in the data immediately implies spectral change involving two species.

4.5 Halochromism and solvatochromism

Chromic phenomena are referring to a color change of a dye due to external stimulus. There are several types of chromism including response to light (photochromism), temperature (thermochromism), electricity (electrochromism) and many others [139]. Two chromic phenomena are relevant for supramolecular binding. The first is *halochromism*, chromic response due to protonation (interaction with acid), deprotonation (base), or interaction with ions (salt). The effect of acids is sometimes called acidochromism. Halochromic substances have an immediate application as pH indicators, for example, the Universal Indicator is a mixture of several dyes including phenolphthalein, methyl red and others [140]. Halochromic spectral changes stem from the formation of new absorption bands upon binding (i.e. in the host-guest complex). Both free and bound species are present in the sample and their relative populations determine the resulting color.

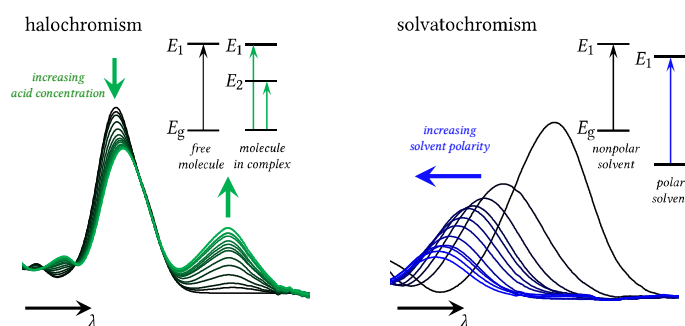


Figure 4.3: Difference between halochromic and solvatochromic spectral changes illustrated with actual UV/vis data. These particular experiments are described and analyzed in Sec. 5.2.

The second chromic event related to supramolecular binding, *solvatochromism*, reflects changes in solvent polarity or weak interactions with solvent molecules (e.g. hydrogen bonding) resulting in shift of the absorption bands, see section 6.2 in [141]. Unlike halochromism, solvatochromism does not produce new absorption bands but produces shift of the absorption maxima, see green and blue arrows in Fig. 4.3, respectively.

Mechanism of solvatochromism consists in different stabilization of ground (E_g) and excited (E_e) energy levels of molecular orbitals [141]. The energy levels of solute are lowered in polar solvent (compared to nonpolar solvent) due to solute-solvent interaction based on permanent and induced electrical dipoles. Polar solvent decreases states with high dipole moment to higher extent than states with low dipole moment. Figure 4.4 illustrates the two types of solvatochromism: (i) positive solvatochromism, where the excited state is stabilized more than the ground state (i.e. $\mu_e > \mu_g$) inducing red (bathochromic) shift with increasing polarity and (ii) negative solvatochromism, where the excited state is stabilized less than the ground state (i.e. $\mu_e < \mu_g$) inducing blue (hypsochromic) shift with increasing polarity. Note that the excited state is in fact Franck-Condon state (E_e^{FC}), i.e. excited electronic states with the position of nuclei and its solvent cage corresponding to the ground state since the optical transitions are much shorter (ca. 10^{-15} s) than vibrational movement of nuclei (ca. 10^{-12} s). A prominent example of a solvatochromic dye is Reichardt's dye [142, 143], used for construction of empirical $E_T(30)$ polarity scale for solvents.

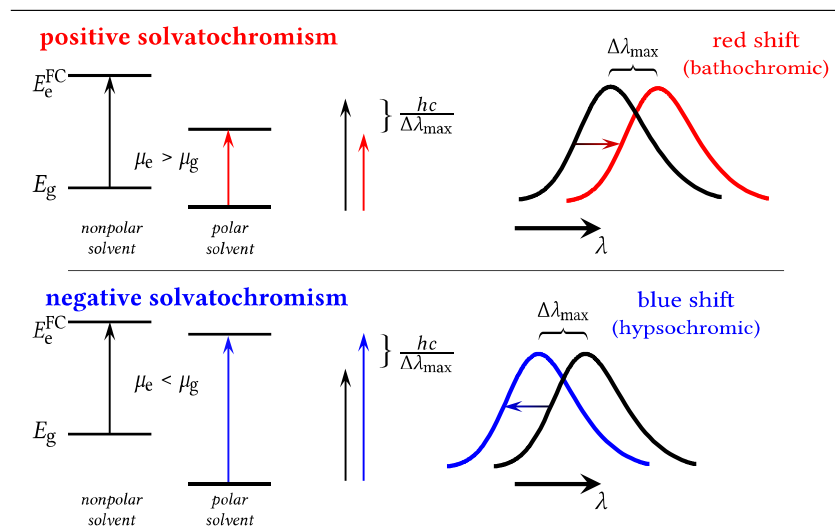


Figure 4.4: Illustration of solvatochromic effects as solvent polarity influence on energy difference between ground state E_g and the Franck-Condon excited state E_e^{FC} .

5 Singular value decomposition

Singular value decomposition (SVD) is a powerful tool for data analysis. In the terms of mathematics, it is a matrix decomposition with good numerical stability and guaranteed existence. It provides a tool for dimensionality reduction, i.e. identification of patterns in high-dimensional data stored in a single matrix and filtering out noise. SVD is closely related to a statistical method called *principal component analysis* (PCA) [144], the terms SVD and PCA are sometimes used interchangeably. There are plenty of SVD- and PCA-based applications, which exploit their power to find patterns in data, for example, in imaging of gene expression in brain [145], face recognition algorithm [146], modal analysis of fluid flows [147, 148], vibration modes of controlled active microparticles [149] and others [150, 151]. Another closely related term is *factor analysis*, which is focused on finding the real factors, which can explain all the variability in the data [152]. Factor analysis stems originally from psychology where it was introduced by Spearman [153] to quantify types of intelligence.

In this chapter, we explain the use of SVD-based methods for analysis of UV/vis spectroscopic data. They serve for determination of the exact number of absorbing species, which can be done without the use of any binding model. The other purpose is determination of the absorption spectra of the individual species, which often cannot be measured separately. This task can be performed either using a particular model of chemical kinetics/binding or even without a model in some cases [154]. Applicability of the SVD-based analysis is limited to a spectroscopic measurement, where the resulting spectrum is linear combination of the individual absorbing components. This includes measurements of UV/vis or IR absorption, Raman scattering [155, 156], circular dichroism and others, but excludes e.g. UV/vis data with solvatochromic shift caused solely by solvent polarity change (see Sec. 5.2.1) and NMR spectra with presence of chemical exchange.

Excellent studying material on both SVD and PCA is provided in book of Brunton and Kutz [157] including series of videlectures.¹ Various applications in chemistry from the perspective of factor analysis are discussed in book of Malinowski [152], comprehensive introduction of SVD in spectroscopy is provided in publications of Henry and Hofrichter [158, 159]. Valuable insights to these methods can be found in publications of Zimányi [160–163].

5.1 Mathematical properties

5.1.1 Basics of SVD

Let us consider a real $m \times n$ matrix \mathbf{A} consisting of experimental data, which has full rank due to the noise. SVD of \mathbf{A} is defined as

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T, \quad (5.1)$$

where \mathbf{U} is a square unitary² $m \times m$ matrix with columns called *left singular vectors*, \mathbf{V} is a square unitary $n \times n$ matrix with columns called *right singular vectors*. The rectangular $m \times n$ matrix \mathbf{S} is diagonal with the non-zero elements termed as *singular values*, the j -th singular value is denoted as s_j . The number of the singular values is $\min(m, n)$ depending on the shape of \mathbf{A} . All singular values are non-negative and follow descending order, i.e. $s_j > s_{j+1}$. For convenience, we denote³

¹Available at <https://www.youtube.com/playlist?list=PLMrJAKhIeNNSVjnsvig1FoY2nXildDCcv>.

²A unitary matrix \mathbf{U} satisfies $\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{1}$. Its columns form therefore an orthonormal basis as well as its rows.

³In this chapter uppercase bold symbols stand for matrices, lowercase bold symbols stand for column vectors.

j -th column of \mathbf{A} as \mathbf{a}_j , j -th column of \mathbf{U} as \mathbf{u}_j and j -th column of \mathbf{V} as \mathbf{v}_j . Vectors \mathbf{a}_j represent single measurements, single images, or single acquired spectra, depending on the type of data. To summarize, left singular vectors \mathbf{u}_j form an orthonormal basis for the single measurements \mathbf{a}_j , right singular vectors multiplied by the corresponding singular value $s_j \mathbf{v}_j$ are coordinates of the basis vectors and s_j themselves indicate, how significant the j -th basis vector is for reconstruction of \mathbf{a}_j .

$$\begin{array}{c}
 \text{full SVD} \qquad \qquad \qquad \text{economy SVD} \\
 \mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{V}^T = \mathbf{U}_n \mathbf{S}_n \mathbf{V}^T \\
 \left\{ \begin{array}{c} \left(\begin{array}{c} \vdots \\ \mathbf{a}_1 \mathbf{a}_2 \dots \mathbf{a}_n \\ \vdots \end{array} \right) = \left(\begin{array}{c} \vdots \\ \mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_m \\ \vdots \end{array} \right) \left(\begin{array}{c} s_1 & & & \\ & s_2 & & \\ & & \ddots & \\ & & & 0 \\ & & & & s_n \\ & & & & & 0 \\ & & & & & & 0 \end{array} \right) \left(\begin{array}{c} \vdots \\ \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_n^T \\ \vdots \end{array} \right) \\
 \underbrace{\hspace{1.5cm}}_n \quad \underbrace{\hspace{1.5cm}}_m \quad \underbrace{\hspace{1.5cm}}_n \quad \underbrace{\hspace{1.5cm}}_n
 \end{array} \right\}^n \\
 \left\{ \begin{array}{c} \left(\begin{array}{c} \vdots \\ \mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_n \\ \vdots \end{array} \right) \left(\begin{array}{c} s_1 & & & \\ & s_2 & & \\ & & \ddots & \\ & & & 0 \\ & & & & s_n \\ & & & & & 0 \end{array} \right) \left(\begin{array}{c} \vdots \\ \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_n^T \\ \vdots \end{array} \right) \\
 \underbrace{\hspace{1.5cm}}_n \quad \underbrace{\hspace{1.5cm}}_n \quad \underbrace{\hspace{1.5cm}}_n
 \end{array} \right\}^n
 \end{array}$$

Figure 5.1: Full and economy SVD of a “tall-skinny” ($m > n$) matrix.

The definition in Eq. (5.1), called *full SVD* is redundant since for a “tall-skinny” matrix ($m > n$), last $m - n$ left singular vectors are multiplied by zeros of \mathbf{S} as illustrated in Fig. 5.1. Let us define truncated matrices $\mathbf{U}_r = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r)$, $\mathbf{V}_r = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r)$ and $\mathbf{S}_r = \text{diag}(s_1, s_2, \dots, s_r)$ (\mathbf{S}_r is a square matrix). Then, the *economy SVD* (equivalent to full SVD) is

$$\mathbf{A} = \mathbf{U}_n \mathbf{S}_n \mathbf{V}^T. \quad (5.2)$$

For a “short-fat” matrix ($m < n$), the economy SVD is analogously

$$\mathbf{A} = \mathbf{U} \mathbf{S}_m \mathbf{V}_m^T. \quad (5.3)$$

Another widely used notation for the SVD decomposition for $p = \min(m, n)$ is

$$\mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{V}^T = s_1 \mathbf{u}_1 \mathbf{v}_1^T + s_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + s_p \mathbf{u}_p \mathbf{v}_p^T, \quad (5.4)$$

where dyadic (tensor) product is used.⁴ This dyadic decomposition is just another notation for the economy SVD. Note that SVD is unique up to the sign in the product $s_j \mathbf{u}_j \mathbf{v}_j^T = s_j (-\mathbf{u}_j) (-\mathbf{v}_j^T)$, however, there exists a method to get rid of this ambiguity [164].

The most useful property of SVD is that the first r terms in Eq. (5.4) are the best possible approximation of the matrix \mathbf{A} than any other matrix of rank r . Using the notation with truncated matrices,

$$\mathbf{A} \approx \mathbf{U}_r \mathbf{S}_r \mathbf{V}_r^T, \quad (5.5)$$

with $r < \min(m, n)$. “Goodness” of SVD approximation is measured with Frobenius norm⁵ $\|\bullet\|$ and it is formulated as the Eckhart-Young theorem [157, 165],

$$\min_{\mathbf{X}, \text{rank}(\mathbf{X})=r} \|\mathbf{A} - \mathbf{X}\| = \|\mathbf{A} - \mathbf{U}_r \mathbf{S}_r \mathbf{V}_r^T\| = s_{r+1}. \quad (5.6)$$

Note the relationship with the singular value s_{j+1} .

SVD can be considered as a method for calculation of the “real” rank of a noise-containing matrix and consequently for separation of real signal from noise. If the rank of a noiseless matrix

⁴Dyadic product of two column vectors \mathbf{a} , \mathbf{b} is defined as $(\mathbf{a}\mathbf{b}^T)_{jk} = a_j b_k$.

⁵Frobenius norm of a real matrix \mathbf{A} is defined as $\|\mathbf{A}\| = \sqrt{\sum_{j=1}^m \sum_{k=1}^n a_{jk}^2}$.

\mathbf{A} is r , only the first r singular values will be significant after addition of noise, assuming the noise is not too big.

Eigenvalues and eigenvectors of matrices $\mathbf{A}\mathbf{A}^\top$ and $\mathbf{A}^\top\mathbf{A}$ can be used to calculate SVD. Since the matrices \mathbf{U} and \mathbf{V} are unitary,

$$\mathbf{A}\mathbf{A}^\top = \mathbf{U}\mathbf{S}\mathbf{V}^\top\mathbf{V}\mathbf{S}^\top\mathbf{U}^\top = \mathbf{U}\mathbf{S}\mathbf{S}^\top\mathbf{U}^\top. \quad (5.7)$$

Therefore using $\mathbf{U}^\top = \mathbf{U}^{-1}$

$$\mathbf{A}\mathbf{A}^\top\mathbf{u}_j = s_j^2\mathbf{u}_j, \quad (5.8)$$

which implies that the left singular vectors \mathbf{u}_j are eigenvectors of $\mathbf{A}\mathbf{A}^\top$ with eigenvalues s_j^2 . Similarly it holds

$$\mathbf{A}^\top\mathbf{A}\mathbf{v}_j = s_j^2\mathbf{v}_j, \quad (5.9)$$

so that the right singular vectors \mathbf{v}_j are eigenvectors of $\mathbf{A}^\top\mathbf{A}$ with eigenvalues s_j^2 . For numerical implementation of SVD, see book of Press et al. [166]. In *Mathematica* software, SVD is calculated as $\{\mathbf{U}, \mathbf{S}, \mathbf{V}\} = \text{SingularValueDecomposition}[\mathbf{A}]$, in *MATLAB*, $[\mathbf{U}, \mathbf{S}, \mathbf{V}] = \text{svd}(\mathbf{A})$ or $[\mathbf{U}, \mathbf{S}, \mathbf{V}] = \text{svd}(\mathbf{A}, 'econ')$ for the economy version.

5.1.2 Pseudoinverse and linear regression

Pseudoinverse is a generalization of matrix inverse for singular or rectangular matrices also called Moore-Penrose inverse. Unlike the inverse where $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{1}$, the pseudoinverse \mathbf{A}^\dagger of a real matrix must satisfy these four criteria [165, 167]

$$\mathbf{A}\mathbf{A}^\dagger\mathbf{A} = \mathbf{A}, \quad (5.10a)$$

$$\mathbf{A}^\dagger\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger, \quad (5.10b)$$

$$(\mathbf{A}\mathbf{A}^\dagger)^\top = \mathbf{A}\mathbf{A}^\dagger, \quad (5.10c)$$

$$(\mathbf{A}^\dagger\mathbf{A})^\top = \mathbf{A}^\dagger\mathbf{A}. \quad (5.10d)$$

Pseudoinverse exists for every matrix and can be expressed by an explicit formula for a full rank matrix (i.e. $\text{rank}(\mathbf{A}) = \min(m, n)$). In the case of tall-skinny ($m > n$) matrix, the full rank ensures the existence of the inverse $(\mathbf{A}^\top\mathbf{A})^{-1}$, and the conditions in Eq. (5.10a,c) imply

$$\mathbf{A}^\dagger = (\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top. \quad (5.11)$$

Similarly for a short-fat ($m < n$) matrix, the conditions in Eq. (5.10b,d) yield

$$\mathbf{A}^\dagger = \mathbf{A}^\top(\mathbf{A}\mathbf{A}^\top)^{-1}. \quad (5.12)$$

Economy SVD can be used to calculate pseudoinverse when Eq. (5.2) or Eq. (5.3) is substituted into the first pseudoinverse condition in Eq. (5.10a) yielding

$$\mathbf{A}^\dagger = \mathbf{V}\mathbf{S}_n^{-1}\mathbf{U}_n^\top \text{ for } m > n, \quad (5.13a)$$

$$\mathbf{A}^\dagger = \mathbf{V}_m\mathbf{S}_m^{-1}\mathbf{U}^\top \text{ for } m < n. \quad (5.13b)$$

Let us consider a system of linear equations defined by a full-rank $m \times n$ matrix \mathbf{A} ,

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad (5.14)$$

which consists of m equations and a vector of n unknowns \mathbf{x} . For a square matrix \mathbf{A} , the solution is obtained simply as $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. If \mathbf{A} is a tall-skinny ($m > n$) matrix, Eq. (5.14) provides an overdetermined system of equations, which cannot be solved exactly. However, an approximate

solution with minimal least squares error $\|\mathbf{Ax} - \mathbf{b}\|^2$ can be calculated using the pseudoinverse. Substitution of Eq. (5.2) and (5.13) into Eq. (5.14) reveals

$$\mathbf{x} = \mathbf{A}^\dagger \mathbf{b} = \mathbf{V} \mathbf{S}_n^{-1} \mathbf{U}_n^\top. \quad (5.15)$$

This solution of an overdetermined system is also called *linear regression* and it is widely used in statistics [168]. On the other hand, if \mathbf{A} is a short-fat ($m < n$) matrix, Eq. (5.14) is an underdetermined system with infinitely many solutions. Among those, pseudoinverse in $\mathbf{x} = \mathbf{A}^\dagger \mathbf{b}$ yields a solution with minimal norm $\|\mathbf{x}\|$.

5.2 Analysis of UV/vis spectral series via SVD

Series of spectral measurements, where one parameter is changed, are an important tool of chemical physics. The parameter of change is usually concentration (titration experiments), time or temperature [48]. The methods presented here are applicable if the spectral changes depend linearly on concentration of absorbing components, in UV/vis spectroscopy it is limited by validity of the Beer-Lambert law (see Sec. 4.2). UV/vis titrations are usually analyzed at a particular wavelength with the biggest spectral change, often at maximum of an absorbing band λ_{\max} (see Sec. 4.3). However, if some of the absorbing components have small absorption coefficient $\epsilon_j(\lambda_{\max})$, their concentration is obtained with large uncertainty. A quick remedy in this case is simultaneous fitting of data at more than one wavelength. However, the more wavelengths are analyzed at once, the more parameters (absorption coefficients $\epsilon(\lambda_j)$) should be fitted. Here comes the SVD based analysis, which enables fitting data at all measured wavelengths at once while reducing the number of fitting parameters. Moreover, it allows elimination of the experimental noise and reconstruction of the individual absorbing species, which is not possible with the simple one-

Table 5.1: Terminology for SVD-based analysis of UV/vis spectral series.

	mathematical description	application on spectral series
\mathbf{u}_j	left singular vector	basis spectrum (factor (for $j \leq N_{\text{spc}}$))
\mathbf{v}_j	right singular vector	amplitude vector
s_j	singular value	singular value
N_{exp}	n	No. of measured experiments
N_{pts}	m	No. of points in one experiment
N_{spc}		No. of individual absorption species (factor dimension)
\mathbf{Z}		matrix of absorbing components spectra
\mathbf{z}_j		spectrum of j -th absorbing component (columns of \mathbf{Z})
\mathbf{P}_{exp}		population matrix constructed from exp. data
\mathbf{P}_{mod}		population matrix calculated from a model
\mathbf{p}_j		populations of j -th component (columns of \mathbf{P})
$s_j \mathbf{v}_j$		combination coefficients
\mathbf{T}		transformation matrix (often denoted as rotation matrix \mathbf{R})
\mathbf{t}_j		component coordinates in comb. coeff. space (columns of \mathbf{T})

wavelength approach. SVD terminology is adjusted for application in UV/vis spectra analysis. The used terms are listed in the following Table 5.1 including not yet defined quantities.

The SVD-based analysis is illustrated with two experiments in this chapter: (i) titration of **Bz₂OxP** with (*R*)-CSA in Fig. 5.2a. Halochromic spectral change in these data comprises of growing a new absorption band at 750 nm and decrease of a 500 nm band (green arrows). Detailed analysis of this measurement is given in Appendix D.1. (ii) Reichardt's dye in solvent mixture of varying polarity in Fig. 5.2b shows solvatochromic shift (blue arrow) that cannot be modeled as linear combination of individual absorbing species (Sec. 4.5).

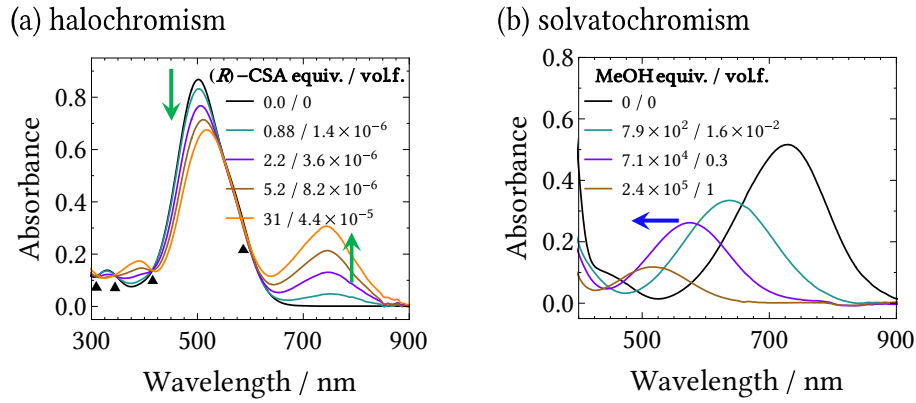


Figure 5.2: UV/vis spectra with two different types of spectral changes, (a) halochromic spectral change, titration of **Bz₂OxP** (7 μ M (initial), CDCl₃) with (*R*)-CSA, dilution during experiment compensated by rescaling to the initial concentration, black triangles denote the isosbestic points (b) solvatochromic spectral change, Reichardt's dye (55 μ M, CDCl₃) in MeOH/chloroform mixture (MeOH increases solvent polarity).

For the purpose of SVD analysis, the measured spectra \mathbf{a}_j are arranged into columns of the absorbance matrix \mathbf{A} , see Fig. 5.3. All spectra must be in the same wavelength range. The

$$\mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{V}^T$$

Figure 5.3: Illustration of the SVD of UV/vis spectral series. Constructed from the data in Fig. 5.2a, spectra in columns \mathbf{a}_j are measured at different acid concentration. Gray values correspond to noise.

abstract basis spectra \mathbf{u}_j will be transformed into the real spectra of absorbing components \mathbf{z}_j , the amplitude vectors \mathbf{v}_j will be transformed into vectors of populations \mathbf{p}_j . Products $s_j \mathbf{v}_j$ are denoted as *combination coefficients* [160] as they combine the basis spectra \mathbf{u}_j to reproduce the matrix \mathbf{A} (Eq. (5.4)). Note that the basis spectrum \mathbf{u}_1 is equal to mean of all measured spectra (possibly multiplied by -1 due to the aforementioned ambiguity in SVD), therefore, it never crosses the x -axis. On the other hand, all other \mathbf{u}_j must always cross the axis in order to preserve orthogonality to \mathbf{u}_1 .

5.2.1 Linear combination of absorbing components

The Beer-Lambert law for mixture of absorbing components in Eq. (4.5) can be expressed in matrix form as $\mathbf{a}_i = \sum_j l\epsilon_j(\mathbf{c}_j)_i$, where \mathbf{a}_i is i -th column of the data matrix \mathbf{A} , ϵ_j is the absorption coefficient of j -th absorbing component and $(\mathbf{c}_j)_i$ is the concentration of j -th absorbing component in i -th measured spectrum (i.e. $\mathbf{A} = l\mathcal{E}\mathbf{C}^\top$). If the total concentration of absorbing components is $[\mathbf{X}]_t = [\mathbf{X}_1] + [\mathbf{X}_2] + \dots + [\mathbf{X}_{N_{\text{spc}}}]$, the concentration \mathbf{c}_j can be rescaled to relative population $\mathbf{p}_j = \mathbf{c}_j/[\mathbf{X}]_t$, and ϵ_j can be rescaled to $\mathbf{z}_j = l\epsilon_j[\mathbf{X}]_t$ (i.e. absorbance at the total concentration).⁶ Thus, the total absorbance \mathbf{A} can be expressed in terms of absorbance of the absorbing components \mathbf{Z} and their populations \mathbf{P} ,

$$\mathbf{A} = \mathbf{Z}\mathbf{P}^\top, \quad (5.16)$$

see Fig. 5.4. Transformation from $\mathbf{A} = l\mathcal{E}\mathbf{C}^\top$ to Eq. (5.16) is convenient for UV/vis data processing.

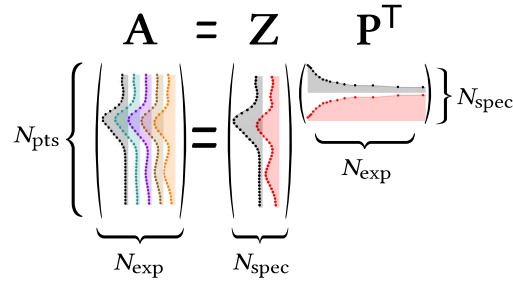


Figure 5.4: Decomposition of halochromic spectra into two absorbing components. Constructed from the data in Fig. 5.2a, spectra in columns \mathbf{a}_j are measured at different acid concentration.

In order to ensure that the populations sum to unity in every column of \mathbf{P} , $[\mathbf{X}]_t$ should be constant throughout the experiment. Otherwise the spectra should be rescaled to the initial concentration.⁷ In our example of halochromic spectra in Fig. 5.2a, the total concentration decreased since the sample was diluted by adding stock solution during the experiment. The spectra were indeed rescaled to the initial concentration since the isosbestic points appeared. Rescaling has little impact on the SVD except for the amplitude vector \mathbf{v}_1 , which is sensitive to total concentration, see Appendix D.1.

5.2.2 Determining the number of absorbing species

One of the key moments in SVD-based analysis is determination of N_{spc} , the number of individual absorbing species in the sample. This task is analogous to finding rank of the signal in a matrix containing signal and noise. The amplitude of s_j is proportional to the significance of \mathbf{u}_j and \mathbf{v}_j in

⁶For illustration, spectrum formed by two absorbing components at concentrations $[\mathbf{X}_1] + [\mathbf{X}_2] = [\mathbf{X}]_t$ and the first measured spectrum \mathbf{a}_1 can be expressed as

$$\mathbf{a}_1 = l\epsilon_1[\mathbf{X}_1] + l\epsilon_2[\mathbf{X}_2] = \underbrace{l\epsilon_1[\mathbf{X}]_t}_{z_1} \underbrace{\frac{[\mathbf{X}_1]}{[\mathbf{X}]_t}}_{(p_1)_1} + \underbrace{l\epsilon_2[\mathbf{X}]_t}_{z_2} \underbrace{\frac{[\mathbf{X}_2]}{[\mathbf{X}]_t}}_{(p_2)_1}.$$

⁷Assume that $(\mathbf{c}_t)_j$ is defined as the total concentration at the j -th measurement, and $(\mathbf{c}_t)_1 = (\mathbf{c}_1)_1 + (\mathbf{c}_2)_1$. Then, the measured spectra are rescaled to the initial concentration using

$$\begin{aligned} \mathbf{a}_1^{\text{resc}} &= \mathbf{a}_1^{\text{orig}}, \\ \mathbf{a}_j^{\text{resc}} &= \mathbf{a}_j^{\text{orig}} \frac{(\mathbf{c}_t)_1}{(\mathbf{c}_t)_j} = \underbrace{l\epsilon_1(\mathbf{c}_t)_1}_{z_1} \underbrace{\frac{(\mathbf{c}_1)_j}{(\mathbf{c}_t)_j}}_{(p_1)_j} + \underbrace{l\epsilon_2(\mathbf{c}_t)_2}_{z_2} \underbrace{\frac{(\mathbf{c}_2)_j}{(\mathbf{c}_t)_j}}_{(p_2)_j} \text{ for } j = 2, \dots, N_{\text{exp}}. \end{aligned}$$

the reconstruction of \mathbf{A} according to Eq. (5.5). Thus, inspection of logarithmic plot of the singular values is a great indicator for obtaining N_{spc} . Figure 5.5a, obtained from our halochromic example (Fig. 5.2a), shows an obvious gap after the second singular value, suggesting two-component spectral change. On the contrary, our solvatochromic example (Fig. 5.2b) does not show any gap between singular values, see Fig. 5.5b. This nonlinear spectral change with bands shifts requires many-dimensional SVD approximation to account for. SVD of uncorrelated Gaussian noise reveals typical exponential decay (straight line in logarithmic plot), see Fig. 5.5c. Similar decay for the noise component is exhibited in Fig. 5.5a (s_3-s_{15}) and Fig. 5.5b (ca. s_5-s_{12}). The highest singular value in (c) provides a crude estimate of noise threshold in Fig. 5.5a (blue line, noise level estimate based on the Fig. 5.2a).

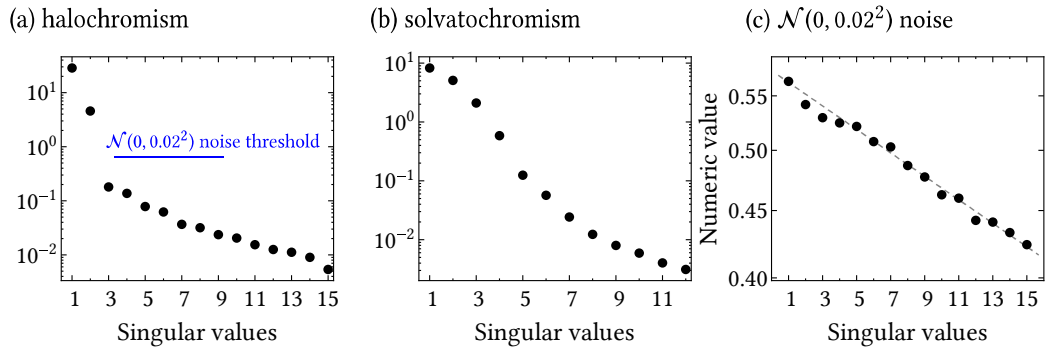


Figure 5.5: Log-plot of singular values with (a) halochromic and (b) solvatochromic spectral changes, (c) uncorrelated noise, approximately following (gray dashed) straight line. Spectral data for (a), (b) are given in Fig. 5.2. Noise data for (c) have normal distribution $\mathcal{N}(0, 0.02^2)$ (mean=0, std.=0.02) and the same dimensions as (a), the noise amplitude was estimated from Fig. 5.5a. The noise threshold, blue line in (a), is given by the highest singular value in (c). Values in (a) and (b) were obtained from the spectra shown in Fig. 5.2a,b, respectively.

The cumulative sum $s_1^2 + s_2^2 + \dots + s_j^2$ is proportional to the amount of data from \mathbf{A} reconstructed by j -dimensional SVD approximation according to Eq. (5.5) (or the amount of reconstructed variance in PCA, cf. Sec. 5.3.2). Thus, the proportion of reconstructed data, also called the *cumulative energy* [157], is defined as

$$\text{cum.en.}(j) = \frac{\sum_{k=1}^j s_k^2}{\sum_{l=1}^{N_{\text{exp}}} s_l^2}. \quad (5.17)$$

In our halochromic example, the cumulative energy for $j = 2$ already accounts for almost 100% of data, see Fig. 5.6a. Another quantity, *residual standard deviation* (RSD) determines the error of j -dimensional SVD approximation as (equation (4.44) in Malinowski [152])

$$\text{RSD}(j) = \sqrt{\frac{\sum_{k=j+1}^{N_{\text{exp}}} s_k^2}{N_{\text{pts}}(N_{\text{exp}} - j)}}. \quad (5.18)$$

If the experimental error is known, it can be directly compared to the RSD value and serve as a criterion for N_{spc} determination. RSD of our halochromic example (Fig. 5.2a) is given in Fig. 5.6b. There is significant drop of RSD for $j = 2$ under the estimated experimental error threshold (std=0.02), which is another indicator of the presence of two absorbing species.

If the signal-to-noise ratio is too low, it is hard to determine N_{spc} just from the visual inspection of singular values, cumulative energy or RSD. If the noise level is known, the noise threshold can be crudely estimated (Fig. 5.5a,c). Estimation of the noise threshold for white noise is given

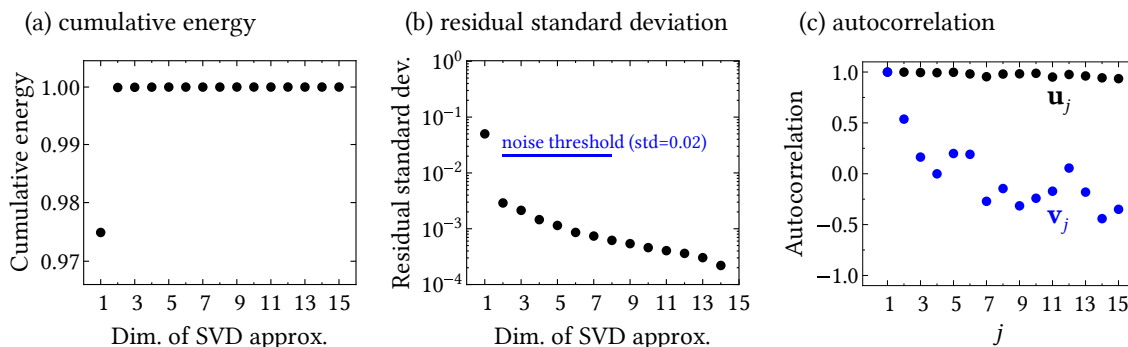


Figure 5.6: Various indicators for N_{spc} determination for our halochromic data (Fig. 5.2a), (a) cumulative energy and (b) residual standard deviation based on the singular values, (c) autocorrelation coefficients calculated for neighbouring points.

by Gavish and Donoho [169] even for unknown noise level. Determination of N_{spc} in UV/vis spectra is limited by the noise from the detector and illumination lamp (dependent on λ [161]) or by imperfect baseline subtraction. Shifts of the absorption bands due to solvatochromic or temperature effects or deviations from the Beer-Lambert law (Sec. 4.2) pose another challenges to SVD analysis. Complications also arise if vectors \mathbf{z}_j or vectors \mathbf{p}_j are approximately linearly dependent [161].

In the presence of uncorrelated noise, the noise-containing vectors \mathbf{u}_j and \mathbf{v}_j for $j > N_{\text{spc}}$ are often less smooth than the signal-containing vectors. The “smoothness” can be quantified as autocorrelation [158], i.e. correlation of a vector \mathbf{u}_j (or \mathbf{v}_j) with the same vector shifted by one data point (or more data points, if the noise is correlated at short distances [161]), in particular,

$$\text{autocorr.}(\mathbf{u}_j) = \sum_{k=1}^{N_{\text{pts}}-1} (\mathbf{u}_j)_k (\mathbf{u}_j)_{k+1}, \quad (5.19)$$

$$\text{autocorr.}(\mathbf{v}_j) = \sum_{k=1}^{N_{\text{exp}}-1} (\mathbf{v}_j)_k (\mathbf{v}_j)_{k+1}. \quad (5.20)$$

In our halochromic example (Fig. 5.2a), the basis spectra are smooth as the autocorrelation is close to unity, Fig. 5.6c. This is probably a consequence of data smoothing during the data processing. On the other hand, only the first two amplitude vectors have a big positive autocorrelation, which can be an indicator for $N_{\text{spc}} = 2$. Henry and Hofrichter [158] introduced a ‘rotation procedure’, which can transform set of p amplitude vectors (although not the first p vectors) to another set of p orthonormal vectors with extremal autocorrelation. Thus, some of the new vectors have bigger autocorrelation on the expense of others, which helps to “push” the signal forward. Nevertheless, we are not going to use this procedure in this work.

5.2.3 Transformation from abstract to real spectra

If all spectra of individual absorbing species can be measured in isolation, their populations can be obtained without SVD as described in Sec. 5.2.5. In the other case, SVD has to be applied together with a model for concentrations of the chemical species (e.g. model based on host-guest binding or chemical kinetics).

First N_{spc} basis spectra can be transformed to the spectra of absorbing components using the

transformation matrix \mathbf{T} ,⁸

$$\mathbf{Z} = \mathbf{U}_{N_{\text{spec}}} \mathbf{T}, \quad (5.21)$$

see Fig. 5.7.

$$\mathbf{Z} = \mathbf{U}_{N_{\text{spec}}} \mathbf{T}$$

Figure 5.7: Reconstruction of real spectra from abstract basis spectra \mathbf{u}_j . Constructed from the data in Fig. 5.2a.

Analogously, populations are reconstructed using the inverse matrix \mathbf{T}^{-1} ,

$$\mathbf{P}^{\text{T}} = \mathbf{T}^{-1} \mathbf{S}_{N_{\text{spec}}} \mathbf{V}_{N_{\text{spec}}}^{\text{T}} \quad (5.22)$$

since $\mathbf{A} = \mathbf{U}_{N_{\text{spec}}} \mathbf{T} \mathbf{T}^{-1} \mathbf{S}_{N_{\text{spec}}} \mathbf{V}_{N_{\text{spec}}}^{\text{T}} = \mathbf{Z} \mathbf{P}^{\text{T}}$, see Fig. 5.8. At this point, we have two ways of expressing

$$\mathbf{P}^{\text{T}} = \mathbf{T}^{-1} \mathbf{S}_{N_{\text{spec}}} \mathbf{V}_{N_{\text{spec}}}^{\text{T}}$$

Figure 5.8: Reconstruction of real populations from abstract amplitude vectors \mathbf{v}_j . Constructed from the data in Fig. 5.2a.

the matrix \mathbf{A} ,

$$\mathbf{A} = \mathbf{U}_{N_{\text{spec}}} \mathbf{S}_{N_{\text{spec}}} \mathbf{V}_{N_{\text{spec}}}^{\text{T}}, \quad (5.23)$$

and

$$\mathbf{A} = \mathbf{Z} \mathbf{P}^{\text{T}} = \mathbf{U}_{N_{\text{spec}}} \mathbf{T} \mathbf{P}^{\text{T}}. \quad (5.24)$$

Comparing these two expressions, we get

$$\mathbf{S}_{N_{\text{spec}}} \mathbf{V}_{N_{\text{spec}}}^{\text{T}} = \mathbf{T} \mathbf{P}^{\text{T}}. \quad (5.25)$$

The left-hand side of this equation is fixed and based solely on SVD. The populations \mathbf{P} follow a system-dependent model derived from thermodynamics or chemical kinetics, see Chapter 2. Hence, the right-hand side of Eq. (5.25) should be adjusted by least squares fitting of elements of \mathbf{T} and model parameters for \mathbf{P} (binding constants or reaction rate coefficients). Thus, we distinguish populations \mathbf{P}_{exp} obtained from the experiment after calculating \mathbf{T} (Eq. (5.22), typically noisy) from populations \mathbf{P}_{mod} obtained from a model (smooth, could be interpolated between experimental points), see Fig. 5.9b. In our halochromic example (Fig. 5.2a), the host-guest binding model with 1:1 stoichiometry (Eq. (2.26)) was used, and the function

$$\chi^2 = \|\mathbf{S}_{N_{\text{spec}}} \mathbf{V}_{N_{\text{spec}}}^{\text{T}} - \mathbf{T} \mathbf{P}_{\text{mod}}^{\text{T}}(K_{\text{HG}})\|^2 \quad (5.26)$$

⁸It is sometimes called ‘rotation procedure’ [152] (denoting the transformation matrix as \mathbf{R}), however, the transformation is not unitary.

was minimized.

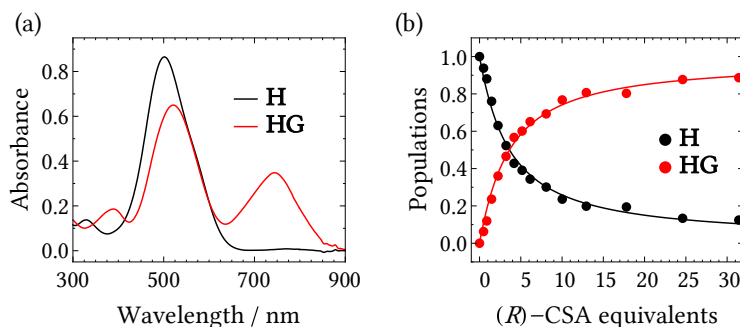


Figure 5.9: Reconstructed (a) spectra of absorbing species Z and (b) their populations P throughout the halochromic spectral change. Points stand for P_{exp} calculated from Eq. (5.22), lines for P_{mod} calculated from a 1:1 host-guest model (Eq. (2.26)).

The reconstructed spectra are shown in Fig. 5.9a. The first measured spectrum \mathbf{a}_1 with no acid added was identified as the first absorbing component \mathbf{z}_1 , therefore, the first column of \mathbf{T} was calculated from Eq. (5.21) and (5.15) as $\mathbf{t}_1 = \mathbf{U}_{N_{\text{spc}}}^\dagger \mathbf{a}_1^\top$ or simply read out from the first column of $\mathbf{S}_{N_{\text{spc}}} \mathbf{V}_{N_{\text{spc}}}^\top$ (combination coefficients of the first measured spectrum, see next section). Quality of the fit can be checked by comparison of populations P_{mod} and P_{exp} , see Fig. 5.9b.

5.2.4 Combination coefficients

Combination coefficients $s_j \mathbf{v}_j$ provide an illustrative “map” of the course of the experiment. Each point in plot of combination coefficients corresponds to a particular measurement. This plot is linear for our halochromic example (Fig. 5.2a) as shown in Fig. 5.10a (full circles). This is a rule for every spectral change involving exactly two absorbing species with constant dye concentration (or rescaled to constant dye concentration, see footnote 7 on page 58). In our case, also the original spectra measured with gradual sample dilution show linear dependence (empty circles) because the dilution was uniform. The change of total dye concentration is mostly captured by the first combination coefficient $s_1 \mathbf{v}_1$, see Fig. D.1f in Appendix D.3. Solvatochromic spectral change (Fig. 5.2b) shows systematic shift without the linear dependence, see Fig. 5.10b. Spectra consisting only of noise do not show any systematic shift, see Fig. 5.10c.

If a measured spectrum \mathbf{a}_j is the spectrum of an isolated individual absorbing species in two-component system, e.g. $\mathbf{a}_j = \mathbf{z}_1$, then Eq. (5.4) implies $\mathbf{z}_1 = \mathbf{u}_1 s_1 (\mathbf{v}_1^\top)_j + \mathbf{u}_2 s_2 (\mathbf{v}_2^\top)_j$. Hence, the combination coefficients are identical to T_{11} and T_{21} (cf. Eq. (5.21)) in this case. In other words, j -th column of \mathbf{T} (i.e. \mathbf{t}_j) constitutes coordinates of the j -th absorbing component in the combination coefficient plot. Coordinates of the first and second absorbing species are shown in Fig. 5.10a (compare the coordinates with Fig. 5.7). Since the measured spectra are linear combination of \mathbf{z}_1 and \mathbf{z}_2 , all experimental points lie on a curve $(T_{11}, T_{21})^\top p_1 + (T_{12}, T_{22})^\top (1 - p_1) = \mathbf{t}_1 p_1 + \mathbf{t}_2 (1 - p_1)$, where $p_1 \in [0, 1]$ is the population of the first component (dashed line in Fig. 5.10a). This is parametric equation for a line segment implying linear dependence in the two-component spectral change.

This linear dependence can be generalized to higher number of absorbing components. For example, spectral dataset with $N_{\text{spc}} = 3$ implies that the combination coefficients lie on a 2D plane inside a triangle given by \mathbf{t}_1 , \mathbf{t}_2 and \mathbf{t}_3 [161]. Also in such system, a spectral change two absorbing components (population of the third one is constant) implies that the corresponding combination coefficients lie on a straight line. Generally, N_{spc} components form an $(N_{\text{spc}} - 1)$ -dimensional object (a convex hull) in N_{spc} -dimensional space of combination coefficients, although small over-

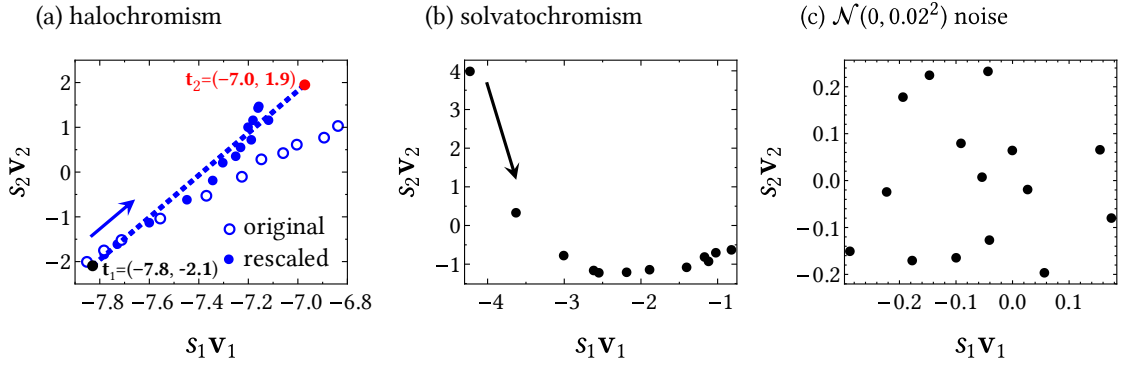


Figure 5.10: Plot of first two combination coefficients for (a) halochromic (data in Fig. 5.2a), (b) solvatochromic (data in Fig. 5.2b) spectral change and (c) spectra generated as $\mathcal{N}(0, 0.02^2)$ noise with dimensions of (a). Course of the experiment is indicated by arrows. In (a) black (red) circle corresponds to combination coefficients of the first (second) real component \mathbf{H} (HG), dashed line denotes the linear path between the two components in two-component spectral change.

lap with other dimensions occurs due to the noise. Combination coefficients corresponding to a spectral change involving r components form an $(r - 1)$ -dimensional object in N_{spc} -dimensional space of combination coefficients.

5.2.5 Decomposition to the absorbing components without SVD

If all spectra of individual components \mathbf{Z} can be measured separately, transformation from abstract to real vectors is not necessary and their populations can be calculated without SVD as

$$\mathbf{P}_{\text{exp}}^{\text{T}} = \mathbf{Z}^{\dagger} \mathbf{A}, \quad (5.27)$$

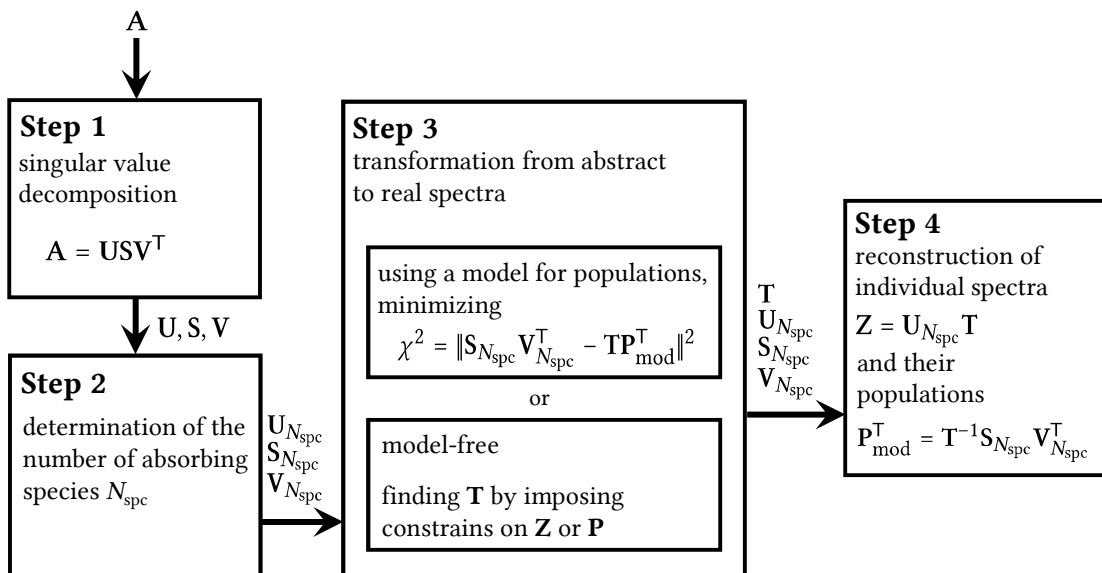
cf. Eq. (5.15). These populations \mathbf{P}_{exp} obtained directly from the experiment serve as binding isotherms and can be fitted with \mathbf{P}_{mod} using a binding or kinetic model to obtain binding constants. In our publication [1], we used this type of analysis and called it ‘decomposition to spectra of individual species’ (DSIS). Instead of Eq. (5.27) the equivalent approach of least squares fitting was used and we minimized

$$\chi^2 = \|\mathbf{A} - \mathbf{Z}\mathbf{P}_{\text{exp}}^{\text{T}}\|^2 \quad (5.28)$$

with all elements of $\mathbf{P}_{\text{exp}}^{\text{T}}$ being fitting parameters. In contrary to pseudoinverse, least squares fitting allows for constrains such as $\sum \mathbf{p}_j = 1$ or $0 \leq \mathbf{p}_j \leq 1$.

5.2.6 SVD analysis – summary

The SVD-based analysis comprises of four basic steps described in Scheme 5.1.



Scheme 5.1: Steps in SVD-based analysis of UV/vis spectral series. Model-free determination of T is explained in [154].

5.3 Analogy with principal component analysis

5.3.1 Calculation of principal components

Principal component analysis (PCA) is a statistical tool identifying directions of maximal variance in multidimensional datasets and consequently finding patterns (clustering of datapoints) in high-dimensional data. We show in Sec. 5.3.2 that PCA is mathematically identical with SVD and can be regarded as statistical interpretation of SVD. In the PCA literature it is common to construct the analyzed matrix as rows of measurements in contrast to our SVD column-wise convention. We follow the usual PCA convention since the common programs as *Mathematica* or *MATLAB* use it too, and we denote the row-wise data matrix $B \equiv A^T$ (in this section B is a $m \times n$ matrix). As shown in Eq. (5.29), measurements \mathbf{b}_j^T (e.g. measured absorption spectra) are arranged in rows. Columns \mathbf{x}_j denote one particular feature (e.g. measured absorbance at one particular wavelength). In statistical application of PCA, the columns \mathbf{x}_j stand for particular analyzed features such as age, height and weight, whereas rows \mathbf{b}_j^T denote entries for a particular person.

$$B = \begin{pmatrix} \dots & \mathbf{b}_1^T & \dots \\ \dots & \mathbf{b}_2^T & \dots \\ \dots & \vdots & \dots \\ \dots & \mathbf{b}_m^T & \dots \end{pmatrix} = \begin{pmatrix} \vdots & \vdots & \dots & \vdots \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \\ \vdots & \vdots & \dots & \vdots \end{pmatrix}. \quad (5.29)$$

Unlike SVD, PCA requires mean-subtracted data $\bar{\mathbf{B}}$ (denoted with bar) defined as

$$\bar{\mathbf{B}} = \mathbf{B} - \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n), \quad (5.30)$$

where $\bar{x}_j = 1/n \sum_{k=1}^m (\mathbf{x}_j)_k$ is arithmetic mean of j -th feature.

PCA is searching for directions α_j in which a linear transformation of mean-centered feature vectors $\bar{\mathbf{x}}_j = \mathbf{x}_j - \bar{x}_j$ captures maximum of measured variance. These direction vectors have unit norm and zero mutual covariance,⁹ i.e. they are uncorrelated. The transformed vectors are denoted as *principal components* (PCs), and the transformation from mean-centered feature vectors reads

$$\underbrace{\begin{pmatrix} \vdots & \vdots & \vdots \\ \text{PC}_1 & \text{PC}_2 & \dots & \text{PC}_n \\ \vdots & \vdots & \vdots \end{pmatrix}}_{m \times n} = \underbrace{\begin{pmatrix} \vdots & \vdots & \vdots \\ \bar{\mathbf{x}}_1 & \bar{\mathbf{x}}_2 & \dots & \bar{\mathbf{x}}_n \\ \vdots & \vdots & \vdots \end{pmatrix}}_{m \times n} \underbrace{\begin{pmatrix} \vdots & \vdots & \vdots \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \end{pmatrix}}_{n \times n}. \quad (5.31)$$

PCs are automatically mean-centered due to the centering of feature vectors. PCs are sometimes denoted as *scores* ($m \times n$ matrix) and the directions α_j as *loadings* ($n \times n$ matrix).

Search for PCs is performed [144] by maximization of $\text{Var}(\text{PC}_j)$ using Lagrange multipliers with normalization conditions $\alpha_j^\top \alpha_j = 1$ and zero covariance between different PCs, i.e. $\text{Cov}(\text{PC}_j, \text{PC}_k) = 0$ (equivalent to $\alpha_j^\top \alpha_k = 0$). It turns out that the solution is equivalent to finding eigenvectors of the covariance matrix $1/(n-1) \bar{\mathbf{B}}^\top \bar{\mathbf{B}}$,

$$\frac{1}{n-1} \bar{\mathbf{B}}^\top \bar{\mathbf{B}} \alpha_j = \lambda_j \alpha_j. \quad (5.32)$$

Since the variance captured by j -th PC is equal to its corresponding eigenvalue λ_j (as shown below), the PCs and corresponding loadings α_j are sorted in descending order according to their eigenvalues,¹⁰ $\lambda_1 > \lambda_2 > \dots > \lambda_n$. The covariance matrix $1/(n-1) \bar{\mathbf{B}}^\top \bar{\mathbf{B}}$ is symmetric (Hermitian), therefore, its eigenvalues are real and its eigenvectors are orthogonal.

5.3.2 Comparison to SVD

SVD of the mean-subtracted data matrix is $\bar{\mathbf{B}} = \mathbf{U} \mathbf{S} \mathbf{V}^\top$. Recalling the calculation of SVD vectors as eigenvalues (Eq. (5.9)), we already know that

$$\bar{\mathbf{B}}^\top \bar{\mathbf{B}} \mathbf{v}_j = s_j^2 \mathbf{v}_j. \quad (5.33)$$

Comparison to Eq. (5.32) implies that $\lambda_j = s_j^2/(n-1)$ and $\alpha_j = \mathbf{v}_j$. Furthermore, comparing SVD in the form $\mathbf{U} \mathbf{S} = \bar{\mathbf{B}} \mathbf{V}$ to Eq. (5.31) yields $\text{PC}_j = \mathbf{u}_j s_j$.

This analogy with SVD allows for a quick proof that the variance captured by PC_j is λ_j , in particular $\text{Var}(\text{PC}_j) = 1/(n-1) \text{PC}_j \text{PC}_j^\top = s_j^2/(n-1) \mathbf{u}_j \mathbf{u}_j^\top = s_j^2/(n-1) = \lambda_j$. Relative amount of captured variance by first j PCs is then calculated as $\sum_{k=1}^j \lambda_k / \sum_{l=1}^n \lambda_l$ (cf. the cumulative energy in Eq. (5.17)). Due to the analogy of PCA and SVD, PCA of real datasets can be calculated using numerical algorithms for SVD.

⁹Covariance is defined as $\text{Cov}(\mathbf{x}, \mathbf{y}) = 1/(n-1) \sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})$ and variance as $\text{Var}(\mathbf{x}) = \text{Cov}(\mathbf{x}, \mathbf{x})$. Covariance calculated from measured data is also called ‘sample covariance’ because the true value is not known. To obtain unbiased value, $n-1$ should be in the denominator rather than n .

¹⁰Since the experimental data contain noise, usually no pair of eigenvalues have the same value.

5.3.3 Application

Standard application of PCA is visualization of high-dimensional data in low-dimensional plot of principal components. This allows to observe clustering of data points, revealing similarities or differences in various data sets. We illustrate PCA analysis with three artificial data sets shown in Fig. 5.11a. Age, height and weight are the observed features ($n = 3$), and each data set has 150 entries (i.e. people, $m = 150$). PCA is calculated from all three data sets combined. Mutual plot of PC_1 and PC_2 (together capturing 92.9% of variance) in Fig. 5.11b projects the 3D data to 2D plot and shows that data sets 1 and 2 mostly overlap, while data set 3 is significantly different. The corresponding loadings α_1 and α_2 represent directions in the combined data set, where the data are maximally “stretched out”. Mutual plot of PC_2 and PC_3 in Fig. 5.11c confirms that PC_3 captures the smallest variance. Note that k -th data point $\bar{\mathbf{b}}_k$ can be projected to j -th PC using the loading α_j according to $(PC_j)_k = \bar{\mathbf{b}}_k^T \cdot \alpha_j$ as follows from Eq. (5.31). In real life applications of PCA, high-dimensional data are processed (i.e. with large number of features), for example, when thousands of genetic markers are analyzed to observe clustering based on common ancestry [170]. Recent developments of PCA are reviewed by Jolliffe and Cadima [171]. In this work, we use PCA to analyze combination coefficients obtained from SVD, the procedure is described in detail in Appendix D.5.

In *Mathematica* software, PCs are obtained by `PC=PrincipalComponents[B]`, in *MATLAB*, the command `[alpha,PC,lambda]=pca(B)` is used to obtain loadings, PC scores and the variances. Both applications use the row-wise data matrix according to Eq. (5.31) and subtract the mean automatically.

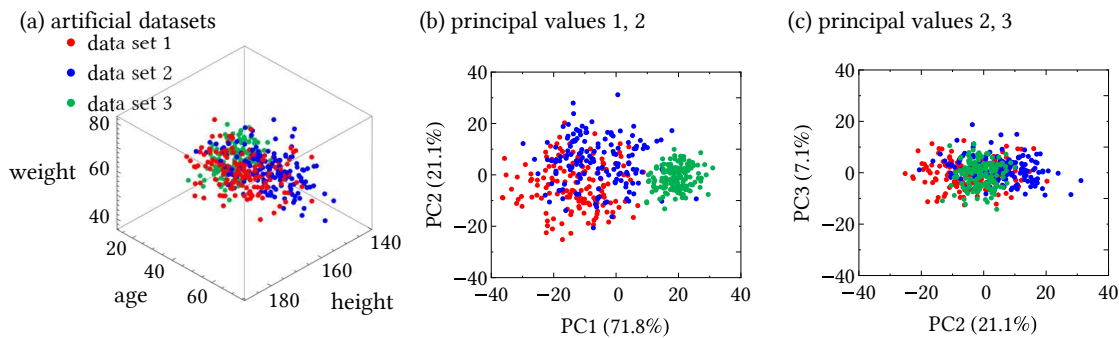


Figure 5.11: Illustration of PCA, (a) three artificially constructed data sets with three features (normally distributed), each of 150 entries. (b) Mutual plot of first two principal components capturing 92.9% of the variance, (c) mutual plot of principal components 2 and 3.

6 Chromic and binding properties of OxP derivatives

As discussed in Sec. 4.5, chromic changes in dye molecules can be caused by protonation (halochromism) or by change of solvent properties (solvatochromism). In this chapter we show our results about halochromism and solvatochromism of the oxoporphyrinogen derivatives **OxP**, **Bz₂OxP** and **Bz₄OxP**. Aside from the complete description of color variation, particular protonated species are identified and the binding constants were determined, where possible. To remind the structures of oxoporphyrinogens, see Fig. 6.1. The **OxP** molecule possesses two calix[4]pyrrole-type binding sites (shortly NH/amine binding sites, black arrows) and four protonation sites at the carbonyl groups (C=O groups, green arrows). N-benylation at the macrocyclic nitrogen atoms of **OxP** leads to stepwise elimination of the two NH binding sites (red crosses) and subsequent serial variation of the chromic properties of the N-benzylated products. The color related properties were studied by UV/vis spectroscopy, however, to verify our conclusions related to protonation, NMR titration experiments (i.e. measurement of spectra at different guest concentrations when the guest is gradually added into the host solution) were conducted. Following sections on chromism and binding mostly follow our article [1] including parts of text, most figures were adopted from [1] with modifications. Several NMR spectra analyzed here and some preliminary results were presented in the master's thesis [45]. Experimental details about the measurements are given in Appendix C.2.

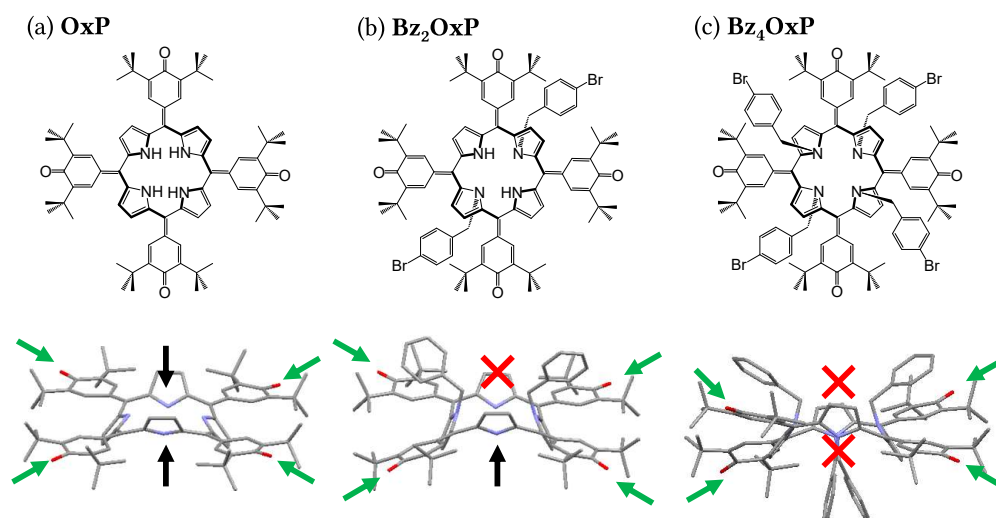


Figure 6.1: Chemical structures (top) and X-ray crystal structures (bottom) of (a) oxoporphyrinogen (**OxP**), (b) bis(4-bromobenzyl)oxoporphyrinogen (**Bz₂OxP**), (c) tetrakis(4-bromobenzyl)oxoporphyrinogen (**Bz₄OxP**) used in this work. X-ray crystal structure in panels (b,c) are of the non-brominated derivatives [172]. Black arrows: accessible calix[4]pyrrole-type binding sites, red crosses: blocked binding sites, green arrows: protonation carbonyl sites.

6.1 Halochromism of **OxP**

Colorimetric response of **OxP** was reported for the first time by Shundo et al. [31]. The cause of color variation upon addition of trifluoroacetic acid (TFA) was the presence of di- and tetraprotonated **OxP** species. The protonation at the hemiquinonoid carbonyl sites was proven by direct observation of the intensity of the C-O-H infrared vibration, see Fig. 6.2. Since the binding steps $\text{HG} \rightarrow \text{HG}_2$ and $\text{HG}_3 \rightarrow \text{HG}_4$ are strongly cooperative (i.e. high interaction parameters α_{12} and α_{34} , see Sec. 2.2.4), mono- and tri- protonated **OxP** species are virtually absent. In our publication [1], we reanalyzed the UV/vis data from the study [31] to assess quantitatively the halochromic properties of **OxP**.

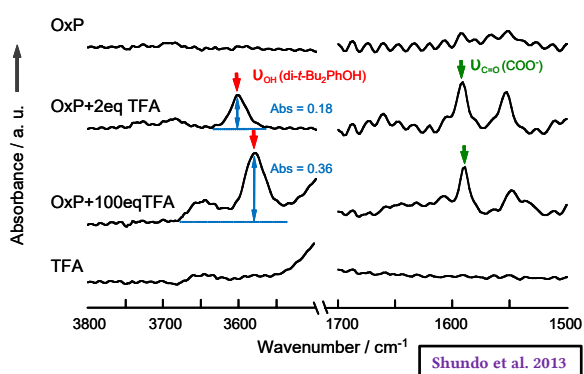


Figure 6.2: Solution (CH_2Cl_2 , 0 °C) state FTIR spectra of **OxP** with various amounts of TFA (0, 2 and 100 equiv. of TFA) and pure TFA. Absorbance of the **OxP** hydroxyl stretching peak is indicated. Appearance of the OH stretching signal is due to first diprotonation (2 equiv. of TFA) then further protonation at higher concentration of TFA. Adopted from [31].

plot further confirms the aforementioned consecutive protonation scheme. In Fig. 6.3g, the OxP^{2+} coordinates do not lie exactly at the intersection due to small change in total dye concentration, see details in Appendix D.2.

As the two spectral changes are separated, the spectra of individual absorbing species, Fig. 6.3c, were identified at the end of each spectral change, thick lines in Fig. 6.3a,b. Since the observed spectral changes are purely of halochromic nature (see Sec. 5.2.1), the measured UV/vis spectra are linear combination of the absorbing components spectra, $\mathbf{A} = \mathbf{Z}\mathbf{P}_{\text{exp}}^T$. Analysis from Sec. 5.2.5 referred to as ‘decomposition into the spectra of individual species’ (DSIS) can be applied. DSIS of \mathbf{a}_{φ_j} , measured spectrum at volume fraction φ_j , is formulated as

$$\mathbf{a}_{\varphi_j} = p_{\text{OxP}}(\varphi_j) \mathbf{z}_{\text{OxP}} + p_{\text{OxP}^{2+}}(\varphi_j) \mathbf{z}_{\text{OxP}^{2+}} + p_{\text{OxP}^{4+}}(\varphi_j) \mathbf{z}_{\text{OxP}^{4+}}, \quad (6.1)$$

where \mathbf{z}_i is the spectrum of i -th absorbing component and $\mathbf{p}_i(\varphi_j)$ its population at DFA volume fraction φ_j . Populations \mathbf{P}_{exp} obtained by DSIS (with constraints $\mathbf{p}_j > 0$ and $\sum \mathbf{p}_j = 1$) are shown in Fig. 6.3d (solid circles). The quality of DSIS is excellent, as the error $\|\mathbf{A} - \mathbf{Z}\mathbf{P}_{\text{exp}}^T\|$ is small, see Fig. D.3 in Appendix D.2. Decomposition according to Eq. (6.1) would not be possible in the presence of solvatochromic shift.

¹The amount of added guest is described by guest equivalents defined as $\text{equiv.} = [\mathbf{G}]_t/[\mathbf{H}]_t$. In addition, volume fraction (vol.f./ φ) is used, especially for later description of solvatochromism (polarity of the sample depends on vol.f. of its components).

Titration spectra¹ in Fig. 6.3a,b show two consecutive spectral changes. The presence of isosbestic points (see Sec. 4.4) indicates consecutive transitions (i.e. only two species are present at a time) $\text{OxP} \rightarrow \text{OxP}^{2+}$ and $\text{OxP}^{2+} \rightarrow \text{OxP}^{4+}$ corresponding to the first and second spectral change, respectively. Singular value decomposition (SVD) analysis of the spectra confirms the presence of three absorbing species based on three significant singular values as shown in Fig. 6.3f (see Sec. 5.2.2). Mutual plot of combination coefficients in Fig. 6.3g,h shows two linear segments connecting all three component coordinates. Since arrangement of combination coefficients on a straight line implies that exactly two components are involved in the particular spectral change (see Sec. 5.2.4), this

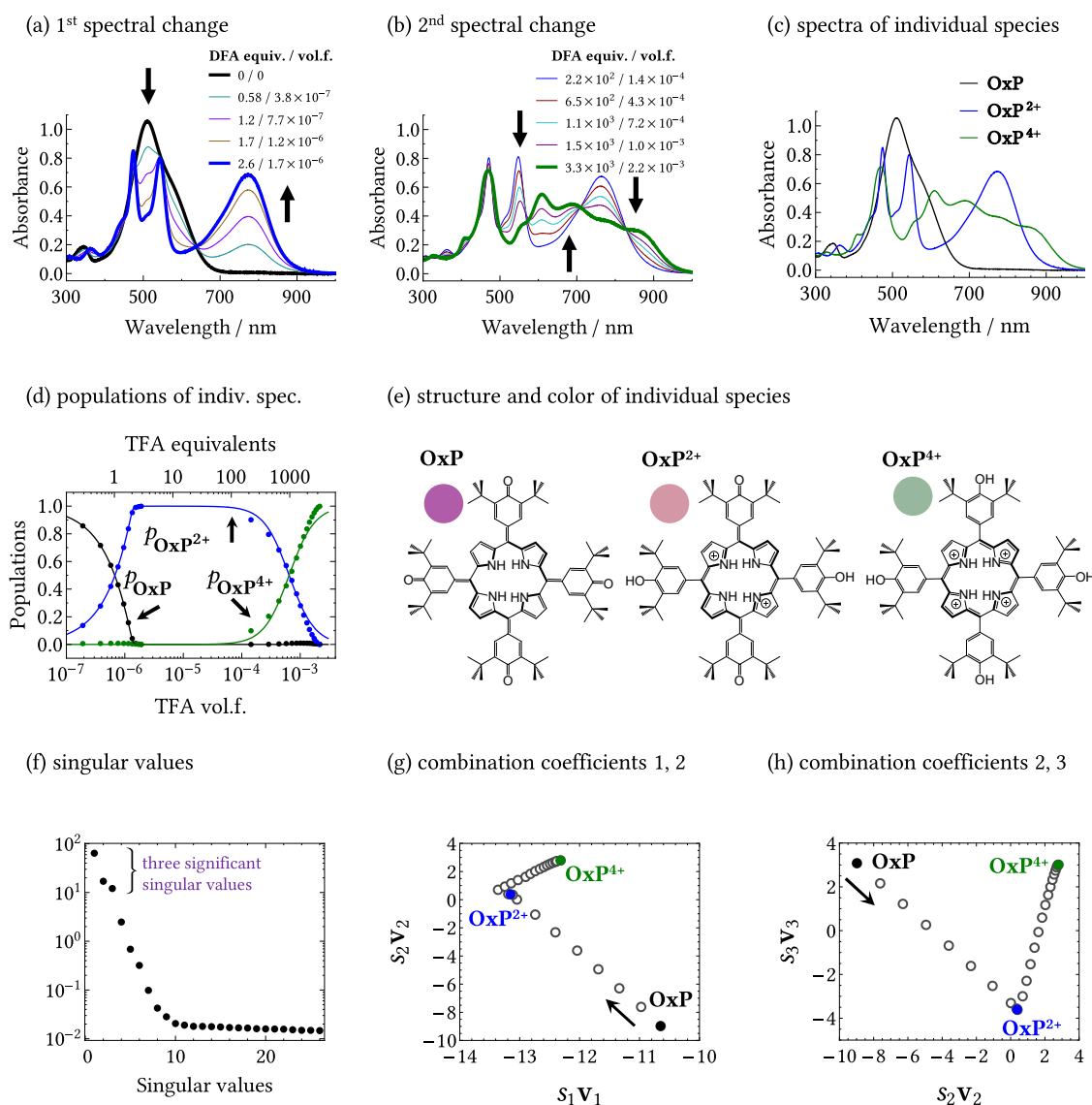


Figure 6.3: UV/vis titration of OxP (10^{-5} M, CH_2Cl_2) with TFA and corresponding analyses. (a, b) First and second sets of spectral changes occurring during titration. These data were first published in [31]. (c) Spectra of individual absorbing species Z identified as the start and endpoints of each spectral change (bold lines in panels (a, b)). (d) Populations \mathbf{P}_{exp} of absorbing species Z as obtained by DSIS (solid circles). Fitting of \mathbf{P}_{exp} to 1:4 binding model provided model populations \mathbf{P}_{mod} (solid lines). (e) Chemical structures and RGB colors of absorbing species (colors calculated from the spectra shown in panel (c), details in Appendix C.2). SVD analysis confirms $N_{\text{spc}} = 3$ via plots of (f) singular values and (g,h) combination coefficients space.

Subsequently, we attempted to describe the interaction of OxP with TFA in terms of a host-guest binding model minimizing $\|\mathbf{P}_{\text{exp}} - \mathbf{P}_{\text{mod}}\|$, where \mathbf{P}_{mod} is a matrix of concentrations reconstructed using a particular binding model. Here, the first spectral change saturates at 2 guest equiv. implying double protonation. The second spectral change is saturated at relatively low concentration of TFA ($\varphi = 0.002$ in CH_2Cl_2), thus, equations governing the chemical equilibria can be used without correction for activity (i.e. assuming $\gamma_j = 1$, see Sec. 2.1.1). A 1:4 host-guest binding model (derived in Sec. 2.2.2 with implementation in Appendix A.4) was adopted to obtain the equilibrium binding constants listed in Table 6.1, Sec. 6.9. The resulting fit is shown in

Fig. 6.3d (solid lines). The diprotonated form is stabilized by binding of two counteranions (on the top and on the bottom, Fig. 6.1a) at the central NH groups. Another source of stabilization is placement of the protonation at opposite carbonyl groups [36] (protonation sites can change during a tautomerization process, see Sec. 7.1.1). The good quality of the data fit also confirms the low stability of any mono- and triprotonated **OxP** species, which are essentially absent due to the ease of further protonation (analogously to tetraphenylporphyrin monocation [39, 173, 174]). The chemical structures of protonated species including the changes in conjugation are given in Fig. 6.3e. Results of UV/vis titration are confirmed by NMR, see Sec. 6.5.

6.2 Halochromism of **Bz₂OxP**

Bz₂OxP differs from **OxP** in that only single counteranion stabilization is allowed due to its double N-alkylation and consequent blocking of the top central binding site shown in Fig. 6.1b. For this reason, a stable monoprotinated **Bz₂OxP⁺** form is readily formed. UV/vis titration of **Bz₂OxP** with difluoroacetic acid (DFA)² reveals three consecutive spectral changes, Fig. 6.4a,b,c. That indicates presence of four individual absorbing species, whose spectra are identified as endpoints of the respective spectral changes, Fig. 6.4d, bold lines in Fig. 6.4a,b,c. The second spectrum was identified as **Bz₂OxP⁺**, the presence of a monoprotinated species is supported by fitting of its binding constant in both UV/vis (see below) and NMR titrations (see Sec. 6.5). Assignment of the remaining individual species as the two- and fourfold protonation products is based on their significant similarities to **z_{OxP²⁺}** and **z_{OxP⁴⁺}** (cf. Fig. 6.3c and 6.7). Similarly to **OxP**, triprotonated species is virtually absent implying strong cooperativity in the **HG₃ → HG₄** binding step. SVD analysis also confirms $N_{\text{spc}} = 4$ based on four significant singular values (see Fig. D.4) and the combination coefficient space (see Fig. 6.4f), which shows consecutive spectral changes as linear segments connecting the component coordinates. The chemical structures of the individual species and their respective colors are given in Fig. 6.4g. Analogously to the **OxP** case, linear combination of these four components accounts for all experimentally observed spectra, thus, DSIS analysis was performed using

$$\mathbf{a}_{\varphi_j} = p_{\text{Bz}_2\text{OxP}}(\varphi_j) \mathbf{z}_{\text{Bz}_2\text{OxP}} + p_{\text{Bz}_2\text{OxP}^+}(\varphi_j) \mathbf{z}_{\text{Bz}_2\text{OxP}^+} + p_{\text{Bz}_2\text{OxP}^{2+}}(\varphi_j) \mathbf{z}_{\text{Bz}_2\text{OxP}^{2+}} + p_{\text{Bz}_2\text{OxP}^{4+}}(\varphi_j) \mathbf{z}_{\text{Bz}_2\text{OxP}^{4+}} \cdot \quad (6.2)$$

Populations \mathbf{P}_{exp} obtained from DSIS are shown in Fig. 6.4e (solid circles). The reconstruction $\mathbf{ZP}_{\text{exp}}^{\text{T}}$ excellently matches the experimental spectra \mathbf{A} (see Fig. D.5) confirming that **Bz₂OxP** is also a halochromic dye.

In order to describe the interactions of **Bz₂OxP** with DFA, the first spectral change (up to $\varphi = 0.001$) was fitted using a 1:1 host-guest binding model, see Fig. 6.4e, solid lines. The quality of fit was very high, see Fig. D.5. Quantities of acid greater than $\varphi = 0.001$ change the polarity of the solvent so that activity cannot be approximated by concentration (Sec. 2.1.1). Consequently, the host-guest binding model is not applicable for higher protonated states of **Bz₂OxP** beyond monoprotination.

UV/vis titration of **Bz₂OxP** with CSA shown in Appendix D.1 revealed only monoprotinated species, higher protonation states are inaccessible due to low solubility of CSA in chloroform, see Appendix C.1. NMR spectroscopy confirms protonation at the carbonyl oxygen atoms of the hemiquinonoid groups, see Sec. 6.5. Binding constants obtained from both UV/vis and NMR titrations are listed in Table 6.1.

²DFA was used because of its nonexchanging hydrogen utilized in NMR spectroscopy analysis, which is lacking in TFA.

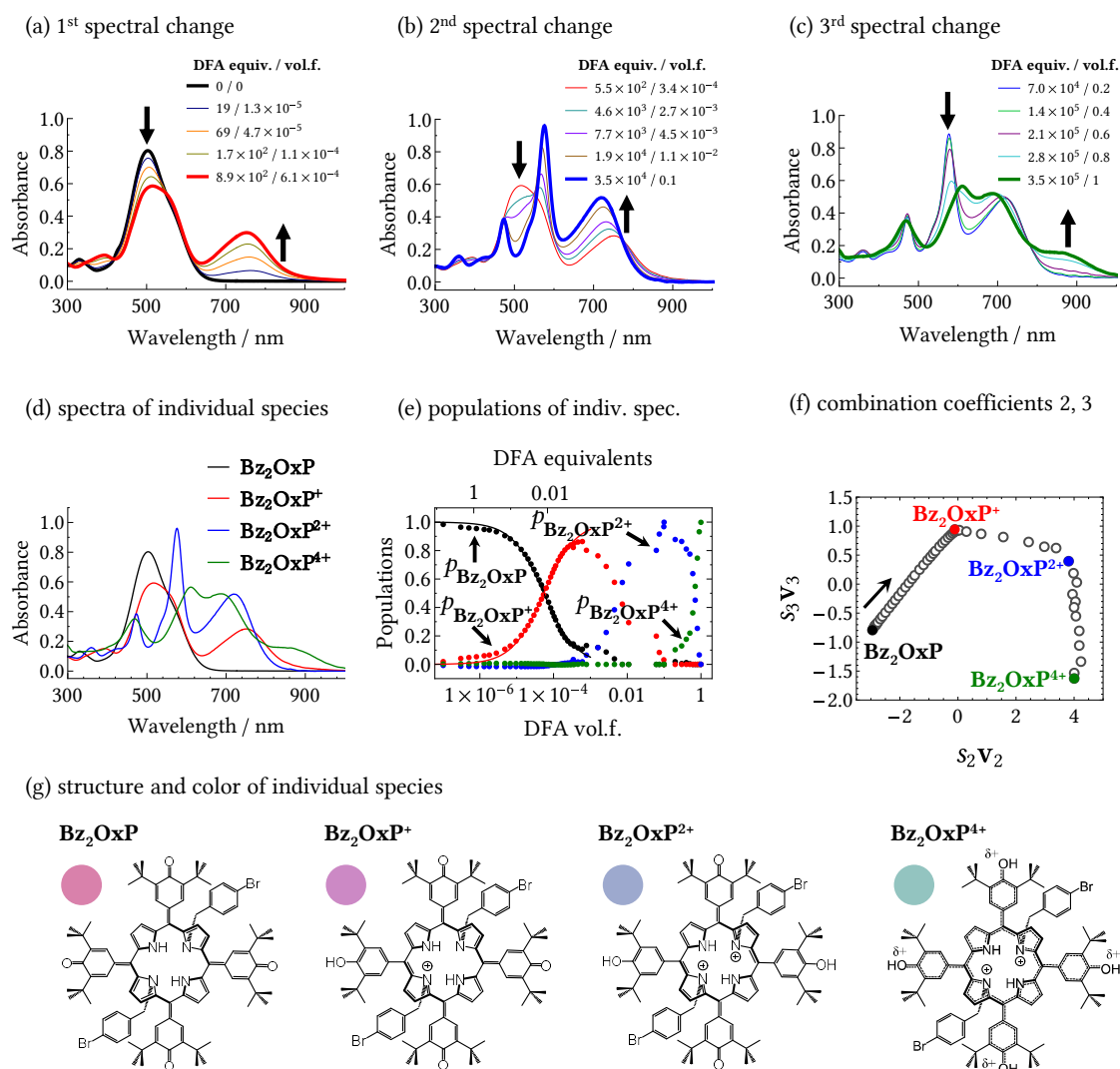


Figure 6.4: UV/vis titration of **Bz₂OxP** (7×10^{-6} M, dilution during measurement compensated by rescaling to initial dye concentration, CDCl_3) with DFA and corresponding analyses. (a, b, c) Three sets of spectral changes occurring during titration. (d) Spectra of individual absorbing species **Z** identified as the start and endpoints of each spectral change (bold lines in panels (a, b, c)). (e) Populations P_{exp} of absorbing species **Z** as obtained by DSIS (solid circles). Fitting of P_{exp} to 1:1 host-guest binding model provided model populations P_{mod} for the first protonation (solid lines). (f) SVD analysis confirms via plot of combination coefficients space, that $N_{\text{spc}} = 4$ and consecutive spectral changes include only two species at once. (g) Chemical structures and RGB colors of absorbing species (colors calculated from the spectra shown in panel (d), details in Appendix C.2).

6.3 Chromism of **Bz₄OxP**

The title of this section avoids term ‘halochromism’ on purpose because **Bz₄OxP** shows both halo- and solvatochromism. Of the compounds studied, this derivative is the least susceptible to protonation due to complete N-alkylation and consequent blockage of both central binding sites, Fig. 6.1c. No stabilizing cation-counteranion interactions are allowed.

UV/vis titration with DFA shows three sets of spectral changes. The first one (up to $\varphi = 0.01$) is a red shift of absorbance maximum (500 nm), i.e. positive solvatochromism (see Sec. 4.5), see Fig. 6.5a (green arrow). This shift was not observed in the previous cases and is attributed to

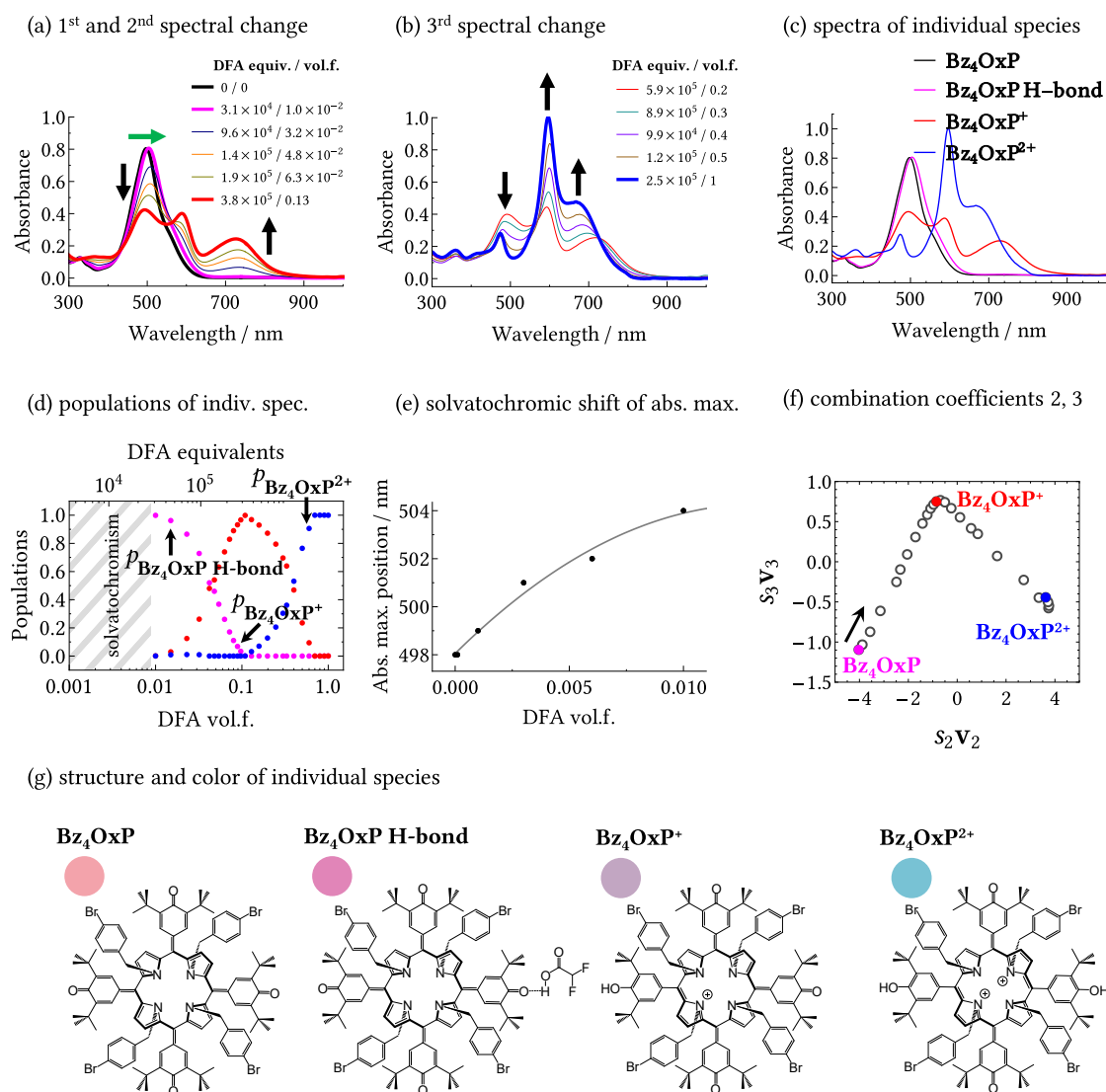


Figure 6.5: UV/vis titration of **Bz₄OxP** (5×10^{-6} M, measured spectra rescaled to initial dye concentration, CDCl_3) with DFA and corresponding analyses. (a) First solvatochromic (green arrow) and second (black arrow) sets of spectral changes occurring during titration, (b) third spectral change. (c) Spectra of individual absorbing species **Z** identified as the start and endpoints of each spectral change (bold lines in panels (a, b)). (d) Populations P_{exp} of absorbing species **Z** (free **Bz₄OxP** excluded due to solvatochromic shift) as obtained by DSIS. (e) Shift of absorbance maxima, gray line is eyeguide. (f) SVD analysis confirms $N_{\text{spc}} = 3$ and consecutive spectral change via plot of combination coefficients space. (g) Chemical structures and RGB colors of absorbing species (colors calculated from the spectra shown in panel (c), details in Appendix C.2).

hydrogen bonding of acid molecules to the carbonyl oxygen atoms. It was shown in [34] that hydrogen bonding with water (in tetrahydrofuran as solvent) was accompanied by a similar spectral shift in the case of both di- and tetra-*N*-substituted oxoporphyrinogens. The solvatochromic effect is weak since the shift of the absorbance maximum is less than 10 nm as shown in Fig. 6.5e. The second and third spectral changes are of typical halochromic nature, see Fig. 6.5a,b (black arrows). Again, the corresponding spectra of individual species were identified as endpoints of the spectral changes, see Fig. 6.5c. Comparing to the individual species of other oxoporphyrinogens, we assigned the last two absorbing components as mono- and diprotonated forms (again proto-

nated at the hemiquinonoid carbonyl sites) as shown in Fig. 6.7b,c. SVD analysis on the dataset excluding initial solvatochromic shift (for $\varphi > 0.01$) confirms consecutive interchange between three components via singular values shown in Fig. D.6a and plot of combination coefficient space shown in Fig. 6.5f.

The same halochromic dataset (as in the case of SVD) was used for DSIS, decomposing the measured spectra into

$$\mathbf{a}_{\varphi_j} = p_{\text{Bz}_4\text{OxP H-bond}}(\varphi_j) \mathbf{z}_{\text{Bz}_4\text{OxP H-bond}} + p_{\text{Bz}_4\text{OxP}^+}(\varphi_j) \mathbf{z}_{\text{Bz}_4\text{OxP}^+} + p_{\text{Bz}_4\text{OxP}^{2+}}(\varphi_j) \mathbf{z}_{\text{Bz}_4\text{OxP}^{2+}}. \quad (6.3)$$

Populations of the halochromic species obtained from DSIS are shown in Fig. 6.5d. As in the previous cases, the quality of DSIS is excellent, see Fig. D.7. Populations \mathbf{P}_{exp} from DSIS cannot be fitted by a simple host-guest binding model, as the high concentration of acid ($\varphi > 0.01$) causes the activity cannot be approximated by concentrations any more.

On the other hand, dye concentrations used in NMR titration are higher than in UV/vis and the first protonation could be achieved at lower DFA vol.f. This enables fitting of the first protonation by 1:1 host-guest binding model presented in Table 6.1. The NMR titration also reveals a significant shift in the pyrrole β -proton resonance, Fig. 6.12c and E.7, which strongly suggests that upon protonation the positive charges are accommodated at the central nitrogen atoms of **Bz₄OxP** as shown in Fig. 6.5g.

6.4 Mechanism of halochromism

Most simple dye systems operate solely on the basis of protonation, or solvent polarity. Molecular design allows for introduction of some complexing moiety to introduce specificity of the analyte-dye interaction [142, 143].

In this context, halochromism of the oxoporphyrinogen derivatives, investigated here, is strongly affected by the presence of central calix[4]pyrrole-type binding site.

The conjugation pathway of **OxP** can be varied upon protonation so the charge is shared over electronegative heteroatoms (i.e. N and O). After double protonation, even reorganization of the chromophore structure through tautomerization is possible. In fact, delocalization of charge may be one of the factors allowing the existence of the highly charged tetracation. The effect of N-alkylation of **OxP** is twofold: first it masks the central binding site(s), which prevents counteranion binding; second, it sterically restricts redistribution of the conjugated structure of the chromophore, which might otherwise further stabilize protonated species.

As shown in Fig. 6.6, the binding site blocking shifts approximate lower threshold of colorimetric response from $\varphi = 10^{-6}$ for **OxP** to $\varphi = 10^{-3}$ for **Bz₂OxP** and further to $\varphi = 0.01$ for **Bz₄OxP**. Such big change in sensitivity is caused by removing not only one,

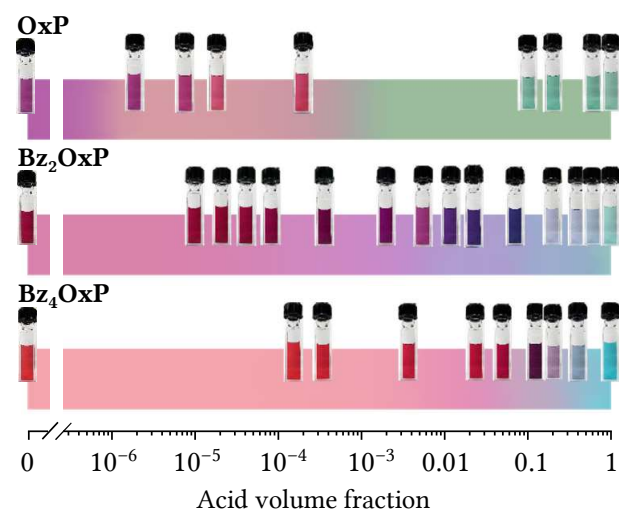


Figure 6.6: Colorimetric response of **OxP** (with TFA), **Bz₂OxP**, and **Bz₄OxP** (with DFA) over the entire concentration range of the acids. Colors shown were obtained by interpolation of RGB values (see Appendix C.3) obtained from the actual measured spectra, and the color scales are overlaid with photographs of the sample cuvettes (from the titration experiments) at the appropriate vol.f.

Such big change in sensitivity is caused by removing not only one,

but two sources of stabilization, as mentioned above. Also the acid vol.f., at which all spectral changes saturate, depends strongly on the N-alkylation. The population plots from DSIS in previous sections show, that **OxP** protonation saturates approximately at $\varphi = 0.002$, for **Bz₂OxP** at $\varphi = 1$ and for **Bz₄OxP** at $\varphi = 0.7$ (only double protonation).

Spectra of the corresponding mono-, di- and tetraprotonated species of each **OxP** derivative are similar in form as shown in Fig. 6.7. This assumption was exploited to assign the degrees of protonation of each absorbing species. Diprotonated species, present in each derivative, show gradual shifts of absorption bands correlated with the degree of N-alkylation (red arrows in Fig. 6.7c).

There exists strong positive cooperativity between the first and second protonation steps in **OxP** because of the stabilization by two counterions (Table 6.1, similar to the case for double protonation of free-base porphyrins [173]). Therefore, monoprotinated spectrum of **OxP** is not observable. Since **Bz₂OxP** is stabilized by only one counteranion, its monoprotinated form is stable. Surprisingly, another strong cooperativity arises between the third and fourth protonation step in both **OxP** and **Bz₂OxP**. As a result, their triprotonated form cannot be observed either.

Bz₄OxP lacks further counterion stabilization, therefore it does not show any strongly cooperative protonation. In addition, its steric crowding at the core obstructs redistribution of charge, therefore **Bz₄OxP** is not completely protonated even in neat acid.

In Fig. 6.3e, 6.4g and 6.5g we have assigned the formal placement of charge at the central nitrogen atoms of the protonated structures. These charges may be shared over all electronegative atoms, albeit with preference for the central pyrrole nitrogens. The hypothesis of charged tetrapyrrole center is supported by significant shifts of pyrrole β -protons resonances in NMR titrations to higher ppm discussed in the next section (green squares in Fig. 6.12).

Based on the formal conjugated structure, charge redistribution can be depicted. Fig. 6.8a shows tautomerization of **OxP²⁺**, Fig. 6.8b clarifies transfer of charge to the molecular center of **Bz₂OxP⁺** upon protonation and Fig. 6.8c illustrates variants of charge placement at the pyrrole nitrogens of **Bz₂OxP²⁺**. The protonation sites dynamically change by a prototropic tautomerization process, which is described in detail in the next chapter in Sec. 7.1.1 and 7.2.3.

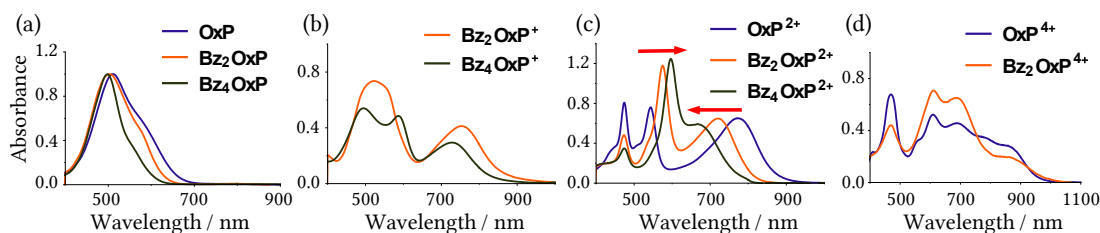
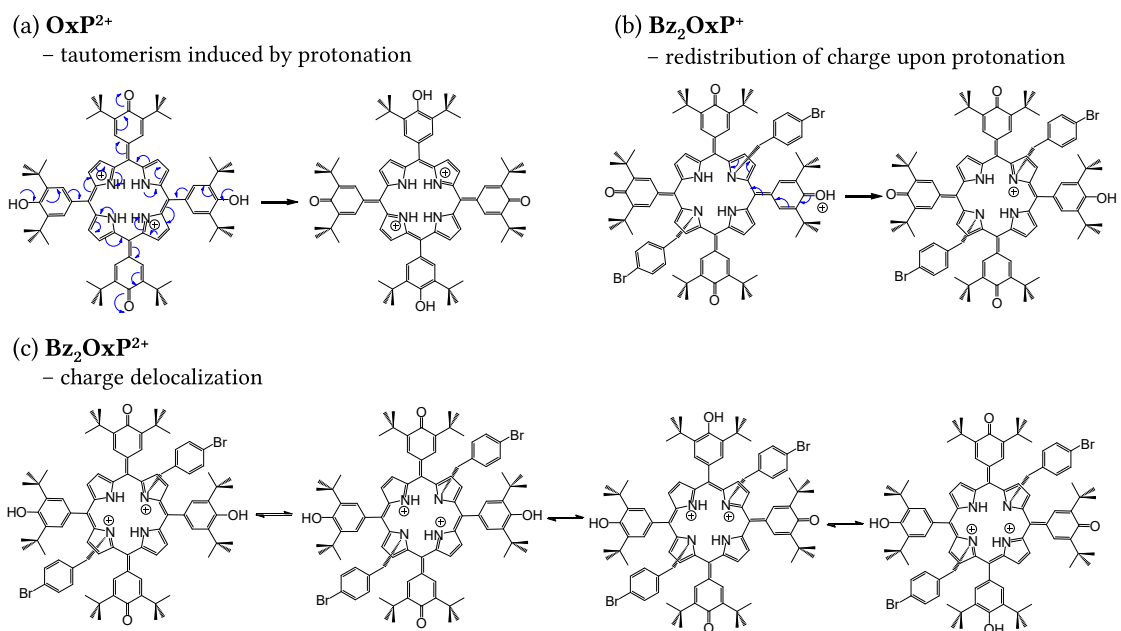


Figure 6.7: Assignment and spectra of neutral and protonated species of **OxP**, **Bz₂OxP** and **Bz₄OxP** in the (a) neutral state and in the (b) monoprotinated, (c) diprotonated and (d) tetraprotonated forms. Spectra of neutral species were normalized to unit absorbance at maximum, protonated species are to scale with the neutral ones.

Figure 6.8: Redistribution of charge upon protonation for (a) OxP and (b,c) Bz_2OxP .

6.5 NMR analysis of host-guest binding

NMR titration experiments are another tool for assessing binding properties of oxoporphyrinogens. Halochromic properties of the investigated oxoporphyrinogen series were tested with DFA as well as with CSA in deuterated chloroform. DFA (liquid) is miscible with CDCl_3 in arbitrary ratio, however the CSA (powder) has limited solubility (Table C.1). Therefore, high degree of protonation can be achieved only at excess of DFA, not in the presence of CSA. Tetramethylsilane (TMS) was used as internal standard. Addition of maximal soluble amount of CSA into solution of **Bz₄OxP** did not result in any spectral change, therefore, these experiments are not shown. Experimental details are given in Appendix C.4, data analysis procedure is described in Appendix C.5. Experimental details about the measurements are given in Appendix C.4.

Assignment of ^1H spectra of the investigated **OxP** derivatives is shown in Fig. 6.9. Signals due to *tert*-butyl-H are denoted by star (★), due to *ortho*-H by circle (●), due to β -H by square (■) and signals of the bromobenzyl groups by triangles (▲). Assignment of **OxP** and **Bz₂OxP** resonances was published in [38] and [40], respectively, assignment of **Bz₄OxP** was conducted with help of HMBC measurement, see Appendix E.18. Despite high number of hydrogen atoms, the resulting spectra contain low number of signals due to symmetry. Spectra of pure acids can be found in Fig. E.1 in Appendix E.

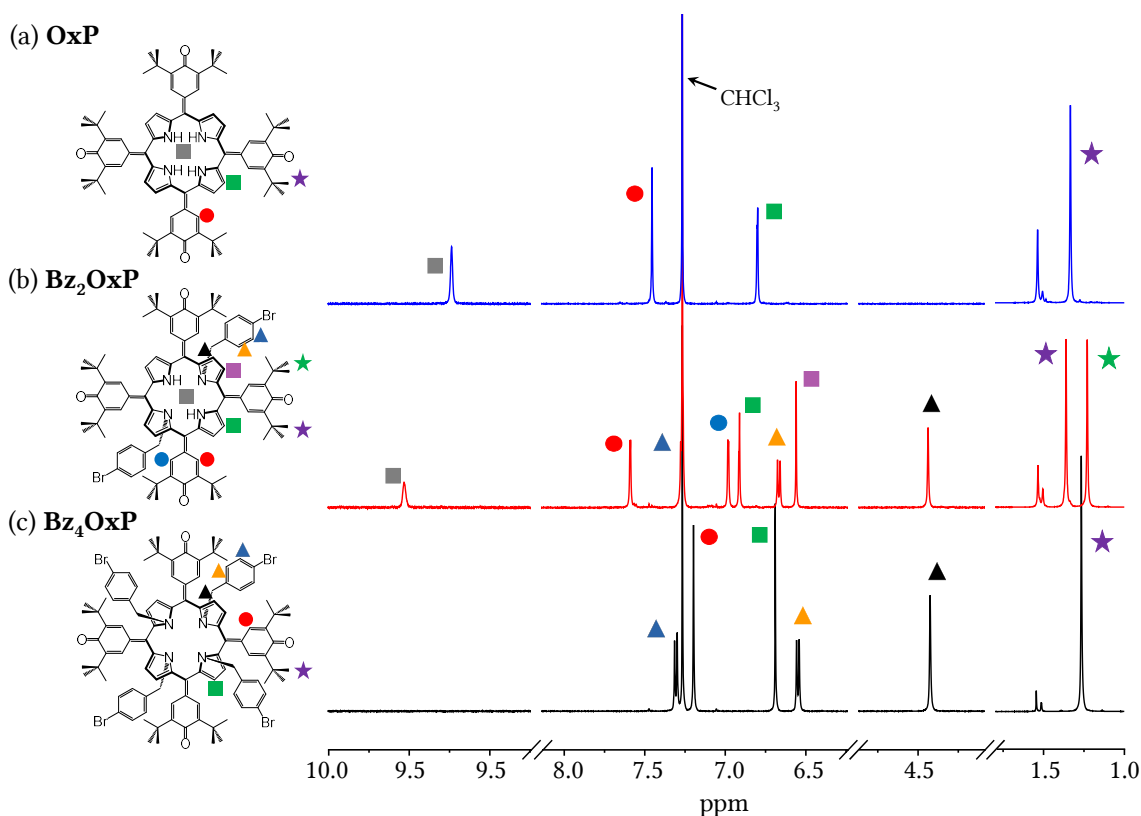


Figure 6.9: Assignment of NMR signals of investigated oxoporphyrinogens without the presence of acid.

NMR chemical shifts of the investigated oxoporphyrinogens are influenced by protonation with consequent charge transfer, by acid anion binding and also by the change of solvent polarity at high acid vol.f. [141]. In the case of **OxP**, new isolated peaks of bound/protonated species arise upon acid addition (both DFA and CSA) due to slow exchange regime. Also the host OH signals from protonation can be observed as shown in Fig. 6.10a (top). Relative area of this resonance,

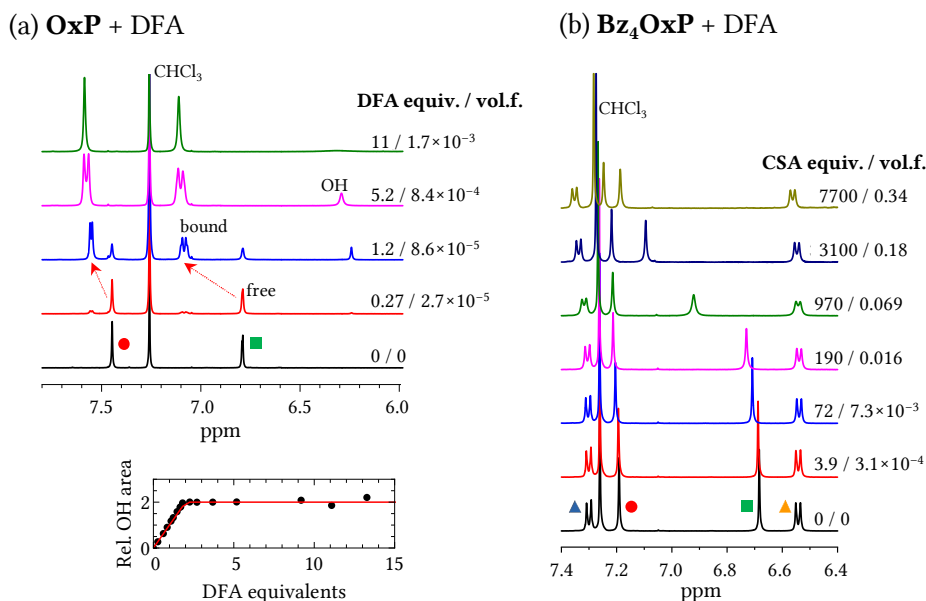


Figure 6.10: Illustration of spectral changes during titration experiments. (a) Titration of OxP with DFA. The spectra (top) show separate peaks of free and bound forms due to *ortho*-H (●) and β -H (■) spins as well as the host OH resonances of protonated host. Relative area of the OH resonances (bottom) normalized to the total area of *tert*-butyl-H peaks, fitted with the function $([HG] + 2[HG_2])/[H]_t$ and yielding parameters $K_1 = 4 \times 10^6$, $K_2 = 8.4 \times 10^6$, compatible with results in Table 6.1. (b) Titration of Bz₄OxP with DFA showing significant shift of the β -H resonance (■).

shown in Fig. 6.10a (bottom), is saturated at the value of 2, which is a direct evidence for two-fold protonation. The assignment to host OH resonance was confirmed by addition of heavy water after which the resonance disappeared, see Appendix E.17. Full titration spectra of OxP with DFA or CSA are shown in Fig. E.2 and E.3 in Appendix E, respectively.

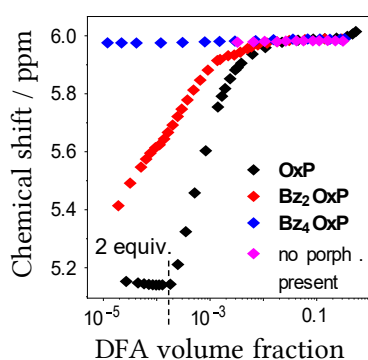


Figure 6.11: Chemical shift of the DFA non-exchanging proton signal (central frequency of the triplet) in the presence and absence of the oxoporphyrinogens investigated for different vol. frac. of DFA.

are in fast exchange between their free and bound states, thus, only the averaged triplet signal is observed. During titration of OxP all acid anions are bound to the oxoporphyrinogen up to 2 equiv. due to the strong binding affinity (◆). For Bz₂OxP, anion binding is not so strong, so

On the other hand, Bz₂OxP does not show new separate signals upon protonation due to intermediate/fast exchange of free and bound species. Only the NH signals are an exception and signals due to free and bound species are separated, see Figure 7.12a in the next chapter. We found a candidate OH signal for direct observation of the protonation, see Appendix E.17.

Titration spectra of Bz₄OxP do not show any formation of separate peaks. The only spectral changes are shifts of *tert*-butyl-H and β -H signals. The significant shift of β -H resonance (■) shown in Fig. 6.10b indicates charge transfer from carbonyl groups to the porphyrinogen center enabled by protonation.

Chemical shift of the center of DFA triplet is plotted in Fig. 6.11 as a function of DFA concentration. The frequency shift reflects acid anion binding rather than protonation of the oxoporphyrinogens. The DFA anions

the corresponding chemical shift is not constant (\blacklozenge). In the case of **Bz₄OxP**, chemical shift of its acid triplet (\blacklozenge) is the same as if no oxoporphyrinogen was present (\blacklozenge). It confirms the lack of acid anion binding to **Bz₄OxP**.

For each titration experiment, binding isotherms were constructed from integrated peak areas (slow exchange) and peak shifts (fast exchange) as described in Sec. 3.4.2. All binding isotherms were simultaneously fitted with host-guest binding model of appropriate stoichiometry (1:4 for **OxP**, 1:1 for **Bz₂OxP** and **Bz₄OxP**), the resulting binding constants are listed in Table 6.1. The binding models were not applicable at acid vol. frac. $\varphi > 0.001$ since the change of solvent polarity, caused by addition of acid, influenced the peak shifts [141]. Moreover, for the same reason, activity cannot be approximated by concentration at high acid concentrations, see Sec. 2.1.1. Relative chemical shifts (i.e. with subtracted initial value) of non-overlapping signals of the oxoporphyrinogens as function of DFA vol.f. are shown in Fig. 6.12. Some of the peak shifts were used as binding isotherms in the fitting procedure, the fitting curves are also shown (solid lines). Of the oxoporphyrinogen peaks, the most significant spectral shifts belong to the pyrrole β -H and NH resonances, supporting the hypothesis of charge localization at the pyrrole nitrogen atoms. Note that even residual solvent peak (black asterisk * in Fig. 6.12) is shifted relative to TMS at high acid. Details about the current data analysis, including peak integration

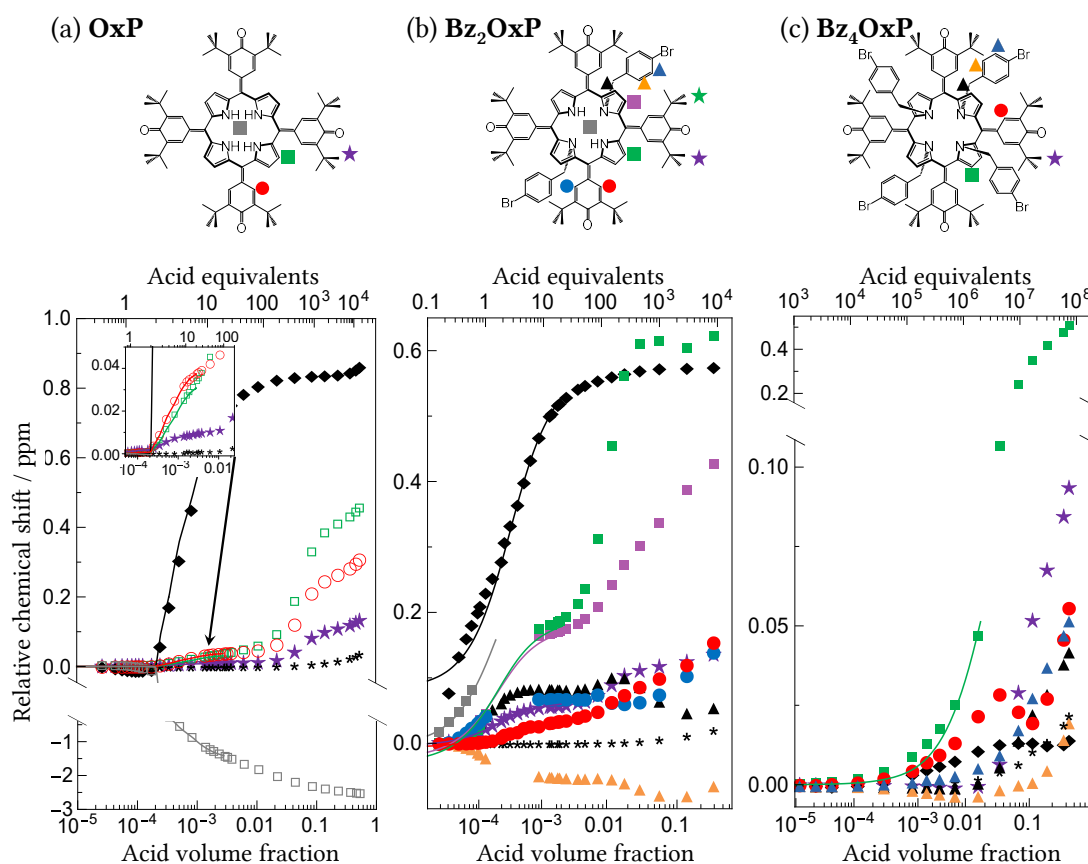


Figure 6.12: Relative chemical shifts of peaks of the investigated oxoporphyrinogens during titrations with DFA. The value of chemical shift of particular resonance at 0 equiv. of DFA is subtracted from corresponding resonance during the titration process. For two distinct tert-butyl peaks or any doublets central chemical shift is displayed as they shift simultaneously. Non-exchangeable DFA resonance (triplet) is denoted by the diamond symbol (\blacklozenge), CHCl_3 is denoted by asterisk (*). Solid lines stand for fits, the corresponding binding constants are in agreement with Table 6.1.

using lineshape fitting, fitting of binding isotherms and the implementation for OxP + (R)-CSA system, are given in Appendix C.5.

6.6 Solvatochromism of OxP

To examine solvatochromic properties (see Sec. 4.5) of OxP and its derivatives, we investigated their behavior in solvents of different polarity, namely in dimethylformamide/chloroform binary mixtures, see Appendix C.2 for experimental details. Increase in concentration of the polar dimethylformamide (DMF) resulted in solvatochromic red shift caused by hydrogen bonding.

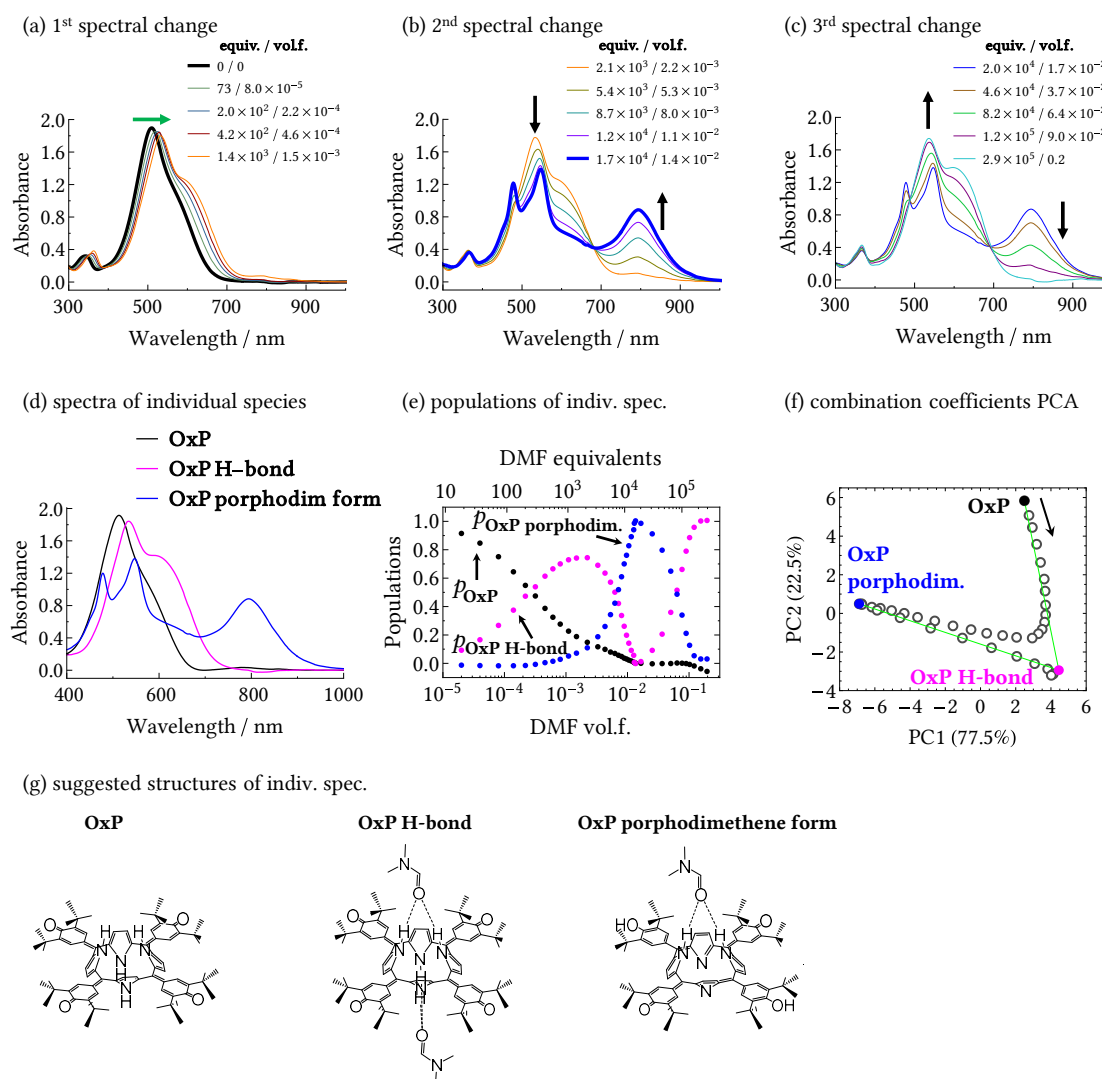


Figure 6.13: Measurement of OxP (1.7×10^{-5} M, spectra rescaled to initial dye concentration to compensate for dilution during experiment, CDCl_3) with DMF and corresponding analyses. (a,b,c) Three sets of spectral changes occurring during titration. Starting point of the first change and endpoint of the second change are identified as individual absorbing species. (d) Individual spectra Z and (e) their populations P_{exp} obtained from coordinates T in combination coefficients space in frame of 3D SVD approximation. (f) PCA of combination coefficients, colored points denote coordinates of absorbing components, intersection of green lines helped to identify the coordinate of OxP H-bond. (g) Suggested structures of the absorbing components.

In the case of **OxP**, unexpected rearrangement of its conjugated system was observed for particular DMF concentration range. Hydrogen bonding between oxoporphyrinogens and solvent molecules has been reported, including carbonyl-containing solvents [30] (UV/vis spectra, X-ray crystallographic structure of **Bz₂OxP** · acetone) and trace water in tetrahydrofuran [34] (UV/vis titration with water, DFT calculations of **Bz₂OxP** · *n*H₂O).

As shown in Fig. 6.13a,b,c, titration of **OxP** with DMF resulted in three distinct spectral changes. The first one involves a positive solvatochromic shift (the original electronic structure does not change), identified as hydrogen bonding with either DMF or water, which is also present in the sample (amount of water increased during titration, see NMR spectra in Fig. E.16). Hydrogen bonding of water to **OxP** induces similar spectral change [34] shown in Fig. 6.14a. The other two spectral changes imply formation of new species with different electronic structure since new absorption bands arise.

In our publication [1] we identified the spectra of absorbing components as endpoints of the three spectral changes (implying four species) and consequently performed DSIS analysis. However, in this work we revisited these results and provide more precise analysis based on combination coefficients. As shown in the SVD analysis in Appendix D.5, the first spectral change, although solvatochromic, can be described as linear combination of two spectra of individual species, thus, for the whole experiment it holds $\mathbf{A} = \mathbf{ZP}^T$ (see Sec. 5.2.1). Singular values indicate the presence of three components, also the first three combination coefficients lie on a 2D plane in their 3D space as confirmed by principal component analysis (PCA, explained in Sec. 5.3), see Fig. D.8 (N_{spc} components form an $(N_{\text{spc}} - 1)$ -dimensional object in the combination coefficients space, although small overlap with other dimensions occurs due to the noise). These are sufficient arguments to prove the presence of exactly three absorbing species.³ Mutual plot of combination coefficients PCs (principal components) in Fig. 6.13f shows that the line defined by the first spectral change intersects the last points of the third spectral change (first and third spectral changes are denoted by green lines). Hence, before the hydrogen bonding in the first spectral change could be saturated, there is a “digression” in the combination coefficients plot identified clearly as formation of a third absorbing component, which looks very distinct from both **OxP** and **OxP H-bond**. Coordinates of **OxP** and this third species were deduced directly from the measured spectra (bold lines in Fig. 6.13a,b) and co-

³In our study [1], we concluded the presence of four absorbing components, two of them almost identical, however, the simpler variant in this work seems more plausible.

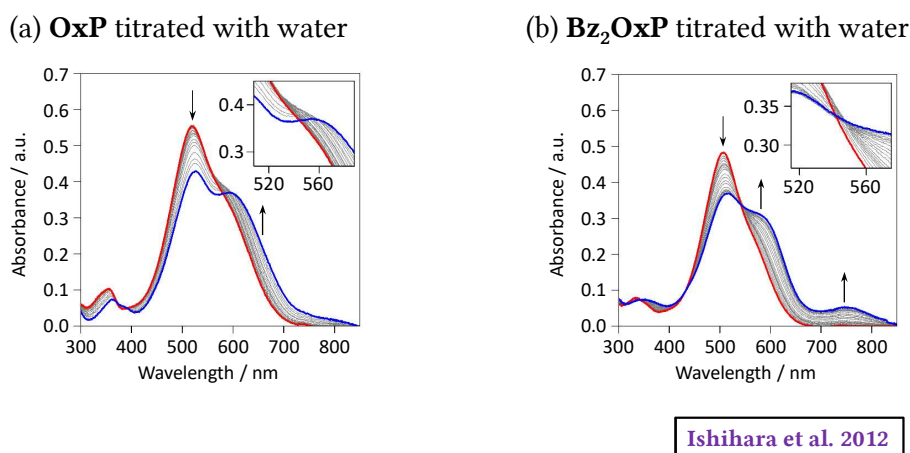


Figure 6.14: UV/vis titration of (a) **OxP** and (b) **Bz₂OxP** with water in tetrahydrofuran solvent. Adopted from [34].

ordinates of **OxP H-bond** were determined as intersection of the green lines in combination coefficients PCA plot. After obtaining the coordinates, individual spectra were calculated (using Eq. (5.21)) in frame of 3D SVD approximation (capturing 99.98% of the signal) and together with their populations plotted in Fig. 6.13d,e. Details of this analysis are given in Appendix D.5.

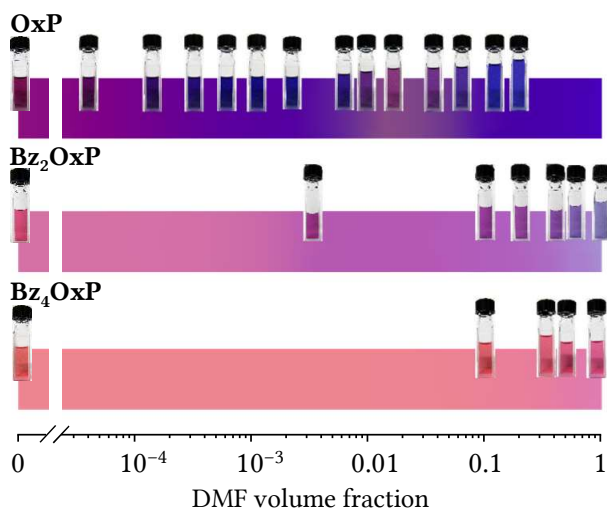


Figure 6.15: Colorimetric response of **OxP**, **Bz₂OxP**, and **Bz₄OxP** over the entire concentration range of DMF/ CDCl_3 mixture. Colors shown were obtained by interpolation of RGB values (see Appendix C.3) obtained from the actual measured spectra, and the color scales are overlaid with photographs of the sample cuvettes (from the titration experiments) at the appropriate vol.f.

Electronic structure of the third absorbing species implies big reorganization of the conjugated system.⁴ Therefore its structure was assigned to a porphodimethene form (cf. Fig. 1.1), **OxP porphodimethene form** in Fig. 6.13g, based on significant similarity to the spectrum of tautomeric **OxP²⁺**. Prototropic tautomerization of the porphodimethene form could be assisted by DMF acting as proton acceptor (due to its basic properties) and water as proton donor. Stabilization of **OxP porphodimethene form** depends on solvent polarity (potentially on the balance of polarity and hydrogen bonding) as it is prevalent only in DMF vol.f. range of $\varphi \in [0.01, 0.05]$. Phenol groups of the potential tautomer might as well form hydrogen bonds with DMF, although this interaction ought to be weakened by steric effects of the tert-butyl groups.

NMR titration of **OxP** by DMF shown in Fig. E.16 reveals no conclusive results, however, it indicates interaction of **OxP** with DMF as their peaks are slightly shifted even at low DMF concentration.

Figure 6.15 provides overview of the colorimetric response of all three oxoporphyrinogens. The vol.f. range of stabilized **OxP porphodimethene form** is clearly visible with naked eye. Clearly, there is no such effect in the case of the N-alkylated oxoporphyrinogen derivatives. Note that the color change in **OxP** caused by DMF was saturated at $\varphi(\text{DMF}) < 0.2$, while for **Bz₂OxP** and **Bz₄OxP** the color change was not saturated even in pure DMF.

6.7 Solvatochromism of **Bz₂OxP**

Measurements of **Bz₂OxP** in DMF/ CDCl_3 at different vol.f. shows two different solvatochromic changes, see Fig. 6.16a, the first one involving only small shift of the original absorbing maxima (green arrow), the second including variation of the original band intensities and formation of another absorbing band at ca. 750 nm (black arrows). SVD analysis indicates presence of three absorbing components (99.99% of signal, see Appendix D.6), which were identified from the PCs of combination coefficients in Fig. 6.16d. The first component **Bz₂OxP** was measured in isolation as the first measured spectrum (in pure chloroform), the second was identified as intersection of the orange lines in Fig. 6.16d delineating first and second spectral change. However, placement of third component on the orange line is not unique as the second spectral change is probably

⁴Precautions were taken to exclude acid formation in the solvents, potentially causing such spectral change by protonation, see Appendix C.2.

not saturated even at $\varphi(\text{DMF}) = 1$. Therefore, we identified range of possible coordinates as a line segment bounded by the last measured spectrum at one side and the position of last reasonable spectral shape (transparent green line segment in Fig. 6.16d). The green point in the same figure denotes coordinates of a representative spectral shape in the middle of the allowed range. Figures 6.16b,c show the absorbing spectra and the their populations during the measurements. Green solid line (or green points) correspond to the green point in (d), green shaded area corresponds to the uncertainty denoted by the transparent line segment in (d). Negative values of the individual spectrum at ca. 700 nm and negative populations at ca. $\varphi(\text{DMF}) = 0.4$ are caused by non-linear effects in the UV/vis data. Details of this analysis are given in Appendix D.6.

The current spectral changes highly resemble those of hydrogen bonding of **Bz₂OxP** with water in tetrahydrofuran reported in [34], see Fig. 6.14b. In the same article, DFT calculations were conducted, revealing different effects of simple hydrogen bonding at single carbonyl (**Bz₂OxP** · H₂O) compared to multiple carbonyl hydrogen bonding or bonding at the central NH group (**Bz₂OxP** · 4H₂O, **Bz₂OxP** · 5H₂O), see Fig. 6.17. Thus, we assign the second absorbing component to **Bz₂OxP** interacting with DMF in unspecified way (possibly hydrogen bonding at the host carbonyl sites or only by increasing solvent polarity) and denote it as **Bz₂OxP form 1**. The third absorbing component is assumed to bind DMF at the poprhyrinogen center based on the comparison with Fig. 6.14b (from [34]) and is denoted as **Bz₂OxP form 2**. The suggested structures are shown in Fig. 6.16e. Colorimetric response of **Bz₂OxP** to the solvent is shown in Fig. 6.15. The color variation is shifted to higher DMF concentration compared to **OxP**.

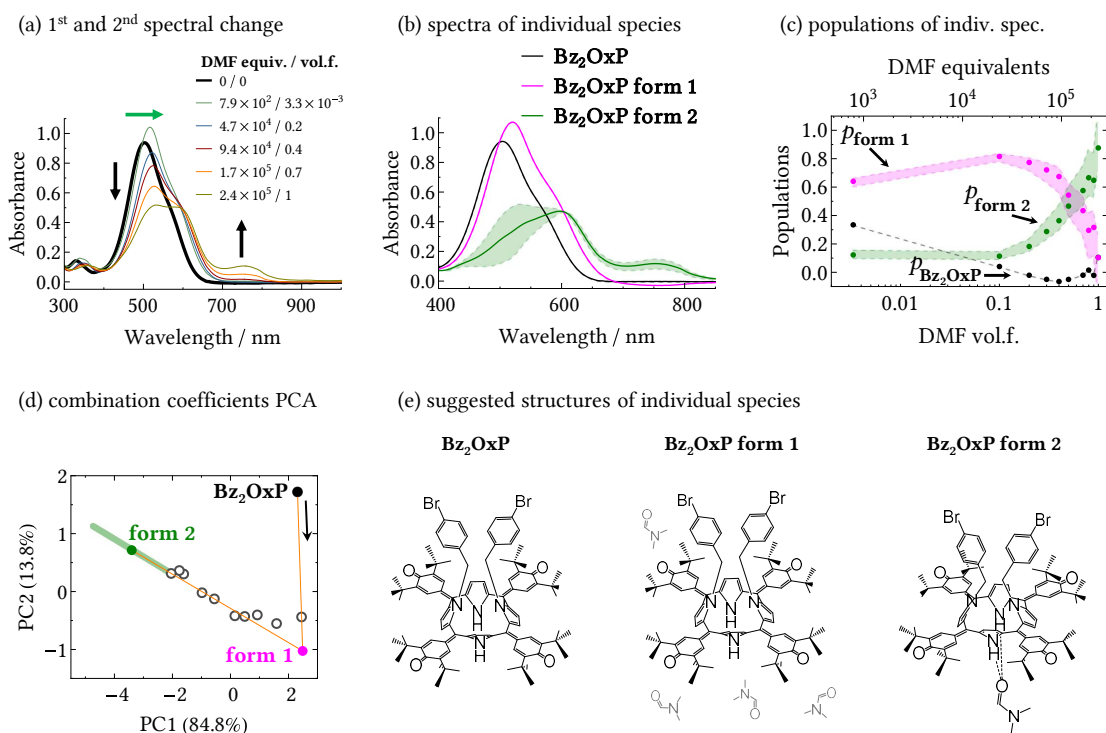


Figure 6.16: UV/vis measurements of **Bz₂OxP** (5.5×10^{-5} M) in DMF/ CDCl_3 mixture and corresponding analyses. (a) Two sets of spectral changes occurring at different DMF vol.f. Starting point of the first change is identified as individual absorbing species (bold line). (b) Individual spectra **Z** and (c) their populations \mathbf{P}_{exp} obtained from coordinates **T** in combination coefficients space in frame of 3D SVD approximation. (d) PCA of combination coefficients, colored points denote coordinates of absorbing components. Intersection of orange lines helped to identify the coordinate of **Bz₂OxP form 1**. (e) Suggested structures of absorbing components.

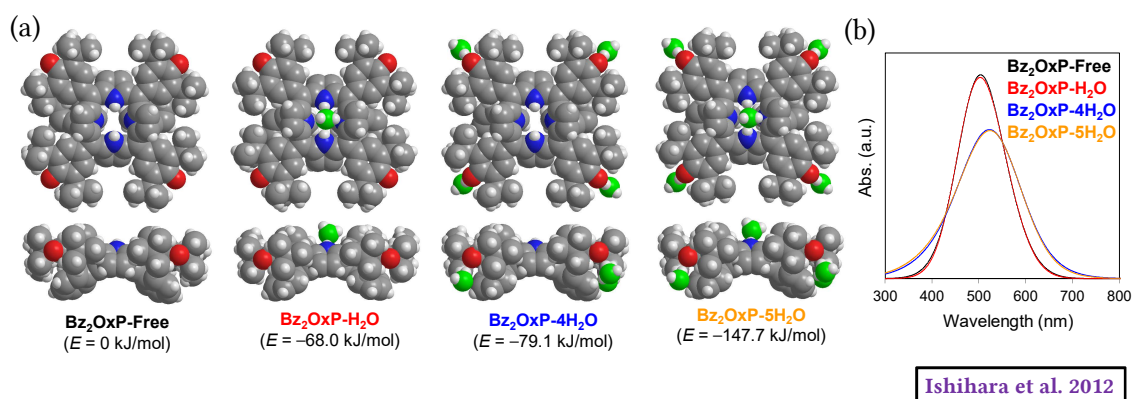


Figure 6.17: DFT calculations of **Bz₂OxP**. Bromine atoms at benzyl substituents were replaced with hydrogen atoms to simplify calculations. (a) Top and side views of calculated hydrogen bonded structures of **Bz₂OxP** (M06-L/6-31G(d,p), 0 K in vacuum). Total binding energies of water molecules were shown. (b) Calculated absorption spectra (TD-M06-2x/6-31++G(d,p)). Adopted from [34].

6.8 Solvatochromism of **Bz₄OxP**

Similarly to solvatochromism in DFA, **Bz₄OxP** shows small solvatochromic shift in DMF, see Fig. 6.18a. Compared to DFA, the red shift caused by DMF requires much higher concentration of the added polar/hydrogen bonding substance as shown in Fig. 6.18b. Due to the small magnitude of the solvatochromic spectral change (≈ 10 nm), SVD cannot unequivocally determine if the spectral change is caused by formation of a new absorbing species. 2D SVD approximation covers 99.994% of the signal suggesting transition between two species, however, combination coefficients shown in Fig. 6.18c do not lie on a straight line as required.⁵ This behavior, if caused by individual absorbing species, requires three components. This hypothetical scenario is further analyzed in Appendix D.7. Colorimetric response of **Bz₄OxP** to the solvent is shown in Fig. 6.15. The color variation is weak and shifted to higher DMF vol.f. compared to **OxP** or **Bz₂OxP**.

⁵Such deviation might be caused by variation in total dye concentration. However, in this experiment, stock solution with DMF contained the same concentration of **Bz₄OxP** as the measured solution, which should prevent any concentration fluctuations.

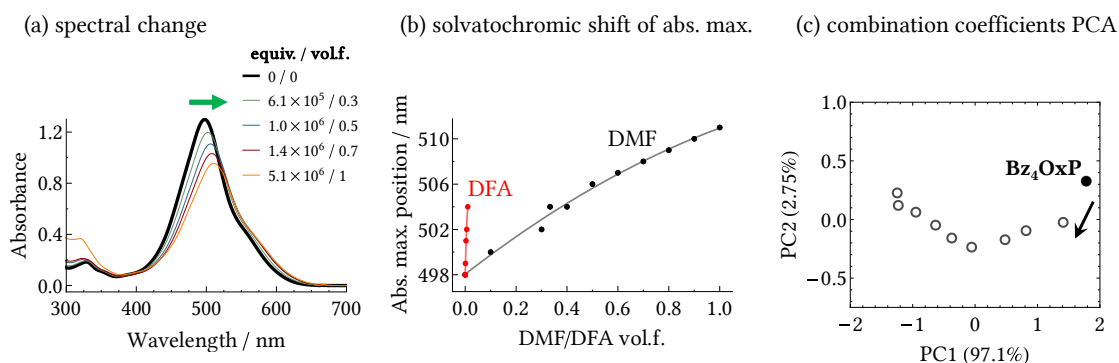


Figure 6.18: UV/vis measurements of **Bz₄OxP** (6.3×10^{-5} M) in DMF/ CDCl_3 mixture and corresponding analyses. (a) The only spectral change occurring at different DMF vol.f. The starting point (in pure chloroform) is identified as an individual absorbing species (bold line). (b) Solvatochromic shift of absorbance maxima induced by DMF (black) or DFA (red, replicated from Fig. 6.5e), solid lines are eyeguides. (c) PCA of combination coefficients.

6.9 Chromism and binding – summary

The investigated oxoporphyrinogens are halochromic dyes operating as colorimetric indicators over broad range of acid contents in nonpolar media ($\varphi(\text{DFA}) \in [10^{-6}, 1]$). This was permitted by up to fourfold protonation of the **OxP** derivatives. The maximum protonation level was tuned by N-alkylation, which gradually shifted the sensitivity, i.e. colorimetric response, to higher acid contents, as protonation of the molecules was increasingly obstructed in the series **OxP**, **Bz₂OxP** and **Bz₄OxP**. The mechanism of operation involves stabilization by counteranion binding and charge delocalization. This stabilization induces strong cooperativity (see Sec. 2.2.4) in the binding step **HG** → **HG₂** (described by the interaction parameter α_{12}). Surprisingly, also the binding step **HG₃** → **HG₄** (α_{34}) is strongly cooperative in **OxP** and **Bz₂OxP** due to unknown reason. For **Bz₄OxP**, stabilization by a counteranion is not permitted, thus, its response is limited to higher acid contents ($\varphi(\text{DFA}) > 0.1$). Schematic overview of the protonated species and the effect of binding site blocking is shown in Fig. 6.19. Note the difference between weak hydrogen bonding of acid to **Bz₄OxP** at low acid contents and protonation of **Bz₄OxP** when more acid is added. UV/vis and NMR titration experiments allowed fitting of binding isotherms (at $\varphi(\text{acid}) < 0.001$) to obtain the corresponding binding constants. At higher acid vol.f., the binding constants could not be obtained because of the unknown activity coefficients of the species in solution (see Sec. 2.1.2) and peak shifts due to increased solvent polarity in NMR measurements. Also, full protonation could not be achieved in CSA due to its limited solubility in chloroform (see Appendix C.1), hence, CSA induces only monoprotation⁶ in **Bz₂OxP** and even no protonation

⁶Peak shifts in NMR titration above ca. 10 CSA equiv. are induced by the increase of solvent polarity, although they might also be attributed to formation of **HG₂** complex. However, UV/vis titration of **Bz₂OxP** with CSA proved

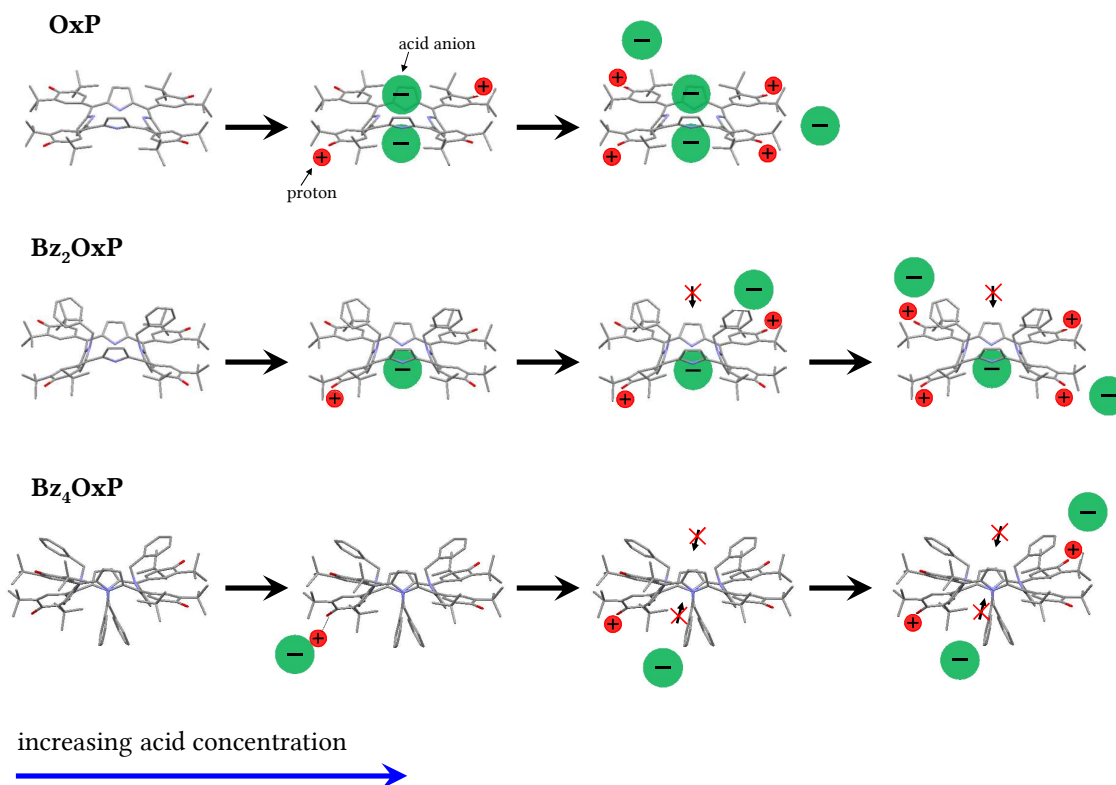


Figure 6.19: Binding site blocking effect of the oxoporphyrinogen molecules on their interaction with acid. Crossed arrows denote blocked sites due to N-alkylation.

Table 6.1: Stepwise binding constants^a (in M^{-1}) and corresponding interaction parameters^b (measure of cooperativity) determined by fitting of the NMR and UV/vis titration experiments. Decadic logarithms ('log') of parameters are given, the mean values correspond to the geometric mean of the maximum and minimum values. Details about the analysis and errors are given in Appendix C.5.

	DFA / TFA		CSA	
	NMR	UV/vis	NMR	UV/vis
OxP				
log K_1	4.5–8.7	2.3–6.0 ^c	4.0–5.0	
log K_1K_2	8.0–16	13–15 ^c	10–12	
log K_3	2.5–4.0	< 1.0 ^c	< 5.0	N/A
log K_3K_4	4.0–6.0	3.6–5.0 ^c	2.3–2.5	
log α_{12}	—	> 3.7 ^c	—	
log α_{34}	—	> 5 ^c	—	
Bz₂OxP				
K_1	$(1.3 \pm 3.0) \times 10^3$	$(1.8 \pm 2.0) \times 10^3$	$(8.0 \pm 5.0) \times 10^4$	$(5.0 \pm 1.0) \times 10^4$
Bz₄OxP				
K_1	10 ± 9	N/A	no spectral change	no spectral change

^aBinding constants for processes at higher acid vol.f. could not be obtained because of the solubility threshold of CSA, the unknown activity coefficients of the species in solution and peak shifts with polarity in NMR measurements. Due to the nature of the binding processes and the models used, only the range of the binding constants could be determined in the case of **OxP**. Also, due to the strong cooperativity, products of binding constants, i.e. K_1K_2 , K_3K_4 , are used to describe the binding steps $H \rightarrow HG_2$ and $HG_2 \rightarrow HG_4$, respectively (cf. Eq. (2.30)). ^bInteraction parameters were obtained as $\alpha_{12} = \frac{8}{3} \frac{K_2}{K_1}$ and $\alpha_{34} = \frac{8}{3} \frac{K_4}{K_3}$ (see Table 2.3). ^cMeasured with TFA instead of DFA.

in **Bz₄OxP**. The results are presented in Table 6.1 including the calculated values of interaction parameters for **OxP** protonation.⁷ In the case of **OxP**, only the range of the binding constants could be determined due to strong cooperativity. Note that the parameters obtained from NMR titrations are less reliable compared to those from UV/vis because shift of NMR peaks is affected by variation of solvent polarity at high acid contents. Since the values of binding constants were determined with uncertainty over several orders, their logarithms are listed, the mean values correspond to the geometric mean of the maximum and minimum values.

Solvatochromic effects in DMF/chloroform mixture were observed. Besides other interactions, central NH binding sites allow hydrogen bonding with DMF carbonyl group. All three oxoporphyrinogens show red shifts of absorption bands (positive solvatochromism, see Sec. 4.5) upon addition of DMF attributed to the increase of polarity or to hydrogen bonding (the unavoidable presence of water might have small influence on the spectra too). Furthermore, **OxP** showed transition to another form (assigned to porphodimethene form) stable only at particular DMF concentration range. Analogously to halochromism, the solvatochromic colorimetric response is also shifted to higher DMF concentration with increasing degree of N-alkylation.

UV/vis titration experiments proved as a powerful tool to investigate proton transfer and formation of hydrogen bonds in the oxoporphyrinogen conjugated systems. Together with SVD and DSIS, they enable determination of the number of individual absorbing species including their spectra and populations. Especially, mutual plot of SVD combination coefficients is a very effective instrument to assess the present spectral changes. NMR measurements provide only supporting evidence (except for direct observation of protonation in **OxP** and possibly in **Bz₂OxP**

only the presence of monoprotonation, see Appendix D.1.

⁷Here, we revised the values of α_{12} and α_{34} published in our article [1] and used the correct prefactor 8/3.

(see Appendix [E.17](#)) since at higher acid contents, all peaks undergo fast chemical exchange and thus providing only limited information from shifts of their resonant frequencies. On the contrary, NMR is an excellent tool to investigate intra- and intermolecular dynamic processes as presented in the next chapter.

7 Dynamic processes and chemical exchange in OxP derivatives

Systems of oxoporphyrinogen hosts in the presence of organic acids (chiral CSA or achiral DFA) show not only dynamic host-guest binding (described in the previous chapter) but also other processes such as rotation of bulky side groups. In this chapter, these processes are described qualitatively and some of them even quantitatively using appropriate models.

The dynamic behavior is investigated using NMR lineshape analysis covered in Sec. 3.3. NMR signal is sensitive to the kinetics of a particular nuclear spin serving as a probe of molecular processes. Although in many cases there is no direct one-to-one correspondence between chemical kinetics (involving chemical species present in the sample) and the observed spin kinetics (involving states in different environments of a reference nuclear spin), comparison of chemical and spin kinetics schemes provides the relationship between reaction and transition rate coefficients.

Two types of ^1H NMR experiments were used to assess the dynamic properties. The first type consists of titration of host with guest at constant temperature (i.e. measurement of spectra at different guest concentrations when the guest is gradually added into the host solution), which was already used to determine binding properties of oxoporphyrinogens in the previous chapter. Here, these NMR spectra are analyzed in more detail. The second type of experiments consists of variable temperature measurements at constant concentrations of host and guest in the sample. Experimental details about the measurements are given in Appendix C.4. Concentration and temperature dependence of measured transition rate coefficients is different for each process. Full description requires a specific kinetic model, which connects the transition rate coefficients to reaction rate coefficients, see Sec. 3.4.1. The latter are independent of concentration and their temperature dependence follows the Eyring equation, see Sec. 2.1.3 and 2.1.4.

Results in Sec. 7.2.3 mainly follow our publication [2] including parts of the text, figures and schemes, other results are still unpublished in a scientific journal (as of 2022). Several NMR spectra analyzed here and some preliminary results were presented in the master's thesis [45].

7.1 Dynamics of OxP

7.1.1 Prototropic tautomerization

When OxP is double protonated by an acid (while forming a host-guest complex with acid anions), the protonation is located at the opposite oxoporphyrinogen C=O sites because protonation of adjacent sites is energetically unfavorable [36]. Hence, two different tautomers denoted as (+)-OxP²⁺ and (-)-OxP²⁺ are formed, see Fig. 7.1. Double protonation reduces the OxP symmetry from D_{2d} to D_2 , thus, the tautomers are formally mutual mirror images, i.e. enantiomers (because D_2 symmetry group lacks improper rotation). These tautomeric forms can be interconverted by so-called *prototropic tautomerization* process characterized by the reaction rate coefficient κ_{taut} . Tautomerization¹ is a special type of isomerization process (structural change preserving atomic composition of the molecule), where an atom or a group of atoms is repositioned within a molecule. In OxP²⁺, the repositioned atoms are protons (H^+) originating from the acid. Figure 7.1 shows that a reference spin of *tert*-butyl-H (denoted by blue circle) changes its environment between states A_1 and A_2 , and thus undergoes chemical exchange characterized by transition rate coefficient k_{taut} (which is related to κ_{taut}).

¹tautomerization <https://goldbook.iupac.org/terms/view/T06253>

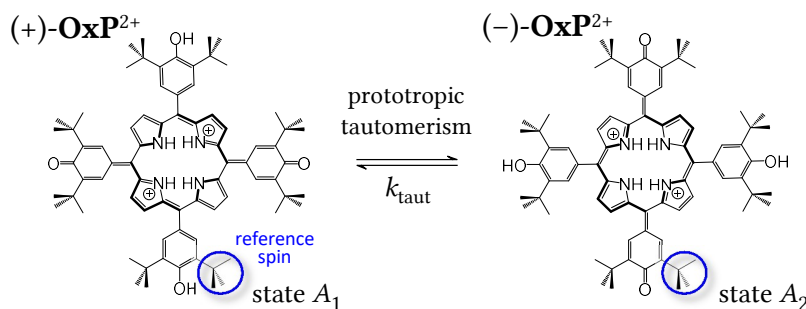


Figure 7.1: Prototropic tautomerization process. The tautomers are different due to position of protonation relative to spatial orientation of central pyrroles (bold bonds are oriented towards the reader). Influence of acid anion is not shown.

Besides protonation, the resulting geometry of the **OxP**-acid host-guest complex is shaped by the two bound acid anions (see Sec. 6.9), which create so-called ‘chiral field’ by their thermal movement around the central binding sites,² which is averaged on the NMR timescale [36, 37]. For our purposes, only the symmetry of chiral field is essential, nevertheless, its exact shape can be rigorously obtained from molecular dynamics modeling for example as angular probability function of a chosen axis in the acid anion [36, 37, 40]. Symmetry of the chiral field is influenced by symmetry of the protonated **OxP**²⁺ host and by symmetry of the acid anion guest. Molecule of DFA possesses mirror symmetry, and thus chiral field of the DFA anion does not change the D_2 symmetry of the host-guest complex as shown in Fig. 7.2a. Here, the chiral field is represented with geometric shapes with proper symmetry. Note that the tautomers (+)-**HG**₂ and (-)-**HG**₂ are mutual mirror images and thus true enantiomers. Number of states of the host nuclei are

²The anion is held close to the charged host because of hydrophobic interaction with the nonpolar solvent. The counterions do not dissociate as they would in a polar solvent.

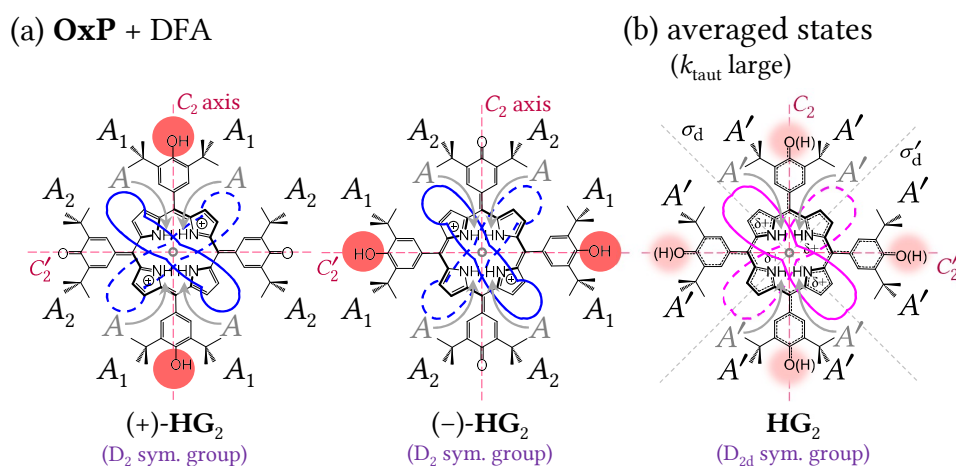


Figure 7.2: Structures of **OxP** + DFA 1:2 host-guest complex, corresponding states of *tert*-butyl-H spins (black labels; also valid for *ortho*-H and β -H spins) and states of NH spins (gray labels). (a) Spin states of two tautomeric forms (enantiomers) (+)-**HG**₂ and (-)-**HG**₂. (b) Spin states of the structure averaged due to fast tautomerization (combination of (+)-**HG**₂ and (-)-**HG**₂ structures). Solid and dashed geometric shapes with proper symmetry represent chiral field in front of and behind the host, respectively. Dashed straight lines denote symmetry elements C_2 (two-fold rotation axis) and σ_d (diagonal mirror plane). Averaging of states implies $\delta_{A'} = \frac{1}{2}(\delta_{A_1} + \delta_{A_2})$, state of NH spins does not change due to the fast exchange.

determined by symmetry of the host-guest complex. Figure 7.2a reveals two distinct states A_1 (closer to protonated site) and A_2 (further from protonated site) for *tert*-butyl-H nuclei. Note that (i) the same notation can be used for *ortho*-H or β -H signals in the vicinity of the corresponding *tert*-butyl-H (denoted by black letters in Fig. 7.2a), and (ii) another independent notation is used for central NH signals since they lie exactly in the mirror symmetry planes of **OxP** (denoted by gray letters in Fig. 7.2a). The situation is more complicated when we consider complexation with a chiral acid lacking mirror symmetry such as CSA, where tautomers (+)-**HG**₂ and (-)-**HG**₂ are not mutual mirror images [36]. Then the host-guest complex can be formed either with two (*R*)-**G**, two (*S*)-**G** or one (*R*)-**G** and one (*S*)-**G** as shown in [37].

When the tautomerization process is fast on the NMR timescale, the chiral field is averaged and D_{2d} symmetry of free **OxP** is restored, see Fig. 7.2b. This averaging is represented as superposition of geometric shapes representing the chiral field in Fig. 7.2a. The combined geometric shapes properly represent symmetry of the averaged chiral field. Fast exchange also induces new averaged spin states denoted by prime (i.e. A') in Fig. 7.2b, cf. replacement states in Sec. 3.3.8.

NMR spectra of titration of **OxP** with DFA are shown in Fig. 7.3a. At low acid concentrations, the reference *tert*-butyl-H spin assumes three states: equally populated A_1 and A_2 in the **HG**₂ complex ($[\text{HG}]$ is low due to strong cooperativity α_{12} , see Sec. 6.9) and C in the free **OxP**. Assignment of states A_1 and A_2 is not confirmed by 2D NMR experiments, however, the presence of protonation close to *tert*-butyl-H spins in state A_1 is expected to alter the Larmor frequency of

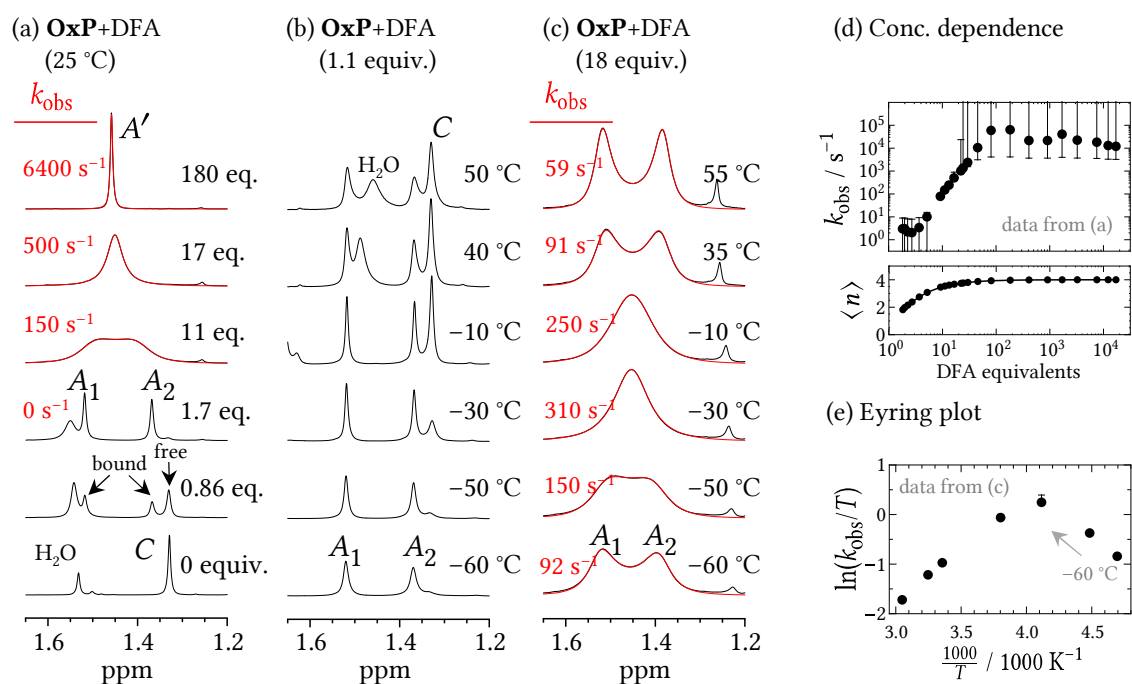


Figure 7.3: ¹H NMR spectral manifestation of prototropic tautomerization at *tert*-butyl signals in **OxP**. (a) Titration of **OxP** with DFA (full spectra in Fig. E.2). States A_1 , A_2 and their average A' refer to bound **OxP** (see Fig. 7.2), state C refers to free **OxP**. (b) Variable temperature measurements at 1.1 equiv. of DFA (full spectra in Fig. E.8). (c) Variable temperature measurements at 18 equiv. of DFA (full spectra in Fig. E.9). (d) Analysis of titration measurement from (a), dependence of k_{obs} on DFA concentration (top). Mean protonation number $\langle n \rangle = ([\text{HG}] + 2[\text{HG}_2] + 3[\text{HG}_3]) + 4[\text{HG}_4]/[\text{H}]_t$ calculated using $K_1 = 4 \times 10^6$, $K_2 = 8.4 \times 10^6$, $K_3 = 680$, $K_4 = 100 \text{ M}^{-1}$, compatible with results in Table 6.1 (bottom). (e) Analysis of variable temperature measurement from (b), the Eyring plot is nonlinear. Values of k_{obs} in (a,c) were obtained from symmetric two-state exchange fit of states A_1 and A_2 (fits are denoted by red line). Fitting procedure and meaning of error bars is explained in Appendix C.6.

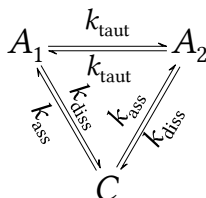
free **OxP** state *C* by greater amount than in state *A*₂ (assuming $|\delta_{A_1} - \delta_C| > |\delta_{A_2} - \delta_C|$ in Fig. 7.3a). At more than 2 equiv. of DFA, free host is absent and the exchange can be modeled and fitted with symmetric two-state exchange with observed transition rate coefficient k_{obs} , see red lines in Fig. 7.3a. Fitting results in Fig. 7.3d (top) show increase of k_{obs} with acid concentration (similarly for CSA in Fig. C.5a). Calculated mean protonation number $\langle n \rangle$ in 7.3d (bottom) indicates that all four protonation sites of **OxP** are saturated at about 10 equiv. of DFA. Complex **HG**₃ is almost not present due to strong cooperativity α_{34} (see Sec. 6.9), thus, the *D*₂ symmetry of **HG**₂ is preserved in the **HG**₄ complex. Hence, binding of third and fourth acid molecules modifies Larmor frequencies of the *A*₁ and *A*₂ states, however, no additional states are observed. Details about the fitting procedure are given in Appendix C.6. Therefore, k_{obs} describes tautomerization of diprotonated **OxP** (k_{taut}) at lower acid concentrations and tautomerization of tetraprotonated **OxP** (\tilde{k}_{taut}) at higher acid concentrations, see the illustration in Table 7.4.

Temperature dependence at 1.1 equiv. of DFA provides little information since the exchange is in slow regime, see Fig. 7.3b. At 18 equiv. of DFA causing fourfold protonation of the host, chemical exchange is in intermediate regime and the lineshape can be fitted as shown in Fig. 7.3c. Fitting results in Fig. 7.3e show that in the range from -60 °C to -30 °C the transition rate coefficient k_{obs} increases but decreases at higher temperatures. Henceforth, this process cannot be described by single energy barrier with constant ΔH^\ddagger and ΔS^\ddagger , see discussion at the end of Sec. 2.1.4.

Transfer of protons between (+)-**HG**₂ and (-)-**HG**₂ is realized in three different ways, see figure 7 in [36]. The first way is proton transfer by free acid molecules, manifested by the increase of k_{obs} with total acid concentration as described above. The second way is mediation by **OxP-OxP** collisions since the transition rate coefficient increases with $[\text{OxP}]_t$. The transition rate coefficient also increases with water concentration, thus, the process is mediated by water (carrying the protons in the form of H_3O^+).

7.1.2 Anion binding at the porphyrinogen center

Behavior of the acid anion is governed by several processes. The first one is *host-guest binding*, transition between free host and the host-guest complex. This process consists of two steps (indiscernible in NMR), protonation of a host carbonyl group and subsequent binding of guest anion at the porphyrinogen center, where the positive charge is located. In addition, intermediate species **HG** is neglected due to high cooperativity α_{12} . Host-guest association is characterized by k_{ass} and dissociation by k_{diss} as shown in Scheme 7.1. Both transition rate coefficients are



Scheme 7.1: Exchange scheme for **OxP** + DFA (or enantiopure CSA) 1:2 host-guest complex applicable to *tert*-butyl-H, β -H and *ortho*-H signals. Intermediate species **HG** is neglected due to high cooperativity α_{12} .

connected by the expression

$$\frac{k_{\text{ass}}}{k_{\text{diss}}} = \frac{p_{\text{bound}}}{p_{\text{free}}}, \quad (7.1)$$

where p_{free} and p_{bound} (i.e. relative fractions with $p_{\text{free}} + p_{\text{bound}} = 1$) are populations of free and bound host, respectively. **OxP** in the presence of either (*R*)-CSA or DFA shows separate peaks

(due to *tert*-butyl-H, β -H and *ortho*-H) due to free and bound species, see Fig. E.2, E.8 and E.3, therefore, the process of host-guest binding is slow on the NMR timescale. Half-symmetric three-state lineshape fitting of *tert*-butyl signals of **OxP** in the presence of 1 equiv. of DFA³ according to Scheme 7.1 yields $k_{\text{ass}}, k_{\text{diss}}, k_{\text{taut}} < 5 \text{ s}^{-1}$.

Another process is *anion exchange* at the porphyrinogen center characterized by k_{ex} , when an incoming anion from free acid pushes out the original bound anion. Anion exchange of enantiopure or achiral acid guests has no effect on the host spin states. However, the presence of both enantiomers of a chiral guest enables anion exchange (*R*)-**G** \leftrightarrow (*S*)-**G**, which affects the spin states by changing the corresponding chiral field, see details in [37, 38]. This process must be fast on the NMR timescale to enable determination of enantiomeric excess as explained in Sec. 3.4.3. In **OxP** + (*R*)-CSA system, the e.e.-induced resonance splitting ($\Delta\delta$) is too small for any practical use, see Fig. E.3. Other processes influencing or influenced by the acid anion behavior are prototropic tautomerization, discussed in the previous section, and macrocyclic inversion described in the next section. **OxP** in the presence of less than 2 equiv. of (*R*)-CSA induces two methyl signals of (*R*)-CSA at 0.57 ppm, see Fig. E.3. This splitting is caused by formation of (+)-**H(R)-G**₂ and (–)-**H(R)-G**₂ diastereomers, which implies slow k_{taut} and k_{ex} (there is no free acid for anion exchange). On the other hand, at more than 2 equiv. of DFA or CSA, acid nuclear spins undergo chemical exchange between free and bound acid states. Since only single averaged peak is detected (DFA triplet at 5.16 ppm, CSA methyls at 0.57 and 0.86 ppm, see Fig. E.2 and E.3), the anion exchange process k_{ex} is intermediate/fast on NMR timescale. Analogously to k_{taut} , the anion exchange process is different (occurs in different protonation states) at lower acid concentrations, where the acid anion binds at the porphyrinogen center (k_{ex}) and at higher acid concentrations without acid anion binding (\tilde{k}_{ex}), see the illustration in Table 7.4.

7.1.3 Macrocyclic inversion

Another dynamic process found in **OxP** is *macrocyclic inversion* of the porphyrinogen macrocycle characterized by k_{flip} . It consists in collective flipping of all pyrrole subunits by 90° and interconverts the tautomeric forms (+)-**HG**₂ and (–)-**HG**₂. This process was reported in [36] for **OxP** + mandelic acid system (in CDCl₃), its progress as suggested in the same article is reproduced in

³The fitted spectrum was measured during the titration of **OxP** with DFA, fitting curve is not shown. Parameters δ_j were obtained at the peak positions, parameter R_2 was obtained from the spectrum at 0 guest equiv. Equation (7.1) was used and the free/bound populations were obtained from peak integration.

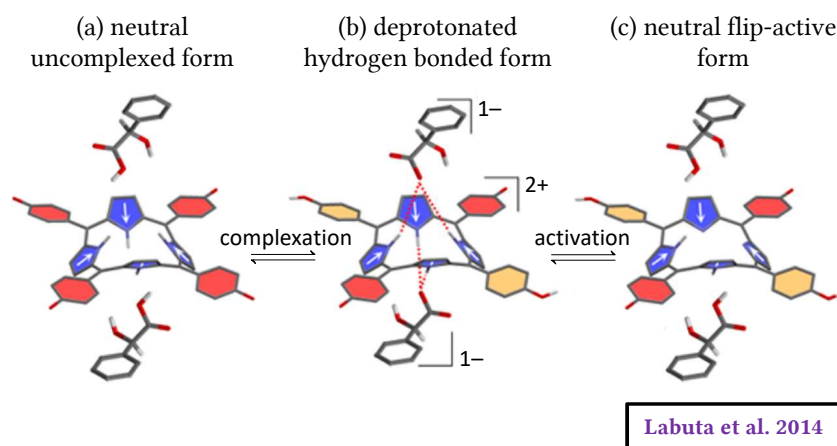


Figure 7.4: Macrocyclic inversion process in **OxP** + mandelic acid system. Adopted from [36].

Figure 7.4a–c. Free **OxP** (a) is unable of the flipping due to steric hindrance at the porphyrinogen center. In the second step (b), **HG₂** complex is formed, yet not enabling the inversion. Finally (c), each of the two acid anions removes a proton from the central NH groups rendering the host neutral, which makes the porphyrinogen center less crowded and activates the flipping process. The transition rate coefficient of flipping (not exceeding 300 s^{-1} at $50\text{ }^{\circ}\text{C}$ for **OxP** in the presence of mandelic acid) decreases with the acid concentration since strongly bound acid anion prevents flipping. Water content in the sample increases the rate because water can push the acid anion away (while keeping the double protonation in place), which enables the ring flip process and preserves its detectability, see figure 7 in [36].

States of **OxP** + (*R*)-CSA (or another enantiopure acid, e.g. (*R*)-mandelic acid) are shown in Fig. 7.5. Each diastereomer (+)-**HG₂** and (–)-**HG₂** induces spin states with different chemical shift as shown in (a). Averaging of these states occurs in three possible ways as shown in (b): (i) Fast prototropic tautomerization (and slow macrocyclic inversion), which does not change D_2 symmetry of the diastereomers, and thus does not change the number of observed peaks. (ii) Fast macrocyclic inversion (and slow prototropic tautomerization), which again does not change the symmetry and the number of observed peaks. (iii) Both prototropic tautomerization and macrocyclic inversion are fast, which averages *tert*-butyl-H (as well as *ortho*-H and β -H) and NH resonances to a single peak. Figure 7.5c shows that the reference *ortho*-H spin (denoted by blue circle) undergoes four-state chemical exchange governed by k_{taut} and k_{flip} . In the case of mandelic acid in [36], all four states were observed as separate peaks (see figure 5 in [36]) for *ortho*-H and β -H nuclear spins. Macrocyclic inversion has little effect on distant *tert*-butyl-H spins, it influences mainly the porphyrinogen center.

Unlike mandelic acid inducing four peaks, the presence of (*R*)-CSA induces a doublet in *ortho*-H and a singlet in β -H in the **HG₂** complex, see Fig. E.3. We suggest that splitting due to *tert*-butyl protonation is very small and averaged by the prototropic tautomerization above 2 equiv. of guest (i.e. k_{taut} is large compared to the corresponding *ortho*-H and β -H splitting, although small compared to splitting of *tert*-butyl-H). Furthermore, splitting by the chiral field (which can be used for e.e. determination, see Sec. 3.4.3) is observable in *ortho*-H⁴ but negligible in β -H. This situation implies the absence/slow regime of macrocycle inversion process, which would otherwise cause averaging of the *ortho*-H resonance to a singlet. This conclusion is also supported by the splitting of bound guest methyl signal at 0.57 ppm below 2 guest equiv. (each signal belongs to different diastereomer), which contradicts the presence of fast/intermediate macrocyclic inversion or prototropic tautomerization. The structure in Fig. 7.5b (left) implying two different *ortho*-H states A'_I and A''_I is relevant for this system. The reason for the absence of macrocyclic inversion is probably the greater binding strength of (*R*)-CSA compared to mandelic acid ($K_1K_2 > 10^8\text{ M}^{-2}$ for CSA compared to $K_1K_2 < 10^7\text{ M}^{-2}$ for mandelic acid, see Sec. 6.9 and [36]) preventing the activation step in Fig. 7.4c.

In **OxP** + DFA system, macrocyclic inversion does not alter the spin states shown in Fig. 7.2a. Hence, this process cannot be detected by NMR in this system. Macrocyclic inversion is not possible in **Bz₂OxP** and **Bz₄OxP** due to steric hindrance.

⁴Spectrum of free host implies $J(\beta\text{-H})=2.5\text{ Hz}$ while splitting due to weak J-coupling of *ortho*-H cannot be even observed, see Fig. E.3. The splitting of *ortho*-H in **HG₂** is 0.008 ppm corresponding to 4 Hz. Thus, this value is unrealistically high to be caused by J-coupling and the splitting is ascribed to chiral field.

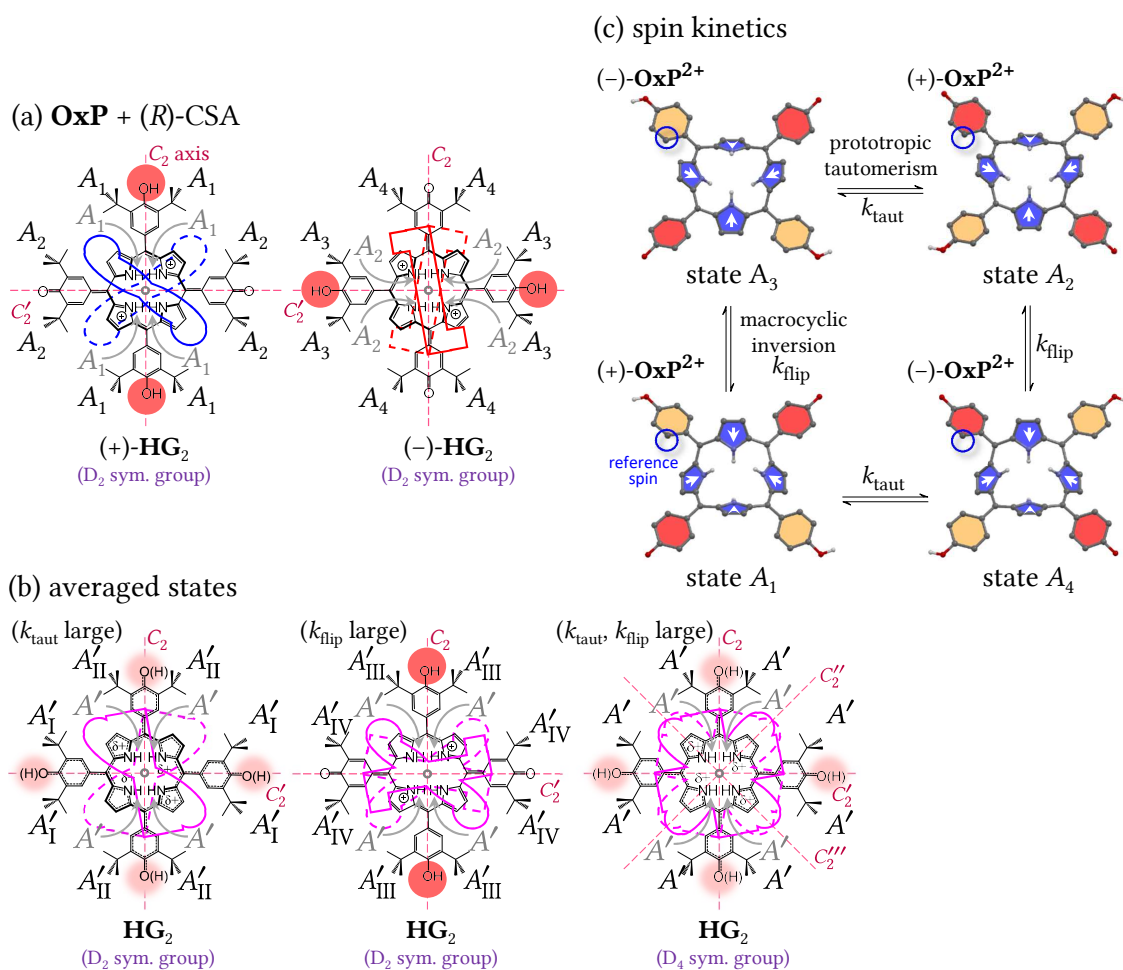


Figure 7.5: Structures of **OxP** + (*R*)-CSA 1:2 host-guest complex, corresponding states of *tert*-butyl-H spins (black labels; also valid for *ortho*-H and β -H spins) and states of NH spins (gray labels). (a) Spin states in the two tautomeric forms (diastereomers) (+)-HG₂ and (-)-HG₂, (b) Averaged structures (combination of (+)-HG₂ and (-)-HG₂ structures) due to fast prototropic tautomerization (left), fast macrocyclic inversion (middle) and due to the combined effect of these two (right). Solid and dashed geometric shapes with proper symmetry represent chiral field in front of and behind the host, respectively. Dashed straight lines denote symmetry elements C₂ (two-fold rotation axis). Fast exchange implies (Eq. (3.62)) in *tert*-butyl-H peaks: $\delta_{A'_I} = p_+ \delta_{A_2} + p_- \delta_{A_3}$, $\delta_{A'_{II}} = p_+ \delta_{A_1} + p_- \delta_{A_4}$, $\delta_{A'_{III}} = p_+ \delta_{A_1} + p_- \delta_{A_3}$, $\delta_{A'_{IV}} = p_+ \delta_{A_2} + p_- \delta_{A_4}$, $\delta_{A'} = \frac{p_+}{2} (\delta_{A_1} + \delta_{A_2}) + \frac{p_-}{2} (\delta_{A_3} + \delta_{A_4})$, where p_+ and p_- are relative populations of the species (+)-H(*R*)-G₂ and (-)-H(*R*)-G₂, respectively. Fast exchange implies in NH peaks: $\delta_{A'} = p_+ \delta_{A_1} + p_- \delta_{A_2}$. (c) Processes of macrocyclic inversion (k_{flip}) and prototropic tautomerization (k_{taut}) in **OxP** in the presence of enantiopure guest. The reference *ortho*-H spin (denoted by blue circle) changes its position among states A₁, A₂, A₃ and A₄ (corresponding to *a'*, *b'*, *a''* and *b''* in [36], respectively). *Tert*-butyl groups and some hydrogen atoms are removed for clarity. Yellow colored hemiquinonoid groups are subject to protonation.

7.2 Dynamics of **Bz₂OxP**

7.2.1 Topomerization at hemiquinonoid groups

*Topomerization*⁵ is a process converting a chemical species into itself, typically rotation of a side group. In protonated **Bz₂OxP**, double bond at the hemiquinonoid group reduces to single bond enabling rotation of the whole group about its *meso*-position bond as denoted by green color in Fig. 7.6. Although the chemical structure does not change, the reference *tert*-butyl nuclear spin (denoted by circle) changes its environment between states *A* and *B*. Consequently, **Bz₂OxP** hemiquinone topomerization is manifested as symmetric two-state exchange in the NMR spectrum (*ortho*-H signals are sensitive to this process as well, but they do not coalesce due to large peak separation). In fact, each of states *A* and *B* consists of multiple substates, which are averaged due to other processes, as will be shown later. Due to this averaging, all four hemiquinonoid groups appear as if they rotate simultaneously on the NMR timescale even when the mean protonation number $\langle n \rangle < 4$. Although hemiquinonoid topomerization almost certainly takes place in **OxP**, it is undetectable because the spin environments are identical (the rotation axis coincides with the *C*₂ symmetry axis, which also lies in the mirror symmetry plane). However, **Bz₂OxP** possesses intrinsic asymmetry due to N-alkylation revealing the topomerization process.

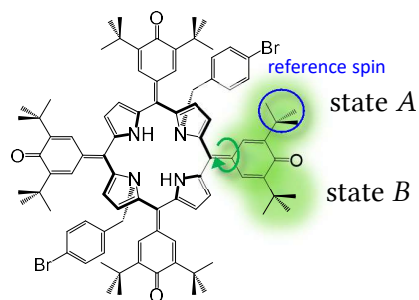


Figure 7.6: Rotation of hemiquinonoid group (topomerization) in protonated **Bz₂OxP**

As expected, the observed transition rate coefficient k_{obs} indicating the hemiquinone topomerization (the relationship between k_{obs} and the “true” reaction/transition rate coefficient $\kappa_{\text{topo}} = k_{\text{topo}}$ is complicated as discussed at the end of this section) depends on acid concentration, see Fig. 7.7a. Initially, increasing the mean protonation number $\langle n \rangle$ on average reduces double bonds of the hemiquinonoid groups to single bonds and k_{obs} increases, see Fig. 7.7f ($\langle n \rangle$ is calculated only up to 20 equiv. because only the value of K_1 is available, see Sec. 6.9). Above ca. 200 equiv. of DFA, rate of the process decreases with acid concentration. The absence of chemical exchange in NMR spectra without any acid (even at high temperatures) proves the necessity of protonation to activate the rotation, see Fig. 7.7b. Concentration dependence at room temperature was observed only in the presence of DFA not CSA (see spectrum in Fig. E.5), energy barrier in the latter case is too high and the effect of topomerization is only observed at higher temperatures as discussed below.

Variable temperature measurements in the presence of constant amount of DFA or CSA (enough to induce $\langle n \rangle \geq 1$ so that free host contribution to the spectral lineshape is negligible) in Fig. 7.7c,d show increase in k_{obs} as determined by lineshape fitting, see details in Appendix C.6. In the system of **Bz₂OxP** + DFA (Fig. 7.7c), the topomerization is fast enough to induce an averaged state *T'* formed from *A'* and *B'*, see the corresponding structures in Fig. C.7 in Appendix C.6. Temperature dependence of k_{obs} is linear in Eyring plot, see Fig. 7.7g,h, thus, the hemiquinone

⁵topomerization <https://goldbook.iupac.org/terms/view/T06396>

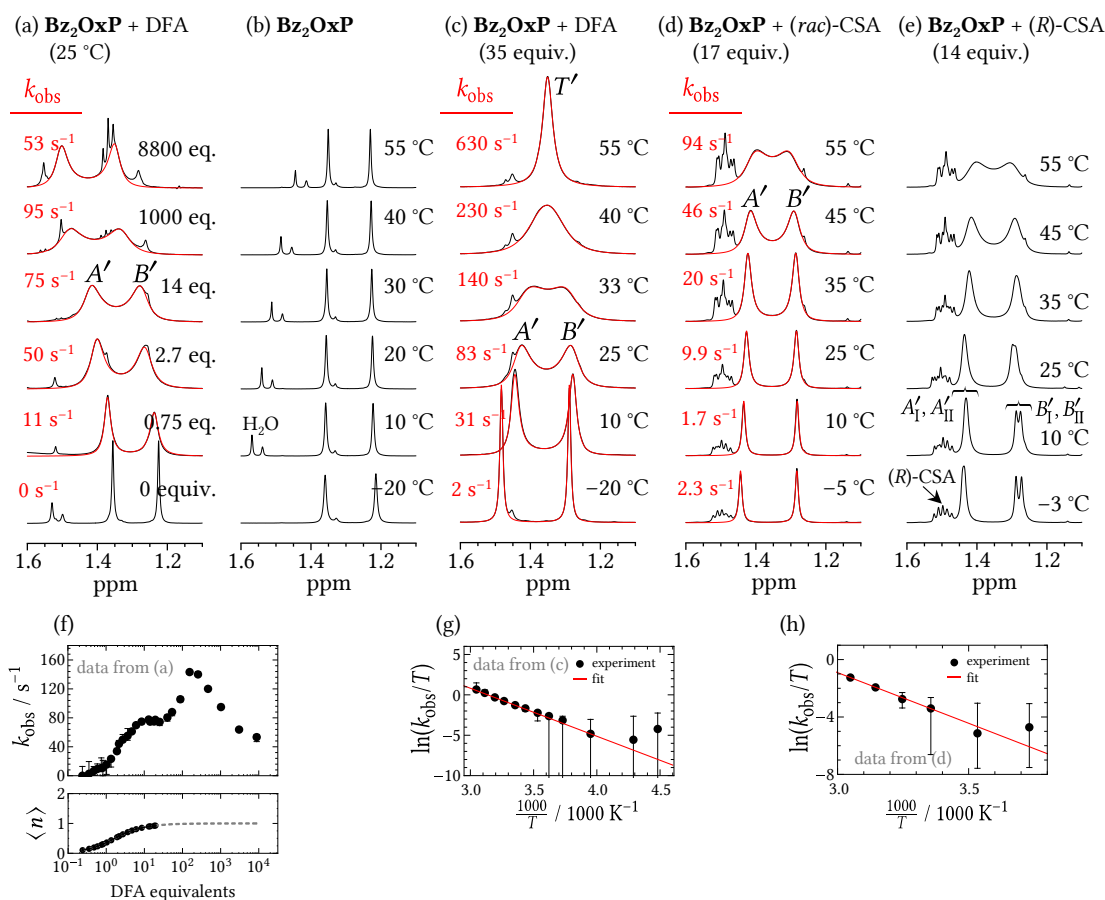


Figure 7.7: Analysis of NMR *tert*-butyl signals of **Bz₂OxP** in various conditions reflecting rotation of the hemiquinonoid groups. (a) Titration of **Bz₂OxP** with DFA (full spectra in Fig. E.4). (b) Variable temperature measurement of **Bz₂OxP**. There is no chemical exchange in the absence of acid. (c) Variable temperature measurements of **Bz₂OxP** with 35 equiv. of DFA (full spectra in Fig. E.11). (d) Variable temperature measurements of **Bz₂OxP** with 17 equiv. of (*rac*)-CSA (full spectra in Fig. E.15). (e) Variable temperature measurements of **Bz₂OxP** with 14 equiv. of (*R*)-CSA (full spectra in Fig. E.13). (f) Concentration dependence of k_{obs} for titration of **Bz₂OxP** with DFA obtained from lineshape fitting of spectra in panel (a) (top) and the corresponding mean protonation number ($\langle n \rangle = [\text{HG}]/[\text{H}]_t$) calculated using $K_1 = 1260 \text{ M}^{-1}$ (bottom). (g) Eyring plot of **Bz₂OxP** + DFA system obtained from lineshape fitting of spectra in panel (c). (h) Eyring plot of **Bz₂OxP** + (*rac*)-CSA system obtained from lineshape fitting of spectra in panel (d). Details about meaning of error bars, lineshape fitting procedure and Eyring plot fitting are given in Appendix C.6. Structures corresponding to states in **Bz₂OxP** in the presence of DFA or (*R*)-CSA are given in Fig. C.7 in Appendix C.6 or in Fig. 7.8, respectively.

topomerization process can be described by a single Gibbs energy barrier $\Delta G^\ddagger = \Delta H^\ddagger - T\Delta S^\ddagger$ (Sec. 2.1.4). Weighted linear fit of the Eyring plots quantifies the Gibbs energy barrier as shown in Table 7.1. Room temperature (298 K) entropic contribution to the Gibbs energy barrier ($-T\Delta S^\ddagger$)

Table 7.1: Results of Eyring plot fitting for topomerization of hemiquinonoid group of **Bz₂OxP** in the presence of acid. The fitting procedure is described in Appendix C.6.

acid	$\Delta H^\ddagger / \text{kJ.mol}^{-1}$	$\Delta S^\ddagger / \text{J.mol}^{-1}.\text{K}^{-1}$	$\Delta G^\ddagger (25^\circ\text{C}) / \text{kJ.mol}^{-1}$	$-T\Delta S^\ddagger (25^\circ\text{C}) / \text{kJ.mol}^{-1}$
DFA	49.4 ± 1.4	42.0 ± 5.0	36.8 ± 2.0	-12.7 ± 1.4
CSA	58.7 ± 2.6	29.0 ± 8.0	50.1 ± 3.5	-8.6 ± 2.4

is smaller than the enthalpic contribution (ΔH^\ddagger), however, it is not negligible as would be expected for rotation processes with no bond breaking [175]. Hence, the structure of protonated complex is rearranged during the topomerization process. Relevance of using k_{obs} in the Eyring plot (instead of reaction rate coefficient κ_{topo}) is discussed at the end of this section.

Comparison of **Bz₂OxP** spectra with racemic CSA and its (*R*)-enantiomer in Fig. 7.7d,e reveals an important difference, the presence of (*R*)-CSA induces four *tert*-butyl states, while (*rac*)-CSA only two. To understand this effect caused by CSA chirality, it is necessary to analyze all possible *tert*-butyl spin states of the monoprotonated **Bz₂OxP** host (states of higher protonation are negligible for **Bz₂OxP** + CSA system, see Sec. 6.9).

Figure 7.8a shows host structures with four different protonation sites in **Bz₂OxP** + (*R*)-CSA system. They comprise of two chemically distinct tautomers (+)-**HG** and (–)-**HG**. The labels (+) and (–) have been assigned arbitrarily (without the influence of guest anion, protonated species (+)-**H**⁺ and (–)-**H**⁺ are mutual mirror images, i.e. enantiomers). Asymmetric acid anions create different chiral field at the porphyrinogen center of (+)-**HG** and (–)-**HG**, which is schematically denoted by blue or red asymmetric shapes, thus, NMR spectra of (+)-**HG** (*tert*-butyl states A_1, \dots, A_4) and (–)-**HG** (*tert*-butyl states A_5, \dots, A_8) tautomers are different.⁶ First and third structures in Fig. 7.8a (i.e. (+)-**H(R)-G**) are chemically identical, but a reference *tert*-butyl nuclear spin in both structures is located in different environments, similarly for (–)-**H(R)-G**. Note that states B_j are always in the vicinity of N-bromobenzyl groups contrary to A_j states. Figure 7.8b shows four different protonation states in **Bz₂OxP** + (*S*)-CSA system. The structure of (–)-**H(S)-G** is a mirror image of (+)-**H(R)-G** and the structure of (+)-**H(S)-G** is a mirror image of (–)-**H(R)-G** as schematically denoted by the shape and orientation of the chiral field. This implies that, for example, spin state A_1 is located at (+)-**H(R)-G** as well as (–)-**H(S)-G** and similarly for other states.

Besides hemiquinone topomerization, other processes change the environment of a reference *tert*-butyl spin, see next two sections for details. These processes are characterized by k_{taut} , k_{taut}^* and k_{ex} , the corresponding chemical exchange is fast on the NMR scale above –20 °C, thus, the spin states and their NMR signals are averaged, see Fig. E.4, E.5 and E.6. In the case of **Bz₂OxP** in the presence of (*R*)-CSA, all four structures in Fig. 7.8a are combined resulting to averaged chiral field inducing C_2 symmetry of the host-guest complex, see Fig. 7.8c (left). Averaged states A_{I} and A_{II} are formed from substates A_1, \dots, A_8 , averaged states B_{I} and B_{II} are formed from substates B_1, \dots, B_8 . These averaged states A_{I} , A_{II} , B_{I} and B_{II} are observed in Fig. 7.7e for **Bz₂OxP** in the presence of (*R*)-CSA (two resonances overlap at 1.44 ppm). The presence of racemate combines eight states from Fig. 7.8a,b resulting to averaged chiral field inducing C_{2v} symmetry of the host-guest complex, see Fig. 7.8c (right). This yields only two states *A* and *B* and explains the spectra in Fig. 7.7d for **Bz₂OxP** in the presence of (*rac*)-CSA. Note that the aforementioned sensitivity to chirality follows the mechanism of determination of enantiomeric excess by prochiral solvating agents as described in Sec. 3.4.3. Spectra of **Bz₂OxP** in the presence of DFA in Fig. 7.7c include only two *tert*-butyl states *A'* and *B'* because the **Bz₂OxP**·CSA complex has C_{2v} symmetry (induced by the mirror symmetry of DFA), see the corresponding structures in Fig. C.7 in Appendix C.6.

At –60 °C and low number of acid equivalents, all k_{topo} , k_{ex} , k_{taut} and k_{taut}^* are small enough that the exchange enters intermediate regime and some of the states A_1, \dots, A_8 and B_1, \dots, B_8 are directly visible in NMR although being broadened, see Fig. 7.8d. Note that chirality does not play a role and low-temperature spectra in the presence of (*R*)-CSA or (*rac*)-CSA are identical.

Determination of energy barrier of topomerization process requires knowledge of κ_{topo} temperature dependence. Our previous analysis in Table 7.1 made a simple assumption $\kappa_{\text{topo}} = k_{\text{obs}}$, however, according to Fig. 7.7d k_{obs} depends on acid concentration, and thus it is not equal

⁶As in the case of **OxP**, the chiral field is formed since the charged species interact strongly together due to hydrophobicity of the CDCl₃ solvent.

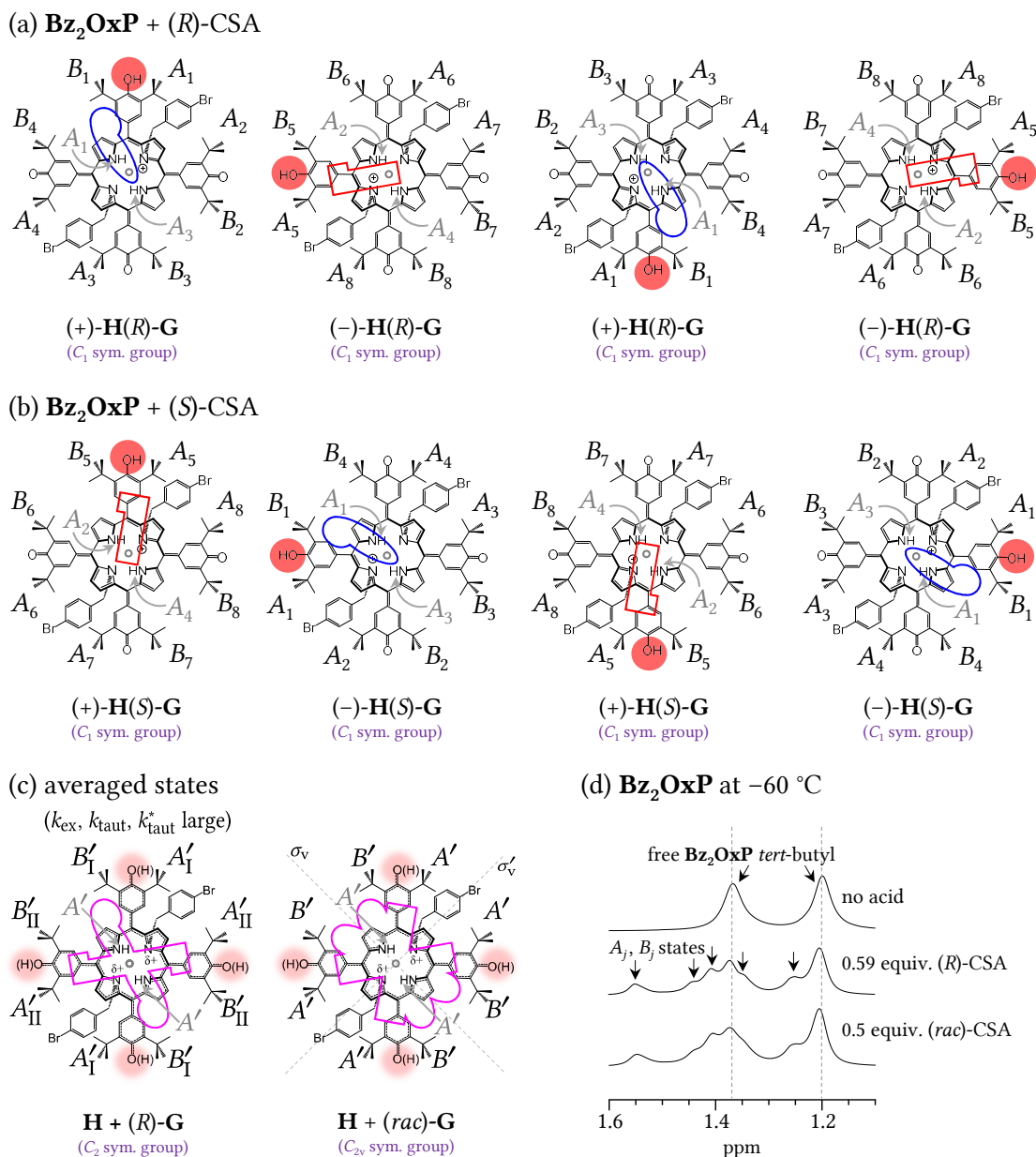


Figure 7.8: Structures of **Bz₂OxP** + CSA complex, corresponding states of *tert*-butyl-H spins (black labels; also valid for *ortho*-H and β -H spins) and states of NH spins (gray labels). (a) Spin states in **Bz₂OxP**⁺·(*R*)-CSA[–] complex, (b) spin states in **Bz₂OxP**⁺·(*S*)-CSA[–] complex. (c) Fast prototropic tautomerization processes cause averaging of chiral field in **Bz₂OxP** + (*R*)-CSA system (consists of averaged states from (a)) and **Bz₂OxP** + (*rac*)-CSA system (consists of averaged states from (a) and (b)), also requires fast anion exchange process). Geometric shapes with proper symmetry represent chiral field created by the acid anion. Dashed straight lines denote symmetry elements σ_v (vertical mirror plane). Fast exchange implies (Eq. (3.62)) in *tert*-butyl-H peaks: $\delta_{A'_I} = \frac{p_{+,R}}{2}(\delta_{A_1} + \delta_{A_3}) + \frac{p_{-,R}}{2}(\delta_{A_6} + \delta_{A_8})$, $\delta_{A'_{II}} = \frac{p_{+,R}}{2}(\delta_{A_2} + \delta_{A_4}) + \frac{p_{-,R}}{2}(\delta_{A_5} + \delta_{A_7})$, $\delta_{B'_I} = \frac{p_{+,R}}{2}(\delta_{B_1} + \delta_{B_3}) + \frac{p_{-,R}}{2}(\delta_{B_6} + \delta_{B_8})$, $\delta_{B'_{II}} = \frac{p_{+,R}}{2}(\delta_{B_2} + \delta_{B_4}) + \frac{p_{-,R}}{2}(\delta_{B_5} + \delta_{B_7})$, $\delta_{A'} = \frac{1}{2}(\delta_{A'_I} + \delta_{A'_{II}}) = \frac{p_{+,R}}{4}(\delta_{A_1} + \delta_{A_2} + \delta_{A_3} + \delta_{A_4}) + \frac{p_{-,R}}{4}(\delta_{A_5} + \delta_{A_6} + \delta_{A_7} + \delta_{A_8})$ and $\delta_{B'} = \frac{1}{2}(\delta_{B'_I} + \delta_{B'_{II}}) = \frac{p_{+,R}}{4}(\delta_{B_1} + \delta_{B_2} + \delta_{B_3} + \delta_{B_4}) + \frac{p_{-,R}}{4}(\delta_{B_5} + \delta_{B_6} + \delta_{B_7} + \delta_{B_8})$, where $p_{+,R}$, $p_{-,R}$, $p_{+,S}$ and $p_{-,S}$ are relative populations of the species (+)-H(*R*)-G, (–)-H(*R*)-G, (+)-H(*S*)-G or (–)-H(*S*)-G, respectively. Since some species are exact mirror images of others, and thus possess the same energy, it follows that $p_{+,R} = p_{-,S}$ and $p_{-,R} = p_{+,S}$. Fast exchange implies in NH peaks: $\delta_{A'} = \frac{p_{+,R}}{2}(\delta_{A_1} + \delta_{A_3}) + \frac{p_{-,R}}{2}(\delta_{A_2} + \delta_{A_4})$. (d) Some of the states from (a) and (b) are detected in low-temperature NMR spectra (full spectra in Fig. E.11, E.12 and E.14).

to a reaction rate coefficient in general. We suggest a simple model, which naturally takes the concentration dependence into account,

$$k_{\text{obs}} = \langle n \rangle \kappa_{\text{topo}}. \quad (7.2)$$

This model predicts linear dependence (with zero intercept) of k_{obs} on $\langle n \rangle$ at $\langle n \rangle \leq 1$, which is approximately confirmed experimentally in Fig. 7.9. However, the presence of free host affects the spectral lineshapes and introduces error to k_{obs} determined by two-state fitting as discussed in Appendix C.6. Thus, the observed dependence is not perfectly linear. In the presence of 35 equiv. of DFA or 17 equiv. of CSA $\langle n \rangle \approx 1$, thus, our simple assumption $\kappa_{\text{topo}} = k_{\text{obs}}$ used in Table 7.1 is valid. At more than 200 equiv. of DFA, k_{obs} decreases with concentration. This effect is probably related to increasing steric hindrance due to the high abundance of guest molecules around carbonyl groups.

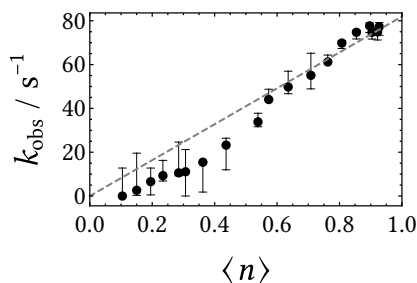


Figure 7.9: Linear correlation between $\langle n \rangle$ and k_{obs} in titration of **Bz₂OxP** with DFA. Dashed line stands for weighted linear fit.

7.2.2 Anion binding at the porphyrinogen center

Processes in **Bz₂OxP** + acid systems involving transfer of the acid anion are analogous to those described for **OxP** in Sec. 7.1.2. The first process is the host-guest binding (accompanied by host protonation), where the association and dissociation is characterized by k_{ass} and k_{diss} (or by concentration-independent κ_{ass} and κ_{diss}). Unlike for **OxP**, there are no separate peaks of free and bound host except for NH signals with bigger separation $|\delta_{\text{bound}} - \delta_{\text{free}}|$, however, other host signals are significantly broadened in the presence of low acid concentration implying intermediate exchange regime of k_{ass} and k_{diss} . Values of k_{ass} and k_{diss} are determined in the next section by lineshape fitting of NH signals.

Another process is the exchange of the bound acid anion characterized by k_{ex} (described in detail for **OxP** in Sec. 7.1.2). This process can also be accompanied by the prototropic tautomerization, see the next section and Appendix C.9. Exchange of achiral anions or identical enantiomers does not alter spin state of a reference spin. However, the spin state is altered when, for example, (*R*)-CSA anion exchanges for (*S*)-CSA anion thus enabling determination of e.e. [45] in the fast exchange regime of k_{ass} , k_{taut} and k_{ex} (exchange broadening must not be present). In theory, all hydrogen signals are split due to e.e. except NH since they lie exactly in the prochiral plane, see Sec. 3.4.3 for details. Actually, only β -protons at alkylated pyrroles (\square in Fig. E.13) provide e.e.-induced splitting $\Delta\delta$ big enough for practical use at room temperature [45], compare Fig. E.13 and E.15. Interestingly, β -protons at non-alkylated pyrroles (\square in Fig. E.13) are J-coupled to each other and also to the NH hydrogen, which results to involved spectral lineshape with badly detectable splitting [40]. *Tert*-butyl resonances split due to e.e. below 10 °C, see Fig. 7.7d,e. Above 10 °C, the hemiquinone topomerism interferes with e.e. splitting through interchange of states $A_{\text{I}} \leftrightarrow B_{\text{I}}$ and $A_{\text{II}} \leftrightarrow B_{\text{II}}$ in Fig. 7.8c and causes merger of the corresponding

resonances. At higher guest concentrations, host-guest stoichiometries higher than 1:1 appear implying guest binding at the C=O sites at the edge of the host, which is characterized by another set of transition rate coefficients \tilde{k}_{ass} , \tilde{k}_{diss} and \tilde{k}_{ex} , see Table 7.5.

7.2.3 Prototropic tautomerization

Prototropic tautomerization, i.e. change of the protonation site in the host-guest complex, is observed at *tert*-butyl resonances in the case of **OxP** + acid systems. However, **Bz₂OxP** + acid systems show *tert*-butyl peak broadening rather than splitting upon protonation due to their small peak separation. Thus, **Bz₂OxP** *tert*-butyl peaks cannot be used as a probe to the tautomerization process. However, the protonation induces partial charge in the porphyrinogen center (see Sec. 6.4), and thus affects state of the central NH spins possessing significant peak separation upon protonation. Hence, NH groups serve as probes of the tautomerization process. In this section, we describe the prototropic tautomerization in the **Bz₂OxP** + (*R*)-CSA system in detail.

The central NH protons of the **Bz₂OxP** host in the presence of (*R*)-CSA (and unavoidable presence of water) are subject to chemical exchange between six environments, which is experimentally manifested in its NMR spectrum as three resonances (some of the resonances are already merged due to fast exchange, see below) in the titration spectra at low $[G]_t$, see Fig. 7.12a (full spectra in Fig. E.5). At low $[G]_t$, the three observable resonances of the reference spin (denoted by green color in Fig. 7.10a) are well separated and the exchange between them is in a slow regime. The peak at 9.3 ppm, denoted as state C' , vanishes at 1 guest equivalent (equivalents

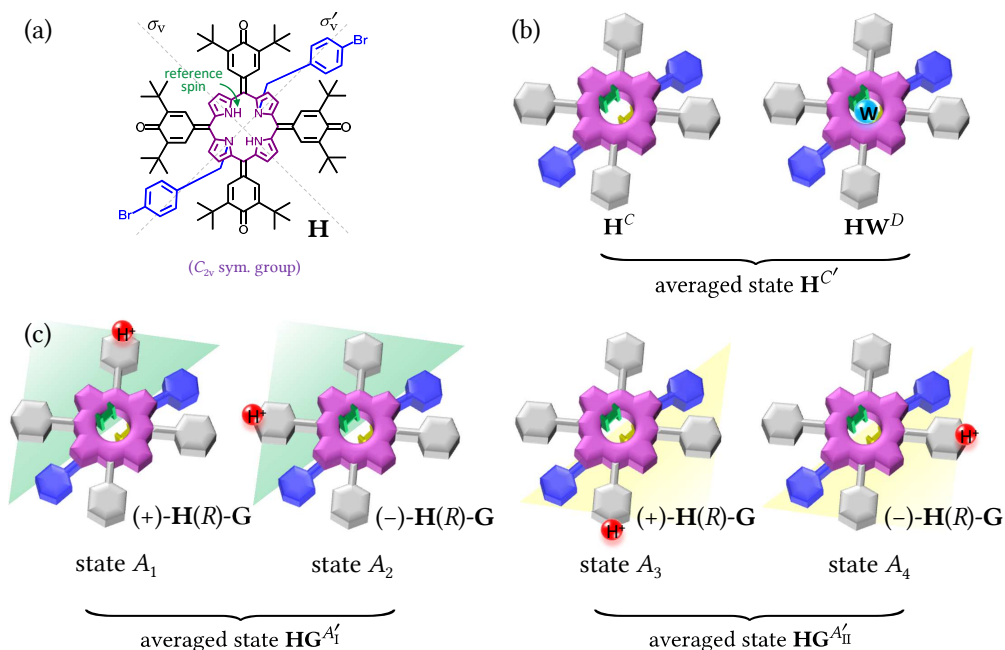


Figure 7.10: Spin states of central NH protons of **Bz₂OxP** host molecule (**H**) in the presence of CSA (**G**). (a) Structure of **H**. Bromobenzyl groups are situated behind the molecule. The spin states are described with respect to reference proton denoted by green arrow. (b) Schematic representation of free host H^C and host-water complex HW^D inducing the spin states denoted in superscript. (c) Protonated host-guest complex in two distinct tautomeric forms (+)-**H(R)-G** and (-)-**H(R)-G** differing in the site of protonation. NH spins of (+)-**HG** tautomer can be present in two spin states HG^{A_1} or HG^{A_3} (similarly NH spins of (-)-**HG** can be in states HG^{A_2} or HG^{A_4}) depending on the carbonyl protonation proximity to the green-labeled reference proton. Different sites of protonation within the green or yellow zones, form the averaged states A_1' and A_{II}' , respectively. The superscripts (e.g. H^C , HG^{A_1}) denote the spin state with respect to the green-labeled reference NH proton.

defined as $[G]_t/[H]_t$). The state C' actually comprises substates C and D averaged due to fast chemical exchange, see Fig. 7.10b. These substates correspond to free host and its complex with water, respectively (details are discussed below). The two resonances at 13.1 and 11.8 ppm correspond to two equally populated (due to the σ'_v symmetry of the host) states denoted as A'_I and A'_{II} , respectively. The state A'_I comprises substates A_1 and A_2 , the state A'_{II} comprises substates A_3 and A_4 , each two substates are averaged due to the fast exchange. States A_1, \dots, A_4 of the green-labeled reference NH spin in Fig. 7.10c correspond to protonated host with protonation at four different C=O sites. Hence, states A_1 and A_3 correspond to (+)-**HG** tautomer and states A_2 and A_4 to (-)-**HG** tautomer (NH spin states of the host in the presence of (*R*)- or (*S*)-form of the guest are also shown in Fig. 7.8a,b including chiral field induced by the acid anion). Considering the mirror symmetry plane σ'_v of the host (shown in Fig. 7.10a), the reference spin is in state A_1 or A_2 (forming A'_I) when the protonation is at the same side of the molecule as the reference spin and in state A_3 or A_4 (forming A'_{II}) when the protonation is on the other side.

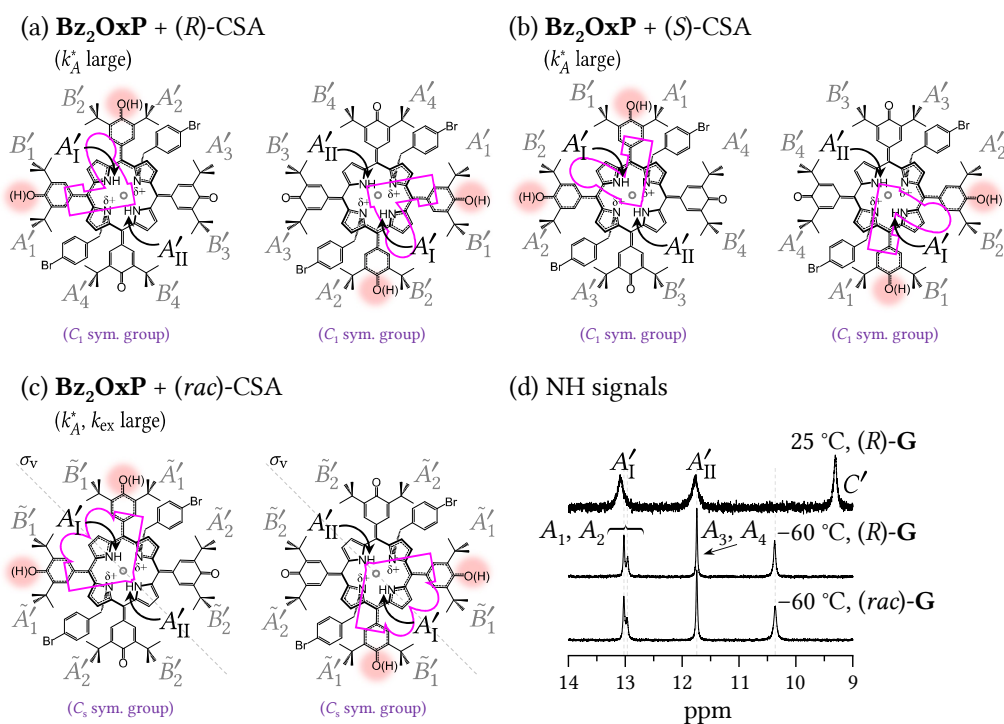


Figure 7.11: (a) Averaged structures induced by fast exchange for the system of **Bz₂OxP** + (*R*)-CSA (large $k_{A'}^*$), (b) **Bz₂OxP** + (*S*)-CSA (large $k_{A'}^*$) and (c) **Bz₂OxP** + (*rac*)-CSA (large $k_{A'}^*$ and k_{ex}). Corresponding states of NH spins (black labels) and states of *tert*-butyl-H spins (gray labels; also valid for *ortho*-H and β -H spins) are formed as combination of spin states in Fig. 7.8a,b. Fast exchange implies (Eq. (3.62)) in NH peaks: $\delta_{A'_I} = p_{+,R}\delta_{A_1} + p_{-,R}\delta_{A_2}$ and $\delta_{A'_{II}} = p_{+,R}\delta_{A_3} + p_{-,R}\delta_{A_4}$. Fast exchange implies in *tert*-butyl-H peaks: $\delta_{A'_I} = p_{+,R}\delta_{A_4} + p_{-,R}\delta_{A_5}$, $\delta_{A'_2} = p_{+,R}\delta_{A_1} + p_{-,R}\delta_{A_6}$, $\delta_{A'_3} = p_{+,R}\delta_{A_2} + p_{-,R}\delta_{A_7}$, $\delta_{A'_4} = p_{+,R}\delta_{A_3} + p_{-,R}\delta_{A_8}$, $\delta_{B'_1} = p_{+,R}\delta_{B_4} + p_{-,R}\delta_{B_5}$, $\delta_{B'_2} = p_{+,R}\delta_{B_1} + p_{-,R}\delta_{B_6}$, $\delta_{B'_3} = p_{+,R}\delta_{B_2} + p_{-,R}\delta_{B_7}$, $\delta_{B'_4} = p_{+,R}\delta_{B_3} + p_{-,R}\delta_{B_8}$, $\delta_{A'_1} = \frac{p_{+,R}}{2}(\delta_{A_1} + \delta_{A_4}) + \frac{p_{-,R}}{2}(\delta_{A_5} + \delta_{A_6})$, $\delta_{A'_2} = \frac{p_{+,R}}{2}(\delta_{A_2} + \delta_{A_3}) + \frac{p_{-,R}}{2}(\delta_{A_7} + \delta_{A_8})$, $\delta_{B'_1} = \frac{p_{+,R}}{2}(\delta_{B_1} + \delta_{B_4}) + \frac{p_{-,R}}{2}(\delta_{B_5} + \delta_{B_6})$, $\delta_{B'_2} = \frac{p_{+,R}}{2}(\delta_{B_2} + \delta_{B_3}) + \frac{p_{-,R}}{2}(\delta_{B_7} + \delta_{B_8})$, where $p_{+,R}$, $p_{-,R}$, $p_{+,S}$ and $p_{-,S}$ are relative populations of the species (+)-**H**(*R*)-**G**, (-)-**H**(*R*)-**G**, (+)-**H**(*S*)-**G** and (-)-**H**(*S*)-**G**, respectively. Since some species are exact mirror images of others, and thus possess the same energy, it follows that $p_{+,R} = p_{-,S}$ and $p_{-,R} = p_{+,S}$. (d) ¹H NMR spectra of NH resonances of host ($8.4 \times 10^{-4} \text{ M}^{-1}$, CDCl₃) with 0.59 equiv. (*R*)-**G** at -60 °C and 25 °C, host with 0.5 equiv. (*rac*)-**G** at 25 °C. The low-temperature spectra show the presence of four states A_1, \dots, A_4 (intensity ratio of the peaks at 13.03 and 12.96 ppm is 74:26), the room-temperature spectrum shows only two averaged states A'_I and A'_{II} .

Transitions between substates A_1, \dots, A_4 are accomplished by the prototropic tautomerization process, mediated by the guest anion (details are discussed below). For simplicity, we characterize this process by transition rate coefficient $k_{A'}$ (corresponding to reaction rate coefficient κ_{taut}), when the protonation changes across the σ'_v plane (i.e. transitions $A_1 \leftrightarrow A_3$, $A_1 \leftrightarrow A_4$, $A_2 \leftrightarrow A_3$ or $A_2 \leftrightarrow A_4$); or $k_{A'}^*$ (corresponding to κ_{taut}^*) otherwise (i.e. transitions $A_1 \leftrightarrow A_2$ or $A_3 \leftrightarrow A_4$). Large value of $k_{A'}^*$ causes averaging of the substates to A'_I and A'_{II} as shown in detail in Fig. 7.11a including the corresponding chiral field (cf. Fig. 7.8a). The presence of (*S*)-CSA induces identical A'_I and A'_{II} states in the host since the NH groups lie in the prochiral plane σ_v , see Fig. 7.11b (cf. Fig. 7.8b). Therefore, also the presence of (*rac*)-CSA induces identical A'_I and A'_{II} states and chemical shift of these averaged states does not depend on enantiomeric excess of guest. Note that averaged states of other hydrogen spins not located in the prochiral plane are in different states in the presence of (*rac*)-CSA compared to (*R*)-CSA. These states are denoted by tilde in Fig. 7.11c and also require fast anion exchange k_{ex} in order to induce single resonance. At higher guest concentrations, the resonance in state C' vanishes and the resonances in states A'_I and A'_{II} start to coalesce and enter intermediate then fast exchange regimes. The appearance of only one resonance implies that fast prototropic tautomerization process restores the C_2 symmetry even in protonated host, see the structures in Fig. 7.8c. Furthermore, titration with (*R*)-CSA produced the same spectral behavior of the NH resonances (see Fig. E.6). Note that in the **Bz₂OxP** + DFA system, the tautomerization process is fast, and thus only one NH resonance is observed in the whole DFA concentration range, the corresponding structures are in Fig. C.7 in Appendix C.6.

All states A_1, \dots, A_4 are directly observed at low temperature as shown in Fig. 7.11d. Since the chiral field induced by the acid anion is different in (+)-**HG** and (-)-**HG** (due to the absence of mirror symmetry of the guest), Larmor frequency of state A_1 differs from that of A_2 , although Larmor frequencies of A_3 and A_4 are coincidentally similar. The presence of (*rac*)-CSA does not alter the corresponding chemical shifts compared to (*R*)-CSA. At room temperature, only averaged states A'_I and A'_{II} are detected due to fast exchange. We were not able to unequivocally assign states A_1, \dots, A_4 (corresponding to structures in Fig. 7.10c) to particular resonances in Fig. 7.11d. However, this information is not essential for analyzing the system's kinetics. Therefore, we assumed that Larmor frequency of the reference NH spin (denoted by green color in Fig. 7.10c) is likely similar when the protonation is in its vicinity (i.e. states A_1 and A_2) contrary to protonation across the symmetry plane (i.e. states A_3 and A_4), which results in assignment shown in Fig. 7.11d.

The observed spectral behavior can be described (in terms of reduced equivalent scheme from Sec. 3.3.8) as the half-symmetric three-state exchange with states A'_I , A'_{II} and C' (Table 3.2b), which is characterized by two independent transition rate coefficients $k_{A'}$ and $k_{A'C'}$ ($k_{A'C'} = k_{A'C'} p_{C'}/p_{A'_I}$ is not independent as follows from Eq. (3.37)). The corresponding populations are $p_{A'_I} = p_{A_1} + p_{A_2}$, $p_{A'_{II}} = p_{A_3} + p_{A_4}$ and $p_{C'} = p_C + p_D$. Resonance positions $\delta_{A'_I}$ and $\delta_{A'_{II}}$ shift upfield (to lower ppm values) at higher acid concentrations due to the increase in polarity of the medium upon addition of acid, see arrow in Fig. 7.12b. The resonance position $\delta_{C'}$ shifts downfield (to larger ppm values) during the titration (while $\delta_{A'_I}$ and $\delta_{A'_{II}}$ still remain constant), see blue arrow in Fig. 7.12b. This is a direct evidence that state C' consists of two substates C and D because the three-state solution in Eq. (3.52) does not allow shifts of peak maxima in the slow exchange regime. In the current model, the peak due to state C' can shift when populations p_C and p_D change ($\delta_{C'} = p_C \delta_C + p_D \delta_D$). It is already known that the host molecule can bind water with a binding constant $K_{\text{HW}} = 240 \pm 35 \text{ M}^{-1}$ [40]. During experiments it is not reasonably possible to prepare CDCl_3 solutions of **H** containing no residual water. Moreover, small quantities of water are added together with the hydrophilic guest during the titration process. Hence, the spin state D most likely corresponds to a host-water complex (**HW**). Competitive binding of guest and water is assumed [40], i.e. complexation with one ligand precludes binding of the other (see

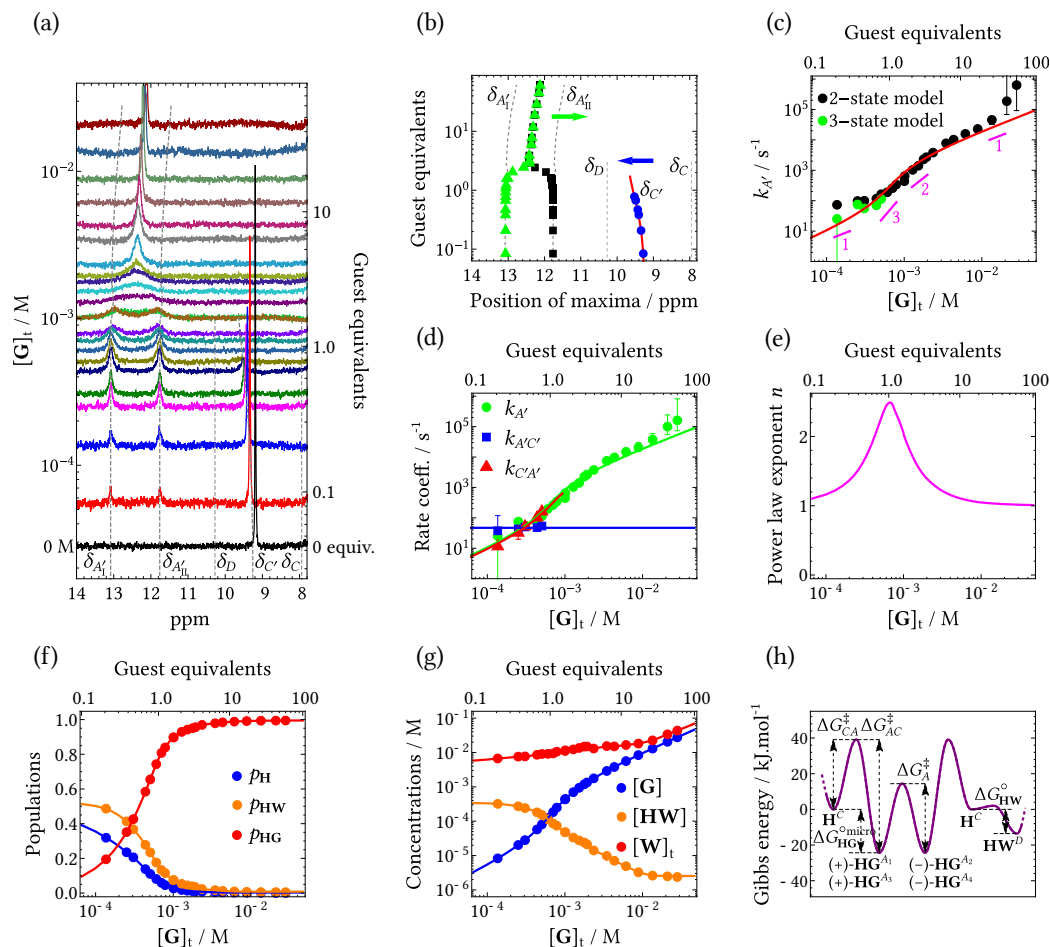


Figure 7.12: Experimental results for the host-guest system of di-bromobenzylated oxoporphyrinogen with (*R*)-camphorsulfonic acid. (a) NH portion of ¹H NMR spectra of host **H** (initial concentration 6×10^{-4} M, CDCl₃) during the titration with guest **G** (*y*-scaling of spectra is adjusted for clarity). Concentration of guest **G** corresponds to the value where the spectrum meets the *y*-axis. (b) Apparent positions of peak maxima during the titration, green and blue arrows denote the shift of maxima due to solvent polarity increase and fast exchange between states **C** (free **H**) and **D** (host-water complex **HW**), respectively. Red line is fit of the frequency $\omega_{C'} = p_C\omega_C + p_D\omega_D$ rescaled to ppm. (c) Concentration dependence of the transition rate coefficient $k_{A'}$ obtained using two-state lineshape fitting on states A'_I and A'_{II} (black circles) and using three-state lineshape fitting on states A'_I , A'_{II} and C' (green circles). The red line is the best fit of the $k_{A'}$ concentration dependence within three-state model (two-state model is used above 1 equiv. due to the disappearance of the **C** state from NMR spectra) using Eq. (7.3a) and (A.3). Magenta lines denote the slope of the red fitting curve. (d) Other transition rate coefficients describing the half-symmetric three-state exchange. The parameter $k_{A'C'}$ was fitted and $k_{C'A'}$ was calculated ($k_{C'A'} = k_{A'C'}p_{C'}/p_{A'_I}$). (e) Concentration dependence of the power law exponent. (f) Concentration dependence of populations of host-related species. (g) Concentration dependence of free guest, host-water complex (both determined from the model in Sec. 2.2.3) and total concentration of water (as determined from water peak integration, red solid line is interpolation). (h) Gibbs energy profile of all chemical species as calculated from the Eyring equation, Eq. (2.20), from the reaction rate coefficients at $T = 298$ K and setting the transition probability $\eta = 1$. Barriers ΔG_{CD}^\ddagger and ΔG_{DC}^\ddagger were not determined since the corresponding exchange process was too fast. The Larmor frequencies of states A'_I and A'_{II} are not unequivocally assigned with respect to the structures in Fig. 7.10c. Error bars in (c) and (d) denote maximum errors, see discussion in Sec. C.7.

Sec. 2.2.3). Because of the high mobility of water (small molecule) and low value of K_{HW} we assume infinitely fast exchange between states *C* and *D*. To summarize, the chemical kinetics description of this system includes competitive host-guest and host-water binding. The observed spin exchange kinetics (for host NH signals) can be described using the half-symmetric three-state model. Fast exchange effectively reduces the number of states to three, enabling the use of the reduced equivalent scheme.

Figure 7.13 shows three different levels of description of chemical kinetics in this system, and the underlying equations are given in Appendix C.8. Figure 7.13a shows the simplified chemical kinetics of the processes involved assuming the presence of host molecule as three distinct chemical species, **H**, **HG** and **HW**. The simplified chemical kinetics scheme is sufficient to describe the competitive host-ligand binding characterized by equilibrium constants K_{HG} and K_{HW} , the corresponding equations are shown in Sec. 2.2.3. The constant K_{HG} accounts for the protonation of any of the four sites. The low-temperature spectrum of **Bz₂OxP** + (*R*)-CSA in Fig. 7.11d suggests that the sites are not equivalent since states *A*₁ and *A*₂ are not equally populated with intensity ratio 74:26. The populations cannot be determined at room temperature. However, we can treat the sites as equivalent (with equal populations) at room temperature with reasonable accuracy since the populations tend to equalize with increasing temperature. Hence, the reaction rate coefficient $4\kappa_{\text{ass}}$ in Fig. 7.13a contains an integer prefactor to account for the four equivalent protonation sites. Protonation of one particular site is characterized by $K_{\text{HG}}^{\text{micro}} = K_{\text{HG}}/4$ (cf. Table 2.2 in Sec. 2.2.4). The microscopic equilibrium constant $K_{\text{HG}}^{\text{micro}}$ is also equal to the ratio of reaction rate coefficients for the molecular processes of protonation and deprotonation, i.e. $K_{\text{HG}}^{\text{micro}} = \kappa_{\text{ass}}/\kappa_{\text{diss}}$.

In fact, the protonation forms two different tautomeric species (+)-**HG** and (-)-**HG** as discussed above, see Fig. 7.10c. This situation is captured by the chemical kinetics scheme in Fig. 7.13b. Both tautomers can be formed in two different ways, hence the integer prefactor in $2\kappa_{\text{ass}}$. As discussed above, chemical species (+)-**HG** and (-)-**HG** can be interconverted by the

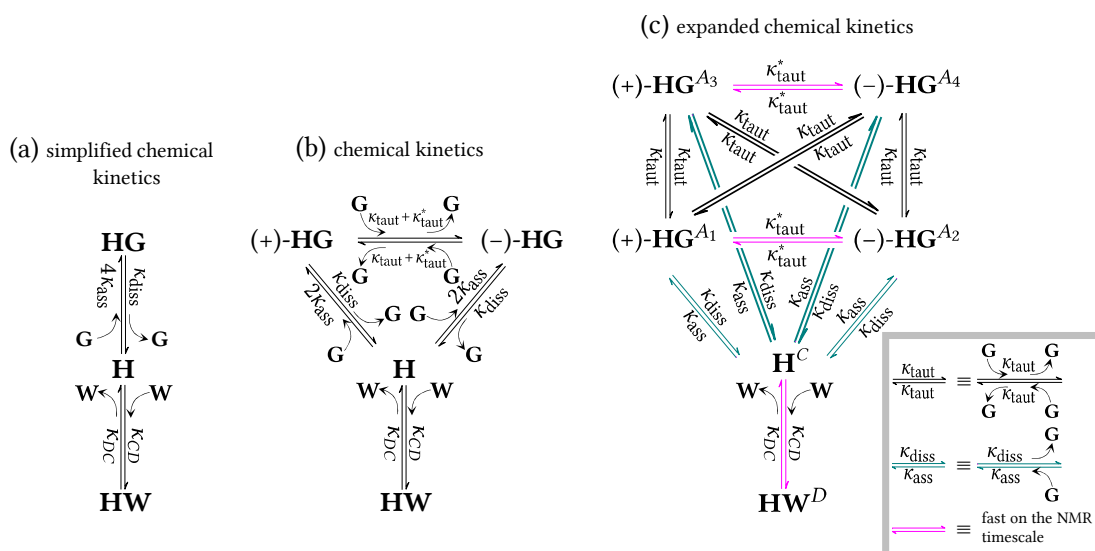


Figure 7.13: Chemical kinetics schemes for the multi-state system of **Bz₂OxP** (host **H**) in the presence of two ligands, (*R*)-CSA (ligand **G**) and water (ligand **W**). (a) Simplified chemical kinetics scheme corresponding to 1:1 **H**:**G** binding with competitive 1:1 **H**:**W** binding (see Sec. 2.2.3 for details). (b) Chemical kinetics scheme describing interconversion of all distinguishable chemical species. (c) Expanded chemical kinetics scheme, equal to the full spin kinetics scheme. All relevant molecular processes and their reaction rate coefficients are shown. Processes denoted by magenta arrows are fast on the NMR timescale.

prototropic tautomerization processes characterized by κ_{taut} or κ_{taut}^* . This interconversion between the two tautomeric forms can occur in two ways implying overall reaction rate coefficient $\kappa_{\text{taut}} + \kappa_{\text{taut}}^*$ (Fig. 7.13b). It is a second order reaction since the guest molecule has to collide with the host-guest complex. Details of these tautomerization processes, including structure of transition states, are given in Appendix C.9.

The chemical kinetics scheme can be further expanded to remove degeneracy from all microstates, see in Fig. 7.13c. Here, all chemically distinct species in all spin states (as listed in Fig. 7.10b,c) are shown as separate entities, so that no prefactors or sums of reaction rate coefficients are present for the reaction rate coefficients. The spin states are denoted in superscripts at the chemical species, e.g. \mathbf{H}^C , \mathbf{HG}^{A_1} . The processes characterized by κ_A^* (denoted by magenta arrows in Fig. 7.13c) have low energy barriers and therefore are fast at room temperature. This makes κ_A^* indeterminable by lineshape analysis. The expanded chemical kinetics scheme is equal to the full spin kinetics scheme and forms a basis for the connection between chemical kinetics and observed spin kinetics as explained in the next paragraph.

Due to fast exchange (large κ_{taut}^* , κ_{CD} and κ_{DC}), the kinetics scheme in Fig. 7.13c must be contracted from six to three states, which results in the spin kinetics scheme in Fig. 7.14a. Then this scheme expressed in reaction rate coefficients can be directly compared to the corresponding scheme expressed in transition rate coefficients in Fig. 7.14b, and subsequently, the relations between reaction and transition rate coefficients can be established. In addition, the concentration dependence of the transition rate coefficients is obtained. Details of this procedure are described in Appendix C.8. The resulting equations are

$$k_{A'} = 2\kappa_{\text{taut}}[\mathbf{G}], \quad (7.3a)$$

$$k_{A'C'} = \kappa_{\text{diss}}, \quad (7.3b)$$

$$k_{C'A'} = 2\kappa_{\text{ass}}[\mathbf{G}]. \quad (7.3c)$$

The transition rate coefficient $k_{A'C'}$, which corresponds to decay of **HG** complex, is independent of concentration. On the other hand, formation of a complex is a bimolecular reaction, therefore $k_{C'A'}$ is proportional to $[\mathbf{G}]$. Also, the guest-mediated prototropic tautomerization is a bimolecular reaction (based on the suggested reaction scheme in Appendix C.9), implying $k_{A'}$ is proportional to $[\mathbf{G}]$.

The use of three-state reduced equivalent spin kinetics scheme enables us to describe the observed spectra with only two transition rate coefficients $k_{A'}$ and $k_{A'C'}$. At guest concentrations higher than 1 equiv. the two-state model with transition rate coefficient $k_{A'}$ is sufficient because

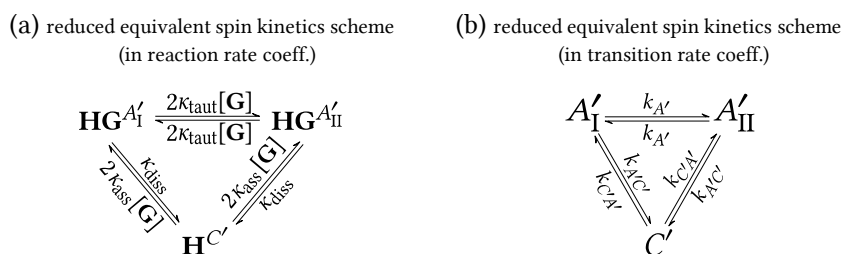


Figure 7.14: Reduced equivalent spin kinetics schemes for the multi-state system of **Bz₂OxP** (host **H**) in the presence of two ligands, (*R*)-CSA (ligand **G**) and water (ligand **W**). Schemes refer to the central NH protons of the host molecule. (a) Spin kinetics in terms of reaction rate coefficients as obtained from contraction of the scheme in Fig. 7.13c. (b) Corresponding spin kinetics in terms of transition rate coefficients. It has the form of half-symmetric three-state exchange. Comparison with (a) gives the relationship between transition and reaction rate coefficients in Eq. (7.3). This scheme represents experimentally observed spin kinetics. It was used for lineshape fitting, see Appendix C.7 for details.

the state C' is unpopulated. During the fitting procedure, the parameter $\delta_{C'}$, which describes the averaged frequency of the state C' , is fixed exactly at the corresponding peak position in the spectra, frequencies of the states A'_I and A'_{II} are also fixed, for details see Appendix C.8. To show the actual usefulness of the three-state model, we have also fitted the states A'_I and A'_{II} with the two-state model over the concentration range studied (fitting parameter $k_{A'}$). From Fig. 7.12c it is obvious that the three-state model systematically shifts $k_{A'}$ to lower values than obtained using the two-state model (presence of the third state C' causes broadening of the peaks due to states A'_I and A'_{II}). Examples of actual fitted spectra are shown in Fig. C.8 in Appendix C.7.

After the values of $k_{A'}$ and $k_{A'C'}$ were obtained from the raw spectra (green and blue points in Fig. 7.12d), fitting of the concentration dependence of $k_{A'}$ (using Eq. (7.3a)) was conducted simultaneously with other changes in spectra. This is shown together with other technical details of the fitting procedure in Appendix C.7. The resulting fitted curve for $k_{A'}$ (red line in Fig. 7.12c) describes the experimental values from the three-state exchange model with good accuracy and confirms the assumption of proportionality to $[G]$ for the guest-mediated prototropic tautomerization, expressed in Eq. (7.3a). The value of equilibrium constant $K_{HG} = (7.4 \pm 3.0) \times 10^4 \text{ M}^{-1}$ was also obtained. This value is comparable with the results of our previous analysis from Sec. 6.9 $K_{HG} = (8.0 \pm 5.0) \times 10^4 \text{ M}^{-1}$ (obtained from NMR binding isotherms) or $K_{HG} = (5.0 \pm 1.0) \times 10^4 \text{ M}^{-1}$ (obtained from UV-vis), which did not take into account the competitive binding of water and the concentration dependence of $k_{A'}$. The concentration dependence of $k_{C'A'}$ was determined using the formula $k_{C'A'} = 2\kappa_{\text{ass}}[G] = K_{HG}k_{A'C'}[G]/2$, see Fig. 7.12d. The value of $k_{C'A'}$ is calculated directly from fitted values of $k_{A'C'}$ up to 1 equiv. (red points in Fig. 7.12d). For higher guest concentrations, the population of state C' is negligible, and values of $k_{A'C'}$ could not be determined by three-state lineshape fitting. The red line in Fig. 7.12d is an extrapolation from the guest concentration dependence using the mean value of $k_{A'C'}$ (blue line). The fitting procedure also provides concentrations of all species present in the sample, see Fig. 7.12f,g. Values of equilibrium constants and reaction rate coefficients are listed in Table 7.2 and an overview of all parameters used during the fitting procedure is given in Table C.3.

The concentration dependence of $k_{A'}$ can also be expressed in the form of a power law as $k_{A'} \propto [G]_t^n$ (for constant $[H]_t$). The exponent n can easily be extracted from the log-log plot in Fig. 7.12c as the gradient (first derivative) of the red curve. The concentration dependence of the power law exponent is shown in Fig. 7.12e. It can be seen that for low and high guest concentrations $n = 1$. However, between these limit cases, the power law exponent reaches values over $n = 3$. It is interesting to mention that the $k_{A'}$ dependence on free guest concentration $[G]$ has a simple linear relationship (see Eq. (7.3a)) while its dependence on total guest concentration $[G]_t$ has a nonlinear form with the largest deviation from linearity around 1 equiv. of total guest concentration (Fig. 7.12e).

Table 7.2: Parameters of Gibbs energy profile at $T = 298 \text{ K}$. In the Eyring equation (Eq. (2.20)), the assumption for transition probability $\eta = 1$ was used.

equilibrium constants		std. Gibbs energy of reaction		mutual relationship
K_{HG}	$(7.4 \pm 3.0) \times 10^4 \text{ M}^{-1}$	ΔG_{HG}°	$-27.8 \pm 1.0 \text{ kJ.mol}^{-1}$	$\Delta G_{HG}^\circ = -RT \ln K_{HG}$
K_{HG}^{micro}	$(1.9 \pm 0.8) \times 10^4 \text{ M}^{-1}$	$\Delta G_{HG}^{\circ \text{micro}}$	$-24.3 \pm 1.0 \text{ kJ.mol}^{-1}$	$\Delta G_{HG}^{\circ \text{micro}} = -RT \ln K_{HG}^{\text{micro}}$
K_{HW}	$240 \pm 35 \text{ M}^{-1}$	ΔG_{HW}°	$-13.6 \pm 0.4 \text{ kJ.mol}^{-1}$	$\Delta G_{HW}^\circ = -RT \ln K_{HW}$
reaction rate coefficients		Gibbs energy of barrier		mutual relationship
κ_{diss}	$47 \pm 6 \text{ s}^{-1}$	$\Delta G_{\text{diss}}^\ddagger$	$63.4 \pm 0.3 \text{ kJ.mol}^{-1}$	$\Delta G_{\text{diss}}^\ddagger = -RT \ln \frac{h\kappa_{\text{diss}}}{k_B T}$
κ_{ass}	$(9 \pm 4) \times 10^5 \text{ M}^{-1}\text{s}^{-1}$	$\Delta G_{\text{ass}}^\ddagger$	$39.0 \pm 1.1 \text{ kJ.mol}^{-1}$	$\Delta G_{\text{ass}}^\ddagger = -RT \ln \frac{h\kappa_{\text{ass}}}{k_B T}$
κ_{taut}	$(10 \pm 1) \times 10^5 \text{ M}^{-1}\text{s}^{-1}$	$\Delta G_{\text{taut}}^\ddagger$	$38.7 \pm 0.2 \text{ kJ.mol}^{-1}$	$\Delta G_{\text{taut}}^\ddagger = -RT \ln \frac{h\kappa_{\text{taut}}}{k_B T}$

The chemical kinetics, according to Fig. 7.13b, can also be viewed in terms of the Gibbs energy landscape as shown in Fig. 7.12h. The energy barriers are calculated from the corresponding reaction rate coefficients using the Eyring equation, Eq. (2.20), at $T = 298$ K with the transition probability $\eta = 1$. Standard reaction Gibbs energies for both **HG** and **HW** complexes were calculated from the equilibrium constants $K_{\text{HG}}^{\text{micro}}$ and K_{HW} , respectively. All parameters of the Gibbs energy profile are listed in Table 7.2. The barrier between states *C* and *D* was not determined since it is very low (i.e. fast exchange regime).

7.3 Dynamic processes – summary

Multiple processes have been identified in **OxP** and **Bz₂OxP**, which is summarized and schematically illustrated in Table 7.4 and 7.5, respectively. Only one variant out of several possible ones is shown. NMR spectra of **Bz₄OxP** in the presence of DFA do not show any dynamic behavior since the tautomeric processes are fast (OH signal due to protonated **Bz₄OxP** is broadened and thus not observable), anion binding at the center is prevented by N-alkylation, and observation of topomerization is impossible due to high symmetry of **Bz₄OxP**. Our analyses quantified some of these processes in terms of transition rate coefficients (summarized in Table 7.3), reaction rate coefficients and energy barrier heights. These results were obtained in two steps: (i) lineshape fitting of NMR spectra to obtain transition rate coefficients and (ii) fitting of binding and exchange models to obtain reaction rate coefficients, binding constants and other parameters included in binding isotherms (see Sec. 3.4.2). These steps are described in detail in Appendix C.7 for the analysis of NH signals in **Bz₂OxP** + (*R*)-CSA titration.

The first process in Table 7.3 is dynamic host-guest binding. This process could not be accurately quantified in **OxP** due to the slow exchange regime, but the corresponding transition rate coefficients were determined in **Bz₂OxP** from NH resonances (see Sec. 7.2.3). Dissociation of the complex is unimolecular reaction, thus, k_{diss} does not depend on concentration, while k_{ass} representing a bimolecular reaction is proportional to $[\text{G}]$. The anion exchange process k_{ex} changes the host spin states only when chiral and not enantiopure guest is present in the sample. This condition was satisfied for **Bz₂OxP** + CSA systems, which revealed fast anion exchange at acid excess (see Sec. 7.2.2 or [45]). Prototropic tautomerization is mediated by free guest molecules, the concentration dependence was measured in both **OxP** and **Bz₂OxP**. In **Bz₂OxP** + (*R*)-CSA system, the concentration dependence was successfully described by a kinetics model (see Sec. 7.2.3 or our article [2]). Macrocyclic inversion is an unexpected dynamic process taking part in **OxP** in the presence of mandelic acid [36]. However, this process was not observed in the **OxP** + CSA system probably due to strong host-guest binding, which prevented formation of the neutral flip-active form. In the presence of DFA it cannot be decided whether this process takes place because the host spin states are not changed due to mirror symmetry of DFA (see Sec. 7.1.3). Finally, rotation of the bulky side groups referred to as hemiquinonoid topomerization was detected in

Table 7.3: Overview of transition rate coefficients. Question mark stands for the inability to obtain the corresponding parameter from current experiments.

process	OxP		Bz₂OxP	
	DFA	CSA	DFA	CSA
host-guest bind.	$k_{\text{ass}}, k_{\text{diss}} < 5 \text{ s}^{-1}$	slow	intermediate	$k_{\text{diss}} = 47 \text{ s}^{-1}, k_{\text{ass}} < 200 \text{ s}^{-1}$
anion exchange	undetectable	?	undetectable	fast at acid excess
prototrop. tautomer.	$k_{\text{taut}} < 10^5 \text{ s}^{-1}$	$k_{\text{taut}} < 2000 \text{ s}^{-1}$?	$k_{\text{taut}} < 2 \times 10^5 \text{ s}^{-1}$
macrocyc. inv.	undetectable	not occurring	impossible	impossible
hemiquin. topomer.	undetectable	undetectable	$k_{\text{topo}} < 150 \text{ s}^{-1}$	slow

Bz₂OxP. This process is activated by protonation, and thus shows complicated dependence on concentration when host-guest complex with higher than 1:1 stoichiometry is formed (see Sec. 7.2.1).

Noteworthy, high symmetry of **OxP** in combination with mirror symmetry of DFA hinders the observability of some processes because the corresponding nuclear spins undergo chemical exchange between identical chemical environments (first column of Table 7.3). On the other hand, N-alkylation lowers the symmetry of **Bz₂OxP**, which in combination with asymmetric (*R*)-CSA allows, in theory, observation of all processes (last column of Table 7.3).

Table 7.4: Schematic overview of all processes in **OxP**. When the processes can be realized by multiple ways, only one instance is shown as illustration. Transition states are not shown.

process	illustration	notes
host-guest binding		in the binding step $H \rightarrow HG_2$
host-guest binding at acid excess		in the binding step $HG_2 \rightarrow HG_4$
anion exchange		shares transition states with prototropic tautomerization; in HG_2
anion exchange at acid excess		either the center anion or the guest at the edge can be exchanged; in HG_4
prototropic tautomerization		mediated by either (i) free guest, (ii) water or (iii) host-host collisions; involved temperature dependence; in HG_2
prototropic tautomerization at acid excess		in HG_4
macrocyclic inversion		activated by the acid binding; in HG_2

Table 7.5: Schematic overview of all processes in **Bz₂OxP**. When the processes can be realized by multiple ways, only one instance is shown as illustration. Proposed transition states of prototropic tautomerization and anion exchange are shown in Sec. C.9.

process	illustration	notes
host-guest binding		in the binding step H → HG
host-guest binding at acid excess		in the binding step HG → HG₂ or HG₂ → HG₄
anion exchange		shares transition states with prototropic tautomerization; in HG
anion exchange at acid excess		either center anion or the guest at the edge can be exchanged; in HG₂ or HG₄
prototropic tautomerization		in HG
prototropic tautomerization at acid excess		in HG₂ or HG₄
hemiquinonoid topomerization		rotation around single bond; activated by protonation; in HG , HG₂ or HG₄

8 Conclusion

OxP, **Bz₂OxP** and **Bz₄OxP** serve as colorimetric sensors for acids in nonpolar media in wide range of acid volume fractions. The halochromic response of **OxP** ($\varphi(\text{DFA}) \in [10^{-6}, 10^{-3}]$) is very sensitive to small amounts of acid (detects ca. 1 ppm of DFA), however, it also saturates at relatively low acid contents. The color response is shifted to higher acid contents in **Bz₂OxP** ($\varphi(\text{DFA}) \in [10^{-4}, 1]$) and **Bz₄OxP** ($\varphi(\text{DFA}) \in [0.01, 0.6]$) due to N-alkylation (Fig. 6.6). The color changes were permitted by protonation and subsequent anion binding (i.e. host-guest binding), which was studied by UV/vis and NMR spectroscopy in Chapter 6. UV/vis spectra were analyzed using SVD to determine the number of absorbing components and using DSIS to obtain the components' populations. Up to fourfold protonation in **OxP** and **Bz₂OxP**, and up to twofold protonation in **Bz₄OxP** was observed, all UV/vis spectra of the individual protonated species were obtained (Fig. 6.7 and Fig. 6.19). Experimental binding isotherms were constructed from both UV/vis and NMR spectroscopic data and fitted with appropriate host-guest binding models (applicable at $\varphi(\text{DFA}) < 0.001$) to obtain binding constants (Table 6.1). It turns out that N-alkylation in **Bz₂OxP** and **Bz₄OxP** reduces the binding affinity in two ways: (i) it blocks the binding sites for acid anions at the porphyrinogen center and (ii) prevents redistribution of charge, which would stabilize the protonated species. Upon host protonation, the acid anions are located near the charged complex because the nonpolar solvent does not allow for full dissociation.

Solvatochromic effects were studied in DMF/chloroform mixture. High contents of DMF increase the solvent polarity and induce hydrogen bonding of DMF to the oxoporphyrinogen host, causing positive solvatochromic shifts in UV/vis spectra of all investigated oxoporphyrinogens. Analogously to halochromism, solvatochromic response threshold is shifted to higher DMF concentrations with increasing degree of N-alkylation (Fig. 6.15). The spectral changes in **OxP** and **Bz₂OxP** could be modeled as interconversion of individual absorbing species, which was further confirmed by SVD. Besides solvatochromic shifts, **OxP** showed transition to another form with distinct absorption bands assigned to porphodimethene form, which was stable only at particular DMF concentration range. Detailed summary of chromic and binding properties is given in Sec. 6.9.

Various molecular dynamic processes are manifested as chemical exchange in NMR spectra of protonated **OxP** and **Bz₂OxP** as shown in Chapter 7. In both oxoporphyrinogens, we identified the processes of anion exchange and prototropic tautomerization, also the host-guest binding was confirmed to be a dynamic process (Fig. 7.4 and 7.5). Protonation also activates rotation of bulky side groups (i.e. process of hemiquinonoid topomerization) as observed in **Bz₂OxP**. This process cannot be detected in **OxP** or **Bz₄OxP** due to their high symmetry. NMR spectra of protonated **Bz₄OxP** do not show any chemical exchange due to its symmetry and fast rate of the molecular processes.

To enable fitting of three-state chemical exchange lineshapes in NMR spectra (in the absence of J-coupling), we calculated the analytical formula, which has not been published before (Sec. 3.3.7). We also devised so-called 'reduced equivalent schemes' for exchange schemes containing a fast-exchanging state in order to reduce the number of parameters for lineshape fitting (Sec. 3.3.8). Furthermore, we derived an approximative formula to determine the coalescence point in symmetric two-state exchange with nonzero transverse relaxation rate (Eq. 3.20).

Using these theoretical instruments, temperature and concentration dependence of transition rate coefficients corresponding to the present molecular processes was obtained. Effect of temperature on the topomerization process in protonated **Bz₂OxP** obeys Eyring equation (Table 7.1), however, transition rate coefficient of tautomerization in protonated **OxP** shows non-

monotonous temperature dependence. Concentration dependence of transition rate coefficients in **Bz₂OxP** systems has been described in detail using a kinetic model derived from competitive host-guest binding (Fig. C.11 and Table 7.2). Note that chirality of the acid influences the number of exchanging states in NMR spectra because the bound acid anion determines overall symmetry of the complex through the so-called chiral field. Detailed summary of dynamic processes in the oxoporphyrinogens is given in Sec. 7.3.

OxP and its derivatives have potential practical applications as colorimetric sensors in non-polar media. N-alkylation modulates the color response so that the full acid volume fraction ($\varphi(\text{DFA}) \in [10^{-6}, 1]$) is covered by either of the compounds. Similar structure of the compounds assures matching solubility, thus, these sensors might be applied as blends analogously to the Universal Indicator [140].

Our NMR study of dynamic processes in the oxoporphyrinogens contributes to understanding the concentration dependence of transition rate coefficients, which is always system-specific. In this regard, models derived from chemical kinetics proved to be useful. We have also shown the advantage of reduced equivalent schemes, derived from analytical exchange lineshapes. The presented methods can be used to describe molecular kinetics in a wide range of systems with a variety of intra- or intermolecular processes, such as those involving host-guest binding, enzymatic reactions or nontrivial conformational dynamics of proteins or other complex molecules.

References

- [1] V. Březina, S. Ishihara, J. Lang, L. Hanyková, K. Ariga, J. P. Hill, and J. Labuta, “Structural Modulation of Chromic Response: Effects of Binding-Site Blocking in a Conjugated Calix[4]pyrrole Chromophore”, [ChemistryOpen](#) **7**, 323–335 (2018).
- [2] V. Březina, L. Hanyková, N. Velychkivska, J. P. Hill, and J. Labuta, “NMR lineshape analysis using analytical solutions of multi-state chemical exchange with applications to kinetics of host–guest systems”, [Sci. Rep.](#), **10**. 1038/s41598-022-20136-4 (2022).
- [3] G. P. Moss, “Nomenclature of tetrapyrroles (Recommendations 1986)”, [Pure and Applied Chemistry](#) **59**, 779–832 (1987).
- [4] R. Meganathan, *Comprehensive Natural Products II: Chemistry and Biology*, edited by L. Mander and H.-W. Liu, Vol. 7 (Elsevier, Mar. 2010).
- [5] P. R. Ortiz de Montellano, “Hemes in Biology”, in [Wiley Encyclopedia of Chemical Biology](#) (American Cancer Society, 2008), pp. 1–10.
- [6] I. A. Rather, S. A. Wagay, M. S. Hasnain, and R. Ali, “New dimensions in calix[4]pyrrole: the land of opportunity in supramolecular chemistry”, [RSC Adv.](#) **9**, 38309–38344 (2019).
- [7] L. W. Schultz, L. Liu, M. Cegielski, and J. W. Hastings, “Crystal structure of a pH-regulated luciferase catalyzing the bioluminescent oxidation of an open tetrapyrrole”, [PNAS](#) **102**, 1378–1383 (2005).
- [8] M. Valiadi and D. Iglesias-Rodriguez, “Understanding Bioluminescence in Dinoflagellates – How Far Have We Come?”, [Microorganisms](#) **1**, 3–25 (2013).
- [9] C. C. Hsia, “Respiratory Function of Hemoglobin”, [New England Journal of Medicine](#) **338**, 239–248 (1998).
- [10] M. Taniguchi, J. S. Lindsey, D. F. Bocian, and D. Holten, “Comprehensive review of photophysical parameters (ϵ , Φ_f , τ_s) of tetraphenylporphyrin (H₂TPP) and zinc tetraphenylporphyrin (ZnTPP) – Critical benchmark molecules in photochemistry and photosynthesis”, [Journal of Photochemistry and Photobiology C: Photochemistry Reviews](#) **46**, 100401 (2021).
- [11] J. Lee, K. T. Crampton, N. Tallarida, and V. A. Apkarian, “Visualizing vibrational normal modes of a single molecule with atomically confined light”, [Nature](#) **568**, 78–82 (2019).
- [12] P. A. Gale, J. L. Sessler, and V. Král, “Calixpyrroles”, [Chem. Commun.](#), 1–8 (1998).
- [13] L. R. Milgrom, “The facile aerial oxidation of a porphyrin”, [Tetrahedron](#) **39**, 3895–3898 (1983).
- [14] P. A. Castelfranco and S. I. Beale, “Chlorophyll Biosynthesis: Recent Advances and Areas of Current Interest”, [Annual Review of Plant Physiology](#) **34**, 241–276 (1983).
- [15] M. O. Senge, A. A. Ryan, K. A. Letchford, S. A. MacGowan, and T. Mielke, “Chlorophylls, Symmetry, Chirality, and Photosynthesis”, [Symmetry](#) **6**, 781–843 (2014).

- [16] R. R. Eitenmiller and W. O. L. Jr, *Vitamin Analysis for the Health and Food Sciences* (CRC Press, Dec. 2010).
- [17] D. Monti, S. Nardis, M. Stefanelli, R. Paolesse, C. Di Natale, and A. D'Amico, "Porphyrin-Based Nanostructures for Sensing Applications", *Journal of Sensors* **2009**, e856053 (2009).
- [18] H. Lee, K.-I. Hong, and W.-D. Jang, "Design and applications of molecular probes containing porphyrin derivatives", *Coord. Chem. Rev., Recent Progress on Fluorescent Probes* **354**, 46–73 (2018).
- [19] S. Ishihara, J. Labuta, W. V. Rossom, D. Ishikawa, K. Minami, J. P. Hill, and K. Ariga, "Porphyrin-based sensor nanoarchitectonics in diverse physical detection modes", *Physical Chemistry Chemical Physics* **16**, 9713–9746 (2014).
- [20] A. C. Sedgwick, T. T. Brewster, T. Wu, X. Feng, S. D. Bull, X. Qian, J. T. Sessler, T. D. James, E. V. Anslyn, and X. Sun, "Indicator displacement assays (IDAs): the past, present and future", *Chemical Society Reviews* **50**, 9–38 (2021).
- [21] P. A. Gale, P. Anzenbacher Jr., and J. L. Sessler, "Calixpyrroles II", *Coordination Chemistry Reviews* **222**, 57–102 (2001).
- [22] K.-H. Pfoertner and T. Oppenländer, "Photochemistry", in *Ullmann's Encyclopedia of Industrial Chemistry* (American Cancer Society, 2012).
- [23] H. Watarai and Y. Kurahashi, "Chiral Recognition of 2-Alkylalcohols with Magnetic Circular Dichroism Measurement of Porphyrin J-Aggregate on Silica Gel Plate", *Anal. Chem.* **88**, 4619–4623 (2016).
- [24] J. M. Merkes, M. Rueping, F. Kiessling, and S. Banala, "Photoacoustic Detection of Superoxide Using Oxoporphyrinogen and Porphyrin", *ACS Sens.* **4**, 2001–2008 (2019).
- [25] J. M. Merkes, L. Zhu, S. B. Bahukhandi, M. Rueping, F. Kiessling, and S. Banala, "Photoacoustic Imaging Probes Based on Tetrapyrroles and Related Compounds", *International Journal of Molecular Sciences* **21**, 3082 (2020).
- [26] J. Kou, D. Dou, and L. Yang, "Porphyrin photosensitizers in photodynamic therapy and its applications", *Oncotarget* **8**, 81591–81603 (2017).
- [27] J. Tian, B. Huang, M. H. Nawaz, and W. Zhang, "Recent advances of multi-dimensional porphyrin-based functional materials in photodynamic therapy", *Coordination Chemistry Reviews* **420**, 213410 (2020).
- [28] J. Helaja, M. Stapelbroek-Möllmann, I. Kilpeläinen, and P. H. Hynninen, "NH Tautomerism in the Natural Chlorin Derivatives", *J. Org. Chem.* **65**, 3700–3707 (2000).
- [29] C. B. Storm and Y. Teklu, "Nitrogen-hydrogen tautomerism in porphyrines and chlorines", *J. Am. Chem. Soc.* **94**, 1745–1747 (1972).
- [30] J. P. Hill, A. L. Schumacher, F. D'Souza, J. Labuta, C. Redshaw, M. R. J. Elsegood, M. Aoyagi, T. Nakanishi, and K. Ariga, "Chromogenic Indicator for Anion Reporting Based on an N-Substituted Oxoporphyrinogen", *Inorg. Chem.* **45**, 8288–8296 (2006).
- [31] A. Shundo, S. Ishihara, J. Labuta, Y. Onuma, H. Sakai, M. Abe, K. Ariga, and J. P. Hill, "Colorimetric visualization of acid–base equilibria in non-polar solvent", *Chem. Commun.* **49**, 6870–6872 (2013).
- [32] M. K. Chahal, J. Labuta, V. Březina, P. A. Karr, Y. Matsushita, W. A. Webre, D. T. Payne, K. Ariga, F. D'Souza, and J. P. Hill, "Knock-on synthesis of tritopic calix[4]pyrrole host for enhanced anion interactions", *Dalton Trans.*, 10.1039/C9DT02365H (2019).

- [33] M. K. Chahal, N. Velychkivska, W. A. Webre, J. Labuta, S. Ishihara, K. Ariga, F. D'Souza, and J. P. Hill, "Increasing the complexity of oxoporphyrinogen colorimetric sensing chromophores: N-alkylation and β -substitution", *J. Porphyrins Phthalocyanines* **23**, 1184–1194 (2019).
- [34] S. Ishihara, J. Labuta, T. Šikorský, J. V. Burda, N. Okamoto, H. Abe, K. Ariga, and J. P. Hill, "Colorimetric detection of trace water in tetrahydrofuran using N,N'-substituted oxoporphyrinogens", *Chem. Commun.* **48**, 3933–3935 (2012).
- [35] S. Ishihara, N. Iyi, J. Labuta, K. Deguchi, S. Ohki, M. Tansho, T. Shimizu, Y. Yamauchi, P. Sahoo, M. Naito, H. Abe, J. P. Hill, and K. Ariga, "Naked-Eye Discrimination of Methanol from Ethanol Using Composite Film of Oxoporphyrinogen and Layered Double Hydroxide", *ACS Appl. Mater. Interfaces* **5**, 5927–5930 (2013).
- [36] J. Labuta, Z. Futera, S. Ishihara, H. Kouřilová, Y. Tateyama, K. Ariga, and J. P. Hill, "Chiral Guest Binding as a Probe of Macrocycle Dynamics and Tautomerism in a Conjugated Tetrapyrrole", *J. Am. Chem. Soc.* **136**, 2112–2118 (2014).
- [37] J. Labuta, S. Ishihara, K. Ariga, and J. P. Hill, "Dynamic Processes in Prochiral Solvating Agents (pro-CSAs) Studied by NMR Spectroscopy", *Symmetry* **6**, 345–367 (2014).
- [38] A. Shundo, J. Labuta, J. P. Hill, S. Ishihara, and K. Ariga, "Nuclear Magnetic Resonance Signaling of Molecular Chiral Information Using an Achiral Reagent", *J. Am. Chem. Soc.* **131**, 9494–9495 (2009).
- [39] J. Labuta, S. Ishihara, A. Shundo, S. Arai, S. Takeoka, K. Ariga, and J. P. Hill, "Chirality Sensing by Nonchiral Porphines", *Chemistry – A European Journal* **17**, 3558–3561 (2011).
- [40] J. Labuta, S. Ishihara, T. Šikorský, Z. Futera, A. Shundo, L. Hanyková, J. V. Burda, K. Ariga, and J. P. Hill, "NMR spectroscopic detection of chirality and enantiopurity in referenced systems without formation of diastereomers", *Nat Commun* **4**, 2188 (2013).
- [41] S. Ishihara, J. Labuta, Z. Futera, S. Mori, H. Sato, K. Ariga, and J. P. Hill, "NMR Spectroscopic Determination of Enantiomeric Excess Using Small Prochiral Molecules", *J. Phys. Chem. B* **122**, 5114–5120 (2018).
- [42] J. Labuta, S. Ishihara, D. T. Payne, K. Takimoto, H. Sato, L. Hanyková, K. Ariga, and J. P. Hill, "Estimation of Enantiomeric Excess Based on Rapid Host–Guest Exchange", *Chemosensors* **9**, 259 (2021).
- [43] K. Norvaiša, J. E. O'Brien, I. Osadchuk, B. Twamley, V. Borovkov, and M. O. Senge, "Importance of molecular symmetry for enantiomeric excess recognition by NMR", *Chemical Communications* **58**, 5423–5426 (2022).
- [44] K. Takimoto, S. Ishihara, J. Labuta, V. Březina, D. T. Payne, J. P. Hill, K. Ariga, M. Sumita, S. Mori, and H. Sato, "Enantiomeric Excess Dependent Splitting of NMR Signal through Dynamic Chiral Inversion and Coligand Exchange in a Coordination Complex", *J. Phys. Chem. Lett.*, 8164–8169 (2020).
- [45] V. Březina, "Spectroscopic and theoretical study of supramolecular complexes of symmetric porphyrins with chiral guests", MA thesis (Univerzita Karlova, Matematicko-fyzikální fakulta, May 2014).
- [46] P. W. Atkins, J. D. Paula, and J. Keeler, *Atkins' Physical Chemistry* (Oxford University Press, 2018).
- [47] D. A. McQuarrie and J. D. Simon, *Physical Chemistry: A Molecular Approach* (University Science Books, Aug. 1997).

- [48] K. A. Connors, *Binding constants: the measurement of molecular complex stability* (Wiley, May 1987).
- [49] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics* (Wiley, Sept. 1985).
- [50] L. Lue, *Chemical thermodynamics* (Ventus Publishing, Place of publication not identified, 2009).
- [51] J. Gmehling, B. Kolbe, M. Kleiber, and J. Rarey, *Chemical Thermodynamics for Process Simulation* (John Wiley & Sons, Mar. 2012).
- [52] R. Srivastava and B. D. Smith, "Total pressure vapor-liquid equilibrium data for benzene + acetonitrile, diethylamine + ethylacetate, and propylamine + diethylamine binary systems", *J. Chem. Eng. Data* **31**, 94–99 (1986).
- [53] T. Hiak, K. Kurihara, and K. Kojima, "Vapor-Liquid Equilibria for Acetone + Chloroform + Methanol and Constituent Binary Systems at 101.3 kPa", *J. Chem. Eng. Data* **39**, 714–719 (1994).
- [54] R. Aveyard, B. J. Briscoe, and J. Chapman, "Activity coefficients and association of n-alkanols in n-octane", *Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases* **69**, 1772–1778 (1973).
- [55] H. Renon and J. M. Prausnitz, "Local compositions in thermodynamic excess functions for liquid mixtures", *AIChE Journal* **14**, 135–144 (1968).
- [56] *Dortmund Data Bank*, <http://www.ddbst.com/>.
- [57] K. J. Mysels, "Textbook errors: VII. The laws of reaction rates and of equilibrium", *J. Chem. Educ.* **33**, 178 (1956).
- [58] E. A. Guggenheim, "Textbook errors: IX. More about the laws of reaction rates and of equilibrium", *J. Chem. Educ.* **33**, 544 (1956).
- [59] K. J. Laidler, "Symbolism and Terminology in Chemical Kinetics", *Pure and Applied Chemistry* **53**, 753–771 (1981).
- [60] P. J. Mulquiney and P. W. Kuchel, *Modelling metabolism with Mathematica: detailed examples including erythrocyte metabolism* (CRC Press, Boca Raton, Fla, 2003).
- [61] K. A. Connors, *Chemical Kinetics: The Study of Reaction Rates in Solution* (Wiley, Aug. 1990).
- [62] K. J. Laidler and M. C. King, "Development of transition-state theory", *J. Phys. Chem.* **87**, 2657–2664 (1983).
- [63] S. Glasstone, K. J. Laidler, and H. Eyring, *The Theory of Rate Processes: The Kinetics of Chemical Reactions, Viscosity, Diffusion and Electrochemical Phenomena* (McGraw-Hill Book Company, Incorporated, 1941).
- [64] V. L. Arcus, E. J. Prentice, J. K. Hobbs, A. J. Mulholland, M. W. Van der Kamp, C. R. Pudney, E. J. Parker, and L. A. Schipper, "On the Temperature Dependence of Enzyme-Catalyzed Rates", *Biochemistry* **55**, 1681–1688 (2016).
- [65] M. W. van der Kamp, E. J. Prentice, K. L. Kraakman, M. Connolly, A. J. Mulholland, and V. L. Arcus, "Dynamical origins of heat capacity changes in enzyme-catalysed reactions", *Nat. Commun.* **9**, 1177 (2018).
- [66] E. P. Kyba, R. C. Helgeson, K. Madan, G. W. Gokel, T. L. Tarnowski, S. S. Moore, and D. J. Cram, "Host-guest complexation. 1. Concept and illustration", *J. Am. Chem. Soc.* **99**, 2564–2571 (1977).

- [67] J. W. Steed and J. L. Atwood, *Supramolecular chemistry*, 2nd ed (Wiley, Chichester, UK, 2009).
- [68] J. T. Edsall, J. Wyman, and J. G. Hoffman, "Thermodynamics, Electrostatics, and the Biological Significance of the Properties of Matter", *Physics Today* **11**, 50–52 (1958).
- [69] I. Klotz and D. Hunston, "Protein interactions with small molecules. Relationships between stoichiometric binding constants, site binding constants, and empirical binding parameters.", *Journal of Biological Chemistry* **250**, 3001–3009 (1975).
- [70] K. A. Connors and D. D. Pendergast, "Microscopic binding constants in cyclodextrin systems: complexation of .alpha.-cyclodextrin with sym-1,4-disubstituted benzenes", *J. Am. Chem. Soc.* **106**, 7607–7614 (1984).
- [71] A. E. Hargrove, Z. Zhong, J. L. Sessler, and E. V. Anslyn, "Algorithms for the determination of binding constants and enantiomeric excess in complex host : guest equilibria using optical measurements", *New Journal of Chemistry* **34**, 348–354 (2010).
- [72] C. A. Hunter and H. L. Anderson, "What is Cooperativity?", *Angewandte Chemie International Edition* **48**, 7488–7499 (2009).
- [73] G. Ercolani, "Assessment of Cooperativity in Self-Assembly", *J. Am. Chem. Soc.* **125**, 16097–16103 (2003).
- [74] E. V. Anslyn and D. A. Dougherty, *Modern Physical Organic Chemistry* (University Science Books, 2006).
- [75] M. I. Stefan, S. J. Edelstein, and N. Le Novère, "Computing phenomenologic Adair-Klotz constants from microscopic MWC parameters", *BMC Systems Biology* **3**, 68 (2009).
- [76] M. H. Levitt, *Spin Dynamics: Basics of Nuclear Magnetic Resonance* (John Wiley & Sons, Nov. 2001).
- [77] J. Cavanagh, W. J. Fairbrother, A. G. Palmer, and N. J. Skelton, *Protein NMR Spectroscopy: Principles and Practice* (Academic Press, 1996).
- [78] R. S. Macomber, *A complete introduction to modern NMR spectroscopy* (Wiley, 1998).
- [79] C. P. Slichter, *Principles of Magnetic Resonance* (Springer Science & Business Media, Mar. 1996).
- [80] B. Barman and A. Petrou, "Measuring the magnetization of a permanent magnet", *American Journal of Physics* **87**, 275–278 (2019).
- [81] H. Günther, *NMR Spectroscopy: Basic Principles, Concepts and Applications in Chemistry* (John Wiley & Sons, Dec. 2013).
- [82] R. E. Hoffman, "Variations on the chemical shift of TMS", *Journal of Magnetic Resonance* **163**, 325–331 (2003).
- [83] F. Bloch, "Nuclear Induction", *Phys. Rev.* **70**, 460–474 (1946).
- [84] F. Bloch, W. W. Hansen, and M. Packard, "The Nuclear Induction Experiment", *Phys. Rev.* **70**, 474–485 (1946).
- [85] D. J. Griffiths, *Introduction to electrodynamics*, Fourth edition (Pearson, Boston, 2013).
- [86] D. Antcliffe and A. C. Gordon, "Metabonomics and intensive care", *Critical Care* **20**, 68 (2016).
- [87] A. D. Bain and G. Duns, "A unified approach to dynamic NMR based on a physical interpretation of the transition probability", *Can. J. Chem.* **74**, 819–824 (1996).

- [88] P. S. Belton, "Chapter 17 Continuous wave and rapid scan correlation NMR", in *Data Handling in Science and Technology*, Vol. 18, edited by D. N. Rutledge, Signal Treatment and Signal Analysis in NMR (Elsevier, Jan. 1996), pp. 362–373.
- [89] Z. Szalay and J. Rohonczy, "Monte Carlo simulation of NMR lineshapes in chemically exchanging spin systems", *Progress in Nuclear Magnetic Resonance Spectroscopy* **56**, 198–216 (2010).
- [90] O. Anamimoghadam, L. O. Jones, J. A. Cooper, Y. Beldjoudi, M. T. Nguyen, W. Liu, M. D. Krzyaniak, C. Pezzato, C. L. Stern, H. A. Patel, M. R. Wasielewski, G. C. Schatz, and J. F. Stoddart, "Discrete Open-Shell Tris(bipyridinium radical cationic) Inclusion Complexes in the Solid State", *J. Am. Chem. Soc.* **143**, 163–175 (2021).
- [91] S. E. Biali, D. A. Nugiel, and Z. Rappoport, "Stable simple enols. Part 19. Steric effects and threshold rotational mechanisms in 1-substituted 2,2-dimesitylethenols", *J. Am. Chem. Soc.* **111**, 846–852 (1989).
- [92] V. Římal, H. Štěpánková, and J. Štěpánek, "Analysis of NMR spectra in case of temperature-dependent chemical exchange between two unequally populated sites", *Concepts in Magnetic Resonance Part A* **38A**, 117–127 (2011).
- [93] J. Kozic, Z. Novák, V. Římal, V. Profant, J. Kuneš, and J. Vinšová, "Conformations, equilibrium thermodynamics and rotational barriers of secondary thiobenzanilides", *Tetrahedron* **72**, 2072–2083 (2016).
- [94] A. Takai, D. J. Freas, T. Suzuki, M. Sugimoto, J. Labuta, R. Haruki, R. Kumai, S.-i. Adachi, H. Sakai, T. Hasobe, Y. Matsushita, and M. Takeuchi, "The effect of a highly twisted CC double bond on the electronic structures of 9,9'-bifluorenylidene derivatives in the ground and excited states", *Org. Chem. Front.* **4**, 650–657 (2017).
- [95] M. Hensmann, G. W. Booker, G. Panayotou, J. Boyd, J. Linacre, M. Waterfield, and I. D. Campbell, "Phosphopeptide binding to the N-terminal SH2 domain of the p85 α subunit of PI 3'-kinase: A heteronuclear NMR study", *Protein Science* **3**, 1020–1030 (1994).
- [96] U. Günther, T. Mittag, and B. Schaffhausen, "Probing Src Homology 2 Domain Ligand Interactions by Differential Line Broadening", *Biochemistry* **41**, 11658–11669 (2002).
- [97] T. Mittag, B. Schaffhausen, and U. L. Günther, "Tracing Kinetic Intermediates during Ligand Binding", *J. Am. Chem. Soc.* **126**, 9017–9023 (2004).
- [98] C. A. Waudby, A. Ramos, L. D. Cabrita, and J. Christodoulou, "Two-Dimensional NMR Lineshape Analysis", *Sci. Rep.* **6**, 1–8 (2016).
- [99] S. Shinya, M. G. Ghinet, R. Brzezinski, K. Furuita, C. Kojima, S. Shah, E. L. Kovrigin, and T. Fukamizo, "NMR line shape analysis of a multi-state ligand binding mechanism in chitosanase", *J. Biomol. NMR* **67**, 309–319 (2017).
- [100] C. A. Waudby, M. Ouvry, B. Davis, and J. Christodoulou, "Two-dimensional NMR lineshape analysis of single, multiple, zero and double quantum correlation experiments", *J. Biomol. NMR* **74**, 95–109 (2020).
- [101] A. D. Bain, "Chemical exchange in NMR", *Progress in Nuclear Magnetic Resonance Spectroscopy* **43**, 63–103 (2003).
- [102] I. R. Kleckner and M. P. Foster, "An introduction to NMR-based approaches for measuring protein dynamics", *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, Protein Dynamics: Experimental and Computational Approaches* **1814**, 942–968 (2011).

- [103] A. Furukawa, T. Konuma, S. Yanaka, and K. Sugase, "Quantitative analysis of protein–ligand interactions by NMR", *Progress in Nuclear Magnetic Resonance Spectroscopy* **96**, 47–57 (2016).
- [104] E. L. King and C. Altman, "A Schematic Method of Deriving the Rate Laws for Enzyme-Catalyzed Reactions", *J. Phys. Chem.* **60**, 1375–1378 (1956).
- [105] T. Hill, *Free Energy Transduction in Biology: The Steady-State Kinetic and Thermodynamic Formalism* (Elsevier, Dec. 2012).
- [106] S. Grilli, L. Lunazzi, A. Mazzanti, D. Casarini, and C. Femoni, "Conformational Studies by Dynamic NMR. 78.1 Stereomutation of the Helical Enantiomers of Trigonal Carbon Diaryl-Substituted Compounds: Dimesitylketone, Dimesitylthio ketone, and Dimesitylethylene", *J. Org. Chem.* **66**, 488–495 (2001).
- [107] S. Grilli, L. Lunazzi, A. Mazzanti, and G. Mazzanti, "Conformational Studies by Dynamic NMR. 79.1 Dimesityl Sulfine Revisited: Detection of the Helical Antipodes and Determination of Their Enantiomerization Pathways", *J. Org. Chem.* **66**, 748–754 (2001).
- [108] S. E. Biali and Z. Rappoport, "Stable simple enols. 3. Static and dynamic NMR behavior of crowded triarylethenols and related compounds. Three-ring flip as the threshold mechanism for enantiomerization of crowded triarylvinyl propellers", *J. Am. Chem. Soc.* **106**, 477–496 (1984).
- [109] D. Gust and K. Mislow, "Analysis of isomerization in compounds displaying restricted rotation of aryl groups", *J. Am. Chem. Soc.* **95**, 1535–1547 (1973).
- [110] H. M. McConnell, "Reaction Rates by Nuclear Magnetic Resonance", *J. Chem. Phys.* **28**, 430–431 (1958).
- [111] S. Alexander, "Exchange of Interacting Nuclear Spins in Nuclear Magnetic Resonance. II. Chemical Exchange", *J. Chem. Phys.* **37**, 974–980 (1962).
- [112] E. L. Kovrigin, "NMR line shapes and multi-state binding equilibria", *J. Biomol. NMR* **53**, 257–270 (2012).
- [113] C. Feng, E. L. Kovrigin, and C. B. Post, "NmrLineGuru: Standalone and User-Friendly GUIs for Fast 1D NMR Lineshape Simulation and Analysis of Multi-State Equilibrium Binding Models", *Sci. Rep.* **9**, 1–14 (2019).
- [114] A. I. Greenwood, M. J. Rogals, S. De, K. P. Lu, E. L. Kovrigin, and L. K. Nicholson, "Complete determination of the Pin1 catalytic domain thermodynamic cycle by NMR lineshape analysis", *J. Biomol. NMR* **51**, 21–34 (2011).
- [115] G. Binsch, *Topics in Stereochemistry*, edited by E. L. Eliel and N. L. Allinger, Vol. 3 (Interscience publishers., 1968).
- [116] U. L. Günther and B. Schaffhausen, "NMRKIN: Simulating line shapes from two-dimensional spectra of proteins upon ligand binding", *J. Biomol. NMR* **22**, 201–209 (2002).
- [117] C. B. Post, "[19] Characterization of enzyme-complex formation by analysis of nuclear magnetic resonance line shapes", in *Methods in Enzymology*, Vol. 240, Part B: Numerical Computer Methods (Academic Press, Jan. 1994), pp. 438–446.
- [118] H. S. Gutowsky and A. Saika, "Dissociation, Chemical Exchange, and the Proton Magnetic Resonance in Some Aqueous Electrolytes", *J. Chem. Phys.* **21**, 1688–1694 (1953).
- [119] C. S. Johnson, *Advances in Magnetic Resonance*, edited by J. S. Waugh, Vol. 1 (Academic Press, 1965).

- [120] R. A. Sack, "A contribution to the theory of the exchange narrowing of spectral lines", *Molecular Physics* **1**, 163–167 (1958).
- [121] J. Fišer, *Úvod do molekulové symetrie: aplikace teorie grup v chemii* (SNTL, 1980).
- [122] D. Parker, "NMR determination of enantiomeric purity", *Chem. Rev.* **91**, 1441–1457 (1991).
- [123] P. L. Rinaldi, "The determination of absolute configuration using nuclear magnetic resonance techniques", *Progress in Nuclear Magnetic Resonance Spectroscopy* **15**, 291–352 (1982).
- [124] R. Fulwood and D. Parker, "1,2-Diphenylethane-1,2-diamine: an effective NMR chiral solvating agent for chiral carboxylic acids", *Journal of the Chemical Society, Perkin Transactions 2* **0**, 57–64 (1994).
- [125] E. J. Ebbers, G. J. A. Ariaans, J. P. M. Houbiers, A. Bruggink, and B. Zwanenburg, "Controlled racemization of optically active organic compounds: Prospects for asymmetric transformation", *Tetrahedron* **53**, 9417–9476 (1997).
- [126] E. L. Eliel, "Prostereoisomerism (prochirality)", in *Organic Chemistry*, Topics in Current Chemistry (1982), pp. 1–76.
- [127] K. R. Hanson, "Applications of the Sequence Rule. I. Naming the Paired Ligands g,g at a Tetrahedral Atom Xggj. II. Naming the Two Faces of a Trigonal Atom Yghi", *J. Am. Chem. Soc.* **88**, 2731–2742 (1966).
- [128] P. Lesot, C. Aroulanda, H. Zimmermann, and Z. Luz, "Enantiotopic discrimination in the NMR spectrum of prochiral solutes in chiral liquid crystals", *Chemical Society Reviews* **44**, 2330–2375 (2015).
- [129] D. Merlet, J. W. Emsley, P. Lesot, and J. Courtieu, "The relationship between molecular symmetry and second-rank orientational order parameters for molecules in chiral liquid crystalline solvents", *J. Chem. Phys.* **111**, 6890–6896 (1999).
- [130] D. A. Skoog, F. J. Holler, and S. R. Crouch, *Principles of Instrumental Analysis* (Cengage Learning, Jan. 2017).
- [131] S. E. Braslavsky, "Glossary of terms used in photochemistry, 3rd edition (IUPAC Recommendations 2006)", *Pure and Applied Chemistry* **79**, 293–465 (2007).
- [132] T. Soderberg, "Organic Chemistry with a Biological Emphasis Volume I", Chemistry Publications (2019).
- [133] S. Voigt, J. Orphal, and J. P. Burrows, "The temperature and pressure dependence of the absorption cross-sections of NO₂ in the 250–800 nm region measured by Fourier-transform spectroscopy", *Journal of Photochemistry and Photobiology A: Chemistry* **149**, 1–7 (2002).
- [134] W. Demtröder, *Atoms, molecules, and photons: an introduction to atomic-, molecular-, and quantum-physics* (Springer, Berlin ; New York, 2006).
- [135] D. B. Siano and D. E. Metzler, "Band Shapes of the Electronic Spectra of Complex Molecules", *J. Chem. Phys.* **51**, 1856–1861 (1969).
- [136] B. Koeppe, P. M. Tolstoy, and H.-H. Limbach, "Reaction Pathways of Proton Transfer in Hydrogen-Bonded Phenol–Carboxylate Complexes Explored by Combined UV–Vis and NMR Spectroscopy", *J. Am. Chem. Soc.* **133**, 7897–7908 (2011).
- [137] P. L. Dawson and J. C. Acton, "22 - Impact of proteins on food color", in *Proteins in Food Processing (Second Edition)*, edited by R. Y. Yada, Woodhead Publishing Series in Food Science, Technology and Nutrition (Woodhead Publishing, Jan. 2018), pp. 599–638.

- [138] A. Sassaroli and S. Fantini, "Comment on the modified Beer–Lambert law for scattering media", *Phys. Med. Biol.* **49**, N255–N257 (2004).
- [139] P. Bamfield and M. Hutchings, *Chromic Phenomena* (Aug. 2018).
- [140] L. S. Foster and I. J. Gruntfest, "Demonstration experiments using universal indicators", *J. Chem. Educ.* **14**, 274 (1937).
- [141] C. Reichardt and T. Welton, *Solvents and Solvent Effects in Organic Chemistry* (John Wiley & Sons, Aug. 2011).
- [142] C. Reichardt, "Solvatochromism, thermochromism, piezochromism, halochromism, and chiro-solvatochromism of pyridinium N-phenoxide betaine dyes", *Chem. Soc. Rev.* **21**, 147–153 (1992).
- [143] V. G. Machado, R. I. Stock, and C. Reichardt, "Pyridinium N-Phenolate Betaine Dyes", *Chem. Rev.* **114**, 10429–10475 (2014).
- [144] I. T. Jolliffe, *Principal Component Analysis*, Second, Springer Series in Statistics (Springer-Verlag, New York, 2002).
- [145] V. M. Brown, A. Ossadtchi, A. H. Khan, S. Yee, G. Lacan, W. P. Melega, S. R. Cherry, R. M. Leahy, and D. J. Smith, "Multiplex Three-Dimensional Brain Gene Expression Mapping in a Mouse Model of Parkinson's Disease", *Genome Res.* **12**, 868–884 (2002).
- [146] G. Zeng, "Facial Recognition with Singular Value Decomposition", in *Advances and Innovations in Systems, Computing Sciences and Software Engineering*, edited by K. Elleithy (2007), pp. 145–148.
- [147] K. Taira, S. L. Brunton, S. T. M. Dawson, C. W. Rowley, T. Colonius, B. J. McKeon, O. T. Schmidt, S. Gordyev, V. Theofilis, and L. S. Ukeiley, "Modal Analysis of Fluid Flows: An Overview", *AIAA Journal* **55**, 4013–4041 (2017).
- [148] I. Scherl, B. Strom, J. K. Shang, O. Williams, B. L. Polagye, and S. L. Brunton, "Robust principal component analysis for modal decomposition of corrupt fluid flows", *Phys. Rev. Fluids* **5**, 054401 (2020).
- [149] U. Khadka, V. Holubec, H. Yang, and F. Cichos, "Active particles bound by information flows", *Nature Communications* **9**, 3864 (2018).
- [150] T. W. Holliday and L. Friedl, "Hominoid humeral morphology: 3D morphometric analysis", *American Journal of Physical Anthropology* **152**, 506–515 (2013).
- [151] M. Martín-Torres, F. d'Errico, E. Santos, A. Álvaro Gallo, N. Amano, W. Archer, S. J. Armitage, J. L. Arsuaga, J. M. Bermúdez de Castro, J. Blinkhorn, A. Crowther, K. Douka, S. Dubernet, P. Faulkner, P. Fernández-Colón, N. Kourampas, J. González García, D. Larraina, F.-X. Le Bourdonnec, G. MacLeod, L. Martín-Francés, D. Massilani, J. Mercader, J. M. Miller, E. Ndiema, B. Notario, A. Pitarch Martí, M. E. Prendergast, A. Queffelec, S. Rigaud, P. Roberts, M. J. Shoaee, C. Shipton, I. Simpson, N. Boivin, and M. D. Petraglia, "Earliest known human burial in Africa", *Nature* **593**, 95–100 (2021).
- [152] E. R. Malinowski, *Factor Analysis in Chemistry* (Wiley, Mar. 2002).
- [153] C. Spearman, "'General Intelligence,' Objectively Determined and Measured", *The American Journal of Psychology* **15**, 201–292 (1904).
- [154] D. T. Payne, J. Labuta, Z. Futera, V. Březina, L. Hanyková, M. K. Chahal, and J. P. Hill, "Molecular rotor based on an oxidized resorcinarene", *Organic Chemistry Frontiers* **9**, 39–50 (2022).

- [155] J. Hanuš, K. Chmelová, J. Štěpánek, P.-Y. Turpin, J. Bok, I. Rosenberg, and Z. Točík, “Raman spectroscopic study of triplex-like complexes of polyuridylic acid with the isopolar, non-isosteric phosphonate analogues of diadenosine monophosphate”, *Journal of Raman Spectroscopy* **30**, 667–676 (1999).
- [156] J. Palacký, P. Mojzeš, and J. Bok, “SVD-based method for intensity normalization, background correction and solvent subtraction in Raman spectroscopy exploiting the properties of water stretching vibrations”, *J. Raman Spectrosc.* **42**, 1528–1539 (2011).
- [157] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control* (Cambridge University Press, Feb. 2019).
- [158] E. R. Henry and J. Hofrichter, “[8] Singular value decomposition: Application to analysis of experimental data”, in *Methods in Enzymology*, Vol. 210, Numerical Computer Methods (Academic Press, Jan. 1992), pp. 129–192.
- [159] E. R. Henry, “The Use of Matrix Methods in the Modeling of Spectroscopic Data Sets”, *Biophysical Journal* **72**, 652–673 (1997).
- [160] L. Zimányi, Á. Kulcsár, J. K. Lanyi, D. F. Sears, and J. Saltiel, “Singular value decomposition with self-modeling applied to determine bacteriorhodopsin intermediate spectra: Analysis of simulated data”, *PNAS* **96**, 4408–4413 (1999).
- [161] L. Zimányi, “Analysis of the Bacteriorhodopsin Photocycle by Singular Value Decomposition with Self-Modeling: A Critical Evaluation Using Realistic Simulated Data”, *J. Phys. Chem. B* **108**, 4199–4209 (2004).
- [162] L. Zimányi, J. Saltiel, L. S. Brown, and J. K. Lanyi, “A Priori Resolution of the Intermediate Spectra in the Bacteriorhodopsin Photocycle: The Time Evolution of the L Spectrum Revealed”, *J. Phys. Chem. A* **110**, 2318–2321 (2006).
- [163] L. Zimányi, S. Thekkan, B. Eckert, A. R. Condren, O. Dmitrenko, L. R. Kuhn, I. V. Alabugin, and J. Saltiel, “Determination of the pKa Values of trans-Resveratrol, a Triphenolic Stilbene, by Singular Value Decomposition. Comparison with Theory”, *J. Phys. Chem. A* **124**, 6294–6302 (2020).
- [164] R. Bro, E. Acar, and T. G. Kolda, “Resolving the sign ambiguity in the singular value decomposition”, *Journal of Chemometrics* **22**, 135–140 (2008).
- [165] G. H. Golub and C. F. V. Loan, *Matrix Computations* (JHU Press, 2013).
- [166] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing* (Cambridge University Press, Sept. 2007).
- [167] R. Penrose, “A generalized inverse for matrices”, *Mathematical Proceedings of the Cambridge Philosophical Society* **51**, 406–413 (1955).
- [168] N. R. Draper and H. Smith, *Applied Regression Analysis* (John Wiley & Sons, Aug. 2014).
- [169] M. Gavish and D. L. Donoho, “The Optimal Hard Threshold for Singular Values is $4/\sqrt{3}$ ”, *IEEE Transactions on Information Theory* **60**, 5040–5053 (2014).
- [170] A. G. Ioannidis, J. Blanco-Portillo, K. Sandoval, E. Hagelberg, J. F. Miquel-Poblete, J. V. Moreno-Mayar, J. E. Rodríguez-Rodríguez, C. D. Quinto-Cortés, K. Auckland, T. Parks, K. Robson, A. V. S. Hill, M. C. Avila-Arcos, A. Sockell, J. R. Homburger, G. L. Wojcik, K. C. Barnes, L. Herrera, S. Berríos, M. Acuña, E. Llop, C. Eng, S. Huntsman, E. G. Burchard, C. R. Gignoux, L. Cifuentes, R. A. Verdugo, M. Moraga, A. J. Mentzer, C. D. Bustamante, and A. Moreno-Estrada, “Native American gene flow into Polynesia predating Easter Island settlement”, *Nature* **583**, 572–577 (2020).

- [171] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments", *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **374**, 20150202 (2016).
- [172] J. P. Hill, I. J. Hewitt, C. E. Anson, A. K. Powell, A. L. McCarty, P. A. Karr, M. E. Zandler, and F. D'Souza, "Highly Nonplanar, Electron Deficient, N-Substituted tetra-Oxocyclohexadienylidene Porphyrinogens: Structural, Computational, and Electrochemical Investigations", *J. Org. Chem.* **69**, 5861–5869 (2004).
- [173] C. J. Medforth, P. E. Berget, J. C. Fettinger, K. M. Smith, and J. A. Shelnut, "Determination of the activation energies for ND tautomerism and anion exchange in a porphyrin monocation", *J. Porphyrins Phthalocyanines* **20**, 307–317 (2016).
- [174] S. Thyagarajan, T. Leiding, S. P. Årsköld, A. V. Cheprakov, and S. A. Vinogradov, "Highly Non-Planar Dendritic Porphyrin for pH Sensing: Observation of Porphyrin Monocation", *Inorg. Chem.* **49**, 9909–9920 (2010).
- [175] D. Casarini, L. Lunazzi, and A. Mazzanti, "Recent Advances in Stereodynamics and Conformational Analysis by Dynamic NMR and Theoretical Calculations", *European Journal of Organic Chemistry* **2010**, 2035–2056 (2010).
- [176] B. Lindbloom, *Spectral calculator*, <http://www.brucelindbloom.com/>.
- [177] H. S. Muddana and M. K. Gilson, "Calculation of Host–Guest Binding Affinities Using a Quantum-Mechanical Energy Model", *J. Chem. Theory Comput.* **8**, 2023–2033 (2012).

List of Abbreviations

CSA	camphorsulfonic acid
CW	continuous wave
DFA	difluoroacetic acid
DMF	dimethylformamide
DSIS	decomposition of spectra into individual species
FID	free inductions decay
FT	Fourier transform
IR	infrared
NMR	nuclear magnetic resonance
PC	principal component
PCA	principal component analysis
RF	radiofrequency
RSD	residual standard deviation
SVD	singular value decomposition
TFA	trifluoroacetic acid
TMS	tetramethylsilan
TPP	tetraphenylporphyrin
UV/vis	ultraviolet/visible

A Overview of binding models

In this section, *Mathematica* code for the calculation of binding isotherms is provided.

A.1 Binding model 1:1

The simple 1:1 stoichiometry enables simple analytical solution. The concentrations are calculated as the following functions

```
(* H:G 1:1 *)
H[K1_,Ht_,Gt_] := (-1 - Gt K1 + Ht K1 + Sqrt[4 Ht K1 + (-1 - Gt K1 + Ht K1)^2])/(2 K1);
HG[K1_,Ht_,Gt_] := (1 + Gt K1 + Ht K1 - Sqrt[4 Ht K1 + (-1 - Gt K1 + Ht K1)^2])/(2 K1);
G[K1_,Ht_,Gt_] := (-1 + Gt K1 - Ht K1 + Sqrt[4 Ht K1 + (-1 - Gt K1 + Ht K1)^2])/(2 K1);
```

Using artificial data, numeric values can be obtained as follows

```
(* creating artificial data: *)
Htdata = ConstantArray[.001, 100];
Gtdata = Subdivide[0, .01, 99];
K1 = 5000;
(* calculating the concentrations *)
Gdata=G[K1, Htdata, Gtdata];
Hdata=H[K1, Htdata, Gtdata];
HGdata=HG[K1, Htdata, Gtdata];
```

A.2 Binding model 1:2

In principle, this model can be solved analytically using the formula for cubic equation. However, numerical approach from [71] can be applied universally as explained in Sec. 2.2.2. The derivation and implementation is shown here.

First, the host-related species (i.e. $[H]$ and $[HG_j]$) and $[G]_t$ are expressed as functions of $[G]$. Then the free guest concentration is calculated numerically from a polynomial equation in the form $[G]_t = [G]_t([G], [H]_t, K_j)$. The equations for stepwise binding constants and total mass balance are

```
(* H:G 1:2 *)
equations = {K1 == HG/(H G), K2 == HG2/(HG G), Ht == H + HG + HG2, Gt == G + HG + 2 HG2};
solutionRule = Solve[And @@ equations, {H, HG, HG2, Gt}][[1]]
```

This solution provides formulae for host-related species

```
H[K1_, K2_, Ht_, G_] := Ht/(1 + G K1 + G^2 K1 K2);
HG[K1_, K2_, Ht_, G_] := G Ht K1/(1 + G K1 + G^2 K1 K2);
HG2[K1_, K2_, Ht_, G_] := G^2 Ht K1 K2/(1 + G K1 + G^2 K1 K2);
```

The concentration $[G]$ can be calculated from a polynomial equation, which can be obtained from the solution for $[G]_t$ by the following procedure

```
polynomEq = (Gt == (Gt /. solutionRule));
polynomEq = MultiplySides[polynomEq, Denominator[(Gt /. solutionRule)]] [[1, 1, 1]];
polynomEq = SubtractSides[polynomEq, polynomEq[[2]]];
polynomEq = Collect[polynomEq, G] (* polynomial equation for [G] *)
```

which yields the `polynomEq` in the form

```
Gt + G (-1 + Gt K1 - Ht K1) - G^3 K1 K2 + G^2 (-K1 + Gt K1 K2 - 2 Ht K1 K2) == 0
```

The code for calculation of [G] is then obtained by

```
G[K1_?NumericQ, K2_?NumericQ, Ht_, Gt_] :=
G /. FindRoot[{Gt + G (-1 + Gt K1 - Ht K1) - G^3 K1 K2 + G^2 (-K1 + Gt K1 K2 - 2 Ht K1 K2)
== 0}, {G, Gt}];
SetAttributes[G, Listable];
```

The command `SetAttributes[G, Listable]` allows accepting values of Ht and Gt as lists of numbers. For the use in fitting procedure, the check for numeric value `?NumericQ` must be added to input arguments.

Using artificial data, numeric values can be obtained as follows

```
(* creating artificial data: *)
Htdata = ConstantArray[.001, 100];
Gtdata = Subdivide[0, .01, 99];
K1 = 5000; K2 = 1000;
(* calculating the concentrations *)
Gdata=G[K1, K2, Htdata, Gtdata];
Hdata=H[K1, K2, Htdata, Gdata];
HGdata=HG[K1, K2, Htdata, Gdata];
HG2data=HG2[K1, K2, Htdata, Gdata];
```

A.3 Binding model 1:3

The equations describing the model are

```
(* H:G 1:3 *)
equations = {K1 == HG/(H G), K2 == HG2/(HG G), K3 == HG3/(HG2 G),
  Ht == H + HG + HG2 + HG3, Gt == G + HG + 2 HG2 + 3 HG3};
solutionRule = Solve[And @@ equations, {H, HG, HG2, HG3, Gt}][[1]]
```

Polynomial equation for [G] is derived analogously to the previous section, the resulting code is

```
H[K1_, K2_, K3_, Ht_, G_] := Ht/(1 + G K1 + G^2 K1 K2 + G^3 K1 K2 K3);
HG[K1_, K2_, K3_, Ht_, G_] := G Ht K1/(1 + G K1 + G^2 K1 K2 + G^3 K1 K2 K3);
HG2[K1_, K2_, K3_, Ht_, G_] := G^2 Ht K1 K2/(1 + G K1 + G^2 K1 K2 + G^3 K1 K2 K3);
HG3[K1_, K2_, K3_, Ht_, G_] := G^3 Ht K1 K2 K3/(1 + G K1 + G^2 K1 K2 + G^3 K1 K2 K3);
G[K1_?NumericQ, K2_?NumericQ, K3_?NumericQ, Ht_, Gt_] :=
  G /. FindRoot[{Gt + G (-1 + Gt K1 - Ht K1) + G^2 (-K1 + Gt K1 K2 - 2 Ht K1 K2)
  - G^4 K1 K2 K3 + G^3 (-K1 K2 + Gt K1 K2 K3 - 3 Ht K1 K2 K3) == 0}, {G, Gt}];
SetAttributes[G, Listable];
```

Using artificial data, numeric values can be obtained as follows

```
(* creating artificial data: *)
Htdata = ConstantArray[.001, 100];
Gtdata = Subdivide[0, .01, 99];
K1 = 5000; K2 = 1000; K3 = 200;
(* calculating the concentrations *)
Gdata=G[K1, K2, K3, Htdata, Gtdata];
Hdata=H[K1, K2, K3, Htdata, Gdata];
HGdata=HG[K1, K2, K3, Htdata, Gdata];
HG2data=HG2[K1, K2, K3, Htdata, Gdata];
HG3data=HG3[K1, K2, K3, Htdata, Gdata];
```


A.4 Binding model 1:4

The equations describing the model are

```
(* H:G 1:4 *)
equations = {K1 == HG/(H G), K2 == HG2/(HG G), K3 == HG3/(HG2 G), K4 == HG4/(HG3 G),
  Ht == H + HG + HG2 + HG3 + HG4, Gt == G + HG + 2 HG2 + 3 HG3 + 4 HG4};
solutionRule = Solve[And @@ equations, {H, HG, HG2, HG3, HG4, Gt}][[1]]
```

Polynomial equation for [G] is derived analogously to previous sections, the resulting code is

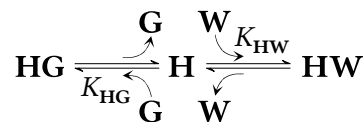
```
H[K1_, K2_, K3_, K4_, Ht_, G_] :=
  Ht/(1 + G K1 + G^2 K1 K2 + G^3 K1 K2 K3 + G^4 K1 K2 K3 K4);
HG[K1_, K2_, K3_, K4_, Ht_, G_] :=
  G Ht K1/(1 + G K1 + G^2 K1 K2 + G^3 K1 K2 K3 + G^4 K1 K2 K3 K4);
HG2[K1_, K2_, K3_, K4_, Ht_, G_] :=
  G^2 Ht K1 K2/(1 + G K1 + G^2 K1 K2 + G^3 K1 K2 K3 + G^4 K1 K2 K3 K4);
HG3[K1_, K2_, K3_, K4_, Ht_, G_] :=
  G^3 Ht K1 K2 K3/(1 + G K1 + G^2 K1 K2 + G^3 K1 K2 K3 + G^4 K1 K2 K3 K4);
HG4[K1_, K2_, K3_, K4_, Ht_, G_] :=
  G^4 Ht K1 K2 K3 K4/(1 + G K1 + G^2 K1 K2 + G^3 K1 K2 K3 + G^4 K1 K2 K3 K4);
G[K1_?NumericQ, K2_?NumericQ, K3_?NumericQ, K4_?NumericQ, Ht_, Gt_] :=
  G /. FindRoot[{Gt + G (-1 + Gt K1 - Ht K1) + G^2 (-K1 + Gt K1 K2 - 2 Ht K1 K2) +
    G^3 (-K1 K2 + Gt K1 K2 K3 - 3 Ht K1 K2 K3) - G^5 K1 K2 K3 K4 +
    G^4 (-K1 K2 K3 + Gt K1 K2 K3 K4 - 4 Ht K1 K2 K3 K4) == 0}, {G, Gt}];
SetAttributes[G, Listable];
```

Using artificial data, numeric values can be obtained as follows

```
(* creating artificial data: *)
Htdata = ConstantArray[.001, 100];
Gtdata = Subdivide[0, .01, 99];
K1 = 5000; K2 = 1000; K3 = 200; K4 = 300;
(* calculating the concentrations *)
Gdata=G[K1, K2, K3, K4, Htdata, Gtdata];
Hdata=H[K1, K2, K3, K4, Htdata, Gdata];
HGdata=HG[K1, K2, K3, K4, Htdata, Gdata];
HG2data=HG2[K1, K2, K3, K4, Htdata, Gdata];
HG3data=HG3[K1, K2, K3, K4, Htdata, Gdata];
HG4data=HG4[K1, K2, K3, K4, Htdata, Gdata];
```

A.5 Competitive binding 1:1 host-ligand binding

The competitive binding of a host **H** with two different ligands **G** and **W** is described by the following scheme with the corresponding governing equations



$$K_{\text{HG}} = \frac{[\text{HG}]}{[\text{H}][\text{G}]}, \quad (\text{A.1a})$$

$$K_{\text{HW}} = \frac{[\text{HW}]}{[\text{H}][\text{W}]}, \quad (\text{A.1b})$$

$$[\text{H}]_t = [\text{H}] + [\text{HG}] + [\text{HW}], \quad (\text{A.1c})$$

$$[\text{G}]_t = [\text{G}] + [\text{HG}], \quad (\text{A.1d})$$

$$[\text{W}]_t = [\text{W}] + [\text{HW}]. \quad (\text{A.1e})$$

A.5.1 First solution

First, the concentrations of host-related species are expressed as functions of $[G]$ and $[W]$ (instead of $[G]_t$ and $[W]_t$) from the governing Eq. (A.1), yielding

$$[H] = \frac{[H]_t}{1 + K_{HG}[G] + K_{HW}[W]}, \quad (\text{A.2a})$$

$$[HG] = \frac{K_{HG}[G][H]_t}{1 + K_{HG}[G] + K_{HW}[W]}, \quad (\text{A.2b})$$

$$[HW] = \frac{K_{HW}[W][H]_t}{1 + K_{HG}[G] + K_{HW}[W]}. \quad (\text{A.2c})$$

In the next step, substitution of Eq. (A.2) into Eq. (A.1c) yields a set of two second-order polynomial equations in $[G]$ and $[W]$ after rearrangement,

$$0 = [G] - [G]_t + [G]^2 K_{HG} - [G][G]_t K_{HG} + [G][H]_t K_{HG} + [G] K_{HW}[W] - [G]_t K_{HW}[W], \quad (\text{A.3a})$$

$$0 = [W] + [G] K_{HG}[W] + [H]_t K_{HW}[W] + K_{HW}[W]^2 - [W]_t - [G] K_{HG}[W]_t - K_{HW}[W][W]_t. \quad (\text{A.3b})$$

This equation is solved numerically in the interval $([G], [W]) \in [0, [G]_t] \times [0, [W]_t]$.

In *Mathematica*, Eq. (A.2) and (A.3) can be derived with the following code

```
(* host-ligand binding H:G 1:1, H:W 1:1 *)
equations = {KHG == HG/(H G), KHW == HW/(H W),
  Ht == H + HG + HW, Gt == G + HG, Wt == W + HW};
solutionRule = Solve[And @@ equations, {H, HG, HW, Gt, Wt}][[1]];
polynomEq1 = (Gt == (Gt /. solutionRule));
polynomEq1 =
  MultiplySides[polynomEq1, Denominator[(Gt /. solutionRule)]] [[1, 1, 1]];
polynomEq1 = SubtractSides[polynomEq1, polynomEq1[[2]]];
polynomEq1 = Collect[polynomEq1, G] (* 1st polynomial equation for [G] and [W] *)
polynomEq2 = (Wt == (Wt /. solutionRule));
polynomEq2 =
  MultiplySides[polynomEq2, Denominator[(Wt /. solutionRule)]] [[1, 1, 1]];
polynomEq2 = SubtractSides[polynomEq2, polynomEq2[[2]]];
polynomEq2 = Collect[polynomEq2, W] (* 2nd polynomial equation for [G] and [W] *)
```

Then the model can be implemented using

```
H[KHG_, KHW_, Ht_, G_, W_] := Ht/(1 + G KHG + KHW W);
HG[KHG_, KHW_, Ht_, G_, W_] := G Ht KHG/(1 + G KHG + KHW W);
HW[KHG_, KHW_, Ht_, G_, W_] := Ht KHW W/(1 + G KHG + KHW W);
G[KHG_?NumericQ, KHW_?NumericQ, Ht_, Gt_, Wt_] :=
  G /. FindRoot[{Gt - G^2 KHG + Gt KHW W + G (-1 + Gt KHG - Ht KHG - KHW W) == 0,
    -KHW W^2 + Wt + G KHG Wt + W (-1 - G KHG - Ht KHW + KHW Wt) == 0},
    {{G, Gt}, {W, Wt}}];
W[KHG_?NumericQ, KHW_?NumericQ, Ht_, Gt_, Wt_] :=
  W /. FindRoot[{Gt - G^2 KHG + Gt KHW W + G (-1 + Gt KHG - Ht KHG - KHW W) == 0,
    -KHW W^2 + Wt + G KHG Wt + W (-1 - G KHG - Ht KHW + KHW Wt) == 0},
    {{G, Gt}, {W, Wt}}];
SetAttributes[G, Listable];
SetAttributes[W, Listable];
```

Using artificial data, numeric values can be obtained as follows

```
(* creating artificial data: *)
Htdata = ConstantArray[.001, 100];
Gtdata = Subdivide[0, .01, 99];
```

```

Wtdata = ConstantArray[.05, 100];
KHG = 5000; KHW = 50;
(* calculating the concentrations *)
Gdata=G[KHG, KHW, Htdata, Gtdata, Wtdata];
Wdata=W[KHG, KHW, Htdata, Gtdata, Wtdata];
Hdata=H[KHG, KHW, Htdata, Gdata, Wdata];
HGdata=HG[KHG, KHW, Htdata, Gdata, Wdata];
HWdata=HW[KHG, KHW, Htdata, Gdata, Wdata];

```

A.5.2 Second solution

Instead of numeric solution of set of two second-order polynomial equations as in the previous section, this method results in numeric solutions of one third-order polynomial equation, which is more numerically stable in some circumstances. Derivation of the equations described in Sec. 2.2.3 is implemented in *Mathematica* in the following code

```

(* host-ligand binding H:G 1:1, H:W 1:1 *)
equations = {KHG == HG/(H G), KHW == HW/(H W), Ht == H + HG + HW,
  Gt == G + HG, Wt == W + HW};
solutionRule = Solve[And @@ equations, {HG, HW, G, W, Ht}][[1]]
polynomEq = (Ht == (Ht /. solutionRule));
polynomEq =
  MultiplySides[polynomEq, Denominator[(Ht /. solutionRule)]] [[1, 1, 1]];
polynomEq = SubtractSides[polynomEq, polynomEq[[2]]];
polynomEq = Collect[polynomEq, H] (* polynomial equation for [H] *)

```

Then the model can be implemented using

```

H[KHG_?NumericQ, KHW_?NumericQ, Ht_, Gt_, Wt_] :=
  H /. FindRoot[{Ht - H^3 KHG KHW + H (-1 - Gt KHG + Ht KHG + Ht KHW - KHW Wt) +
    H^2 (-KHG - KHW - Gt KHG KHW + Ht KHG KHW - KHG KHW Wt) == 0}, {H, Ht}];
HG[KHG_?NumericQ, KHW_?NumericQ, H_, Gt_, Wt_] := Gt H KHG/(1 + H KHG);
HW[KHG_?NumericQ, KHW_?NumericQ, H_, Gt_, Wt_] := H KHW Wt/(1 + H KHW);
G[KHG_?NumericQ, KHW_?NumericQ, H_, Gt_, Wt_] := Gt/(1 + H KHG);
W[KHG_?NumericQ, KHW_?NumericQ, H_, Gt_, Wt_] := Wt/(1 + H KHW);
SetAttributes[H, Listable];

```

Using artificial data, numeric values can be obtained as follows

```

(* creating artificial data: *)
Htdata = ConstantArray[.001, 100];
Gtdata = Subdivide[0, .01, 99];
Wtdata = ConstantArray[.05, 100];
KHG = 5000; KHW = 50;
(* calculating the concentrations *)
Hdata=H[KHG, KHW, Htdata, Gtdata, Wtdata];
HGdata=HG[KHG, KHW, Hdata, Gtdata, Wtdata];
HWdata=HW[KHG, KHW, Hdata, Gtdata, Wtdata];
Gdata=G[KHG, KHW, Hdata, Gtdata, Wtdata];
Wdata=W[KHG, KHW, Hdata, Gtdata, Wtdata];

```

B Overview of NMR spectral lineshapes

B.1 Lineshape formulae and *Mathematica* codes

In this appendix, an overview of basic NMR spectral lineshapes is given. For practical application in fitting of experimental spectra, real parts of the lineshapes $\text{Re}[S(\Omega)]$ are implemented in *Mathematica* functions. Implementation in *MATLAB* can be found in [2]. Spectral lineshapes in Chapter 3 are constructed using angular frequency Ω in units of $\text{rad}\cdot\text{s}^{-1}$, but experimental NMR spectra are usually measured as chemical shift δ in units of ppm. Therefore, conversion formula from ppm to $\text{rad}\cdot\text{s}^{-1}$ is present in all *Mathematica* functions calculating the spectral lineshapes in this appendix, and it reads

$$\Delta\Omega [\text{s}^{-1}] = 2\pi \times \nu_0 [\text{MHz}] \times \Delta\delta [\text{ppm}], \quad (\text{B.1})$$

where ν_0 is the spectrometer frequency in MHz, see Sec. 3.1.2. Be aware that this formula is true only for differences of angular frequencies $\Delta\Omega$ and differences of chemical shift $\Delta\delta$. The conversion factor $\delta\text{To}\Omega\text{ConvFac}$ (equal to $\Delta\Omega/\Delta\delta$) is preset to spectrometer frequency 500.13 MHz (corresponding to $B_0 = 11.7$ T, cf. Eq. (3.4)).

In the presence of chemical exchange, state populations must be specified. In the presented *Mathematica* functions, these are calculated internally from analytical solution of $\mathbf{Kp} = 0$, which implies equilibrium or steady-state populations. Note that the R_2 parameters correspond to transverse relaxation including the inhomogeneous broadening effect (see Sec. 3.1.3). Importantly, area below the curve of all presented spectral lineshapes does not depend on peak position Ω_0 , transverse relaxation rate R_2 , transition rate coefficients k_{ij} or on J-coupling constant J , and it can be obtained as

$$\text{peak area} = \int_{-\infty}^{+\infty} \text{Re}[S(\Omega)] d\Omega = \pi M_0 \text{ (ang. freq. scale)}, \quad (\text{B.2a})$$

$$\text{peak area} = \int_{-\infty}^{+\infty} \text{Re}[S(\delta)] d\delta = \frac{M_0}{2\nu_0 [\text{MHz}]} \text{ (ppm scale)}. \quad (\text{B.2b})$$

Equation (B.2b) is obtained from Eq. (B.2a) using the substitution $\Omega = 2\pi\nu_0 [\text{MHz}]\delta$ from Eq. (B.1). Note that M_0 plays a role of arbitrary proportionality constant in measured spectra, it is not equal to the actual magnetization.

B.1.1 Lorentzian lineshape

Lorentzian lineshape is detected in the absence of chemical exchange and J-coupling. The analytical lineshape formula is

$$S(\Omega) = \frac{M_0}{R_2 - i(\Omega - \Omega_0)} = \frac{M_0}{\alpha}, \quad (\text{B.3})$$

where $\alpha = R_2 - i(\Omega - \Omega_0)$, with the real part

$$\text{Re}[S(\Omega)] = \frac{M_0 R_2}{R_2^2 + (\Omega - \Omega_0)^2}. \quad (\text{B.4})$$

Figure B.1 shows that full width at half-maximum of the Lorentzian lineshape equals to $2R_2$ in the angular frequency scale and 2γ in the chemical shift scale. Thus, R_2 can be calculated from a

NMR spectrum (measured in chemical shift scale) using

$$R_2 [s^{-1}] = 2\pi \times \nu_0 [\text{MHz}] \times \gamma [\text{ppm}], \quad (\text{B.5})$$

resulting from the conversion formula in Eq. (B.1).

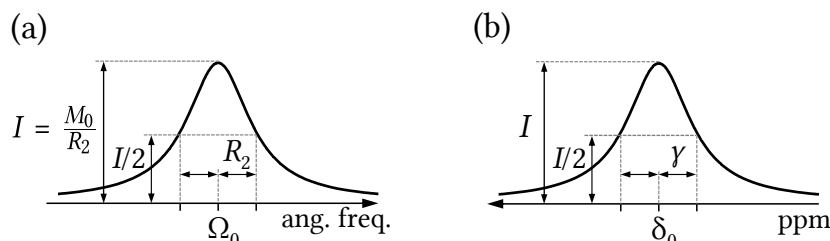


Figure B.1: Illustration of Lorentzian lineshape with the horizontal axis in (a) angular frequency and (b) chemical shift scale.

Code for implementation in *Mathematica* is

```
(* Lorentzian *)
lorentzianFunc[δ_, δ0_, R2_, M0_] :=
Module[{δToΩConvFac, Ω, Ω0},
  δToΩConvFac = 2π*500.13;
  Ω = δ*δToΩConvFac;
  Ω0 = δ0*δToΩConvFac;
  M0 R2/(R2^2 + (Ω - Ω0)^2)];
(* the use is illustrated in the following lines *)
deltaData = Range[0,2,.005];
absorptionSpectrum = lorentzianFunc[deltaData, (*δ0*)1, (*R2*)150, (*M0*)10];
```

Alternative formulation of the Lorentzian lineshape (real part) is

$$\text{Re}[S(\Omega)] = \frac{I\gamma^2}{\gamma^2 + (\delta - \delta_0)^2}, \quad (\text{B.6a})$$

$$\text{peak area} = \int_{-\infty}^{+\infty} \text{Re}[S(\delta)] d\delta = \pi\gamma I \text{ (ppm scale)}, \quad (\text{B.6b})$$

where $I = M_0/R_2$. This is convenient for fitting of spectral lineshapes because variation in γ does not influence the peak height (which always equals I). *Mathematica* implementation of Eq. (B.6a) reads

```
(* Lorentzian - alternative formulation *)
lorentzianAltFunc[δ_, δ0_, γ_, I_] :=
I γ^2/(γ^2 + (δ - δ0)^2);
(* the use is illustrated in the following lines *)
deltaData = Range[0,2,.005];
absorptionSpectrum = lorentzianAltFunc[deltaData, (*δ0*)1, (*γ*)0.04, (*I*)0.25];
```

B.1.2 Two-state exchange lineshape

The analytical formula for two-state exchange lineshape is

$$S_{\text{two-state exch.}}(\Omega) = M_0 \frac{p_A\alpha_B + p_B\alpha_A + k_{AB} + k_{BA}}{\alpha_A\alpha_B + k_{AB}\alpha_B + k_{BA}\alpha_A}, \quad (\text{B.7})$$

where $\alpha_j = R_2^j + i(\Omega - \Omega_j)$ for $j = A, B$. Equilibrium populations are calculated as

$$p_A^{\text{eq}} = \frac{k_{BA}}{k_{AB} + k_{BA}}, \quad (\text{B.8a})$$

$$p_B^{\text{eq}} = \frac{k_{AB}}{k_{AB} + k_{BA}}. \quad (\text{B.8b})$$

The *Mathematica* implementation is

```
(* two-state exchange lineshape *)
twoStateExchFunc[δ_, δ0List_, R2List_, kList_, MO_] :=
Module[{αA, αB, δA, δB, δToΩConvFac, Ω, ΩA, ΩB,
  RA, RB, kAB, kBA, pA, pB, P, Q},
  δToΩConvFac = 2π*500.13;
  {δA, δB} = δ0List;
  Ω = δ*δToΩConvFac;
  ΩA = δA*δToΩConvFac;
  ΩB = δB*δToΩConvFac;
  {kAB, kBA} = kList;
  {RA, RB} = R2List;
  {pA, pB} = {kBA/(kAB + kBA), kAB/(kAB + kBA)};
  αA = RA + I (Ω - ΩA); αB = RB + I (Ω - ΩB);
  P = pA αB + pB αA + kAB + kBA;
  Q = αA αB + kAB αB + kBA αA;
  MO Re[P/Q] ];
(* the use is illustrated in the following lines *)
deltaData = Range[0,2,.005];
absorptionSpectrum = twoStateExchFunc[deltaData, {(*δA*) .7, (*δB*) 1.4},
{(*RA*) 40, (*RB*) 40}, {(*kAB*) 200, (*kBA*) 300}, (*MO*) 10];
```

Any spectral exchange lineshape can be formally written as

$$S(\Omega) = M_0 \frac{\mathcal{A} + i\mathcal{B}}{C + iD} = M_0 \left(\frac{\mathcal{A}C + \mathcal{B}D}{C^2 + D^2} + i \frac{\mathcal{B}C - \mathcal{A}D}{C^2 + D^2} \right). \quad (\text{B.9})$$

Hence, the real part of asymmetric two-state exchange ($k_{AB} \neq k_{BA}$) can be explicitly written as

$$\begin{aligned} \mathcal{A} &= k_{AB} + k_{BA} + p_A R_B + p_B R_A, \\ \mathcal{B} &= p_A (\Omega - \Omega_B) + p_B (\Omega - \Omega_A), \\ \mathcal{C} &= k_{AB} R_B + k_{BA} R_A + R_A R_B - (\Omega - \Omega_A)(\Omega - \Omega_B), \\ \mathcal{D} &= (k_{AB} + R_A)(\Omega - \Omega_B) + (k_{BA} + R_B)(\Omega - \Omega_A) \end{aligned} \quad (\text{B.10})$$

with the corresponding code

```
(* two-state exchange lineshape with explicit real part *)
twoStateExchReFunc[δ_, δ0List_, R2List_, kList_, MO_] :=
Module[{δA, δB, δToΩConvFac, Ω, ΩA, ΩB,
  RA, RB, kAB, kBA, pA, pB, A, B, C, D},
  δToΩConvFac = 2π*500.13;
  {δA, δB} = δ0List;
  Ω = δ*δToΩConvFac;
  ΩA = δA*δToΩConvFac;
  ΩB = δB*δToΩConvFac;
  {kAB, kBA} = kList;
  {RA, RB} = R2List;
  {pA, pB} = {kBA/(kAB + kBA), kAB/(kAB + kBA)};
```

```

A = kAB + kBA + pA RB + pB RA;
B = pA (Ω - ΩB) + pB (Ω - ΩA);
C = kAB RB + kBA RA + RA RB - (Ω - ΩA) (Ω - ΩB);
D = (kAB + RA) (Ω - ΩB) + (kBA + RB) (Ω - ΩA);
M0 (A C + B D)/(C^2 + D^2)];
(* the use is illustrated in the following lines *)
deltaData = Range[0,2,.005];
absorptionSpectrum = twoStateExchReFunc[deltaData,{(*δA*) .7,(*δB*) 1.4},
{(*RA*) 40,(*RB*) 40},{(*kAB*) 200,(*kBA*) 300},(*M0*) 10];

```

In the case of symmetric two-state exchange ($k_{AB} = k_{BA} = k$), the formula simplifies to

$$\begin{aligned}
\mathcal{A} &= (4k + R_A + R_B)/2, \\
\mathcal{B} &= (2\Omega - \Omega_A - \Omega_B)/2, \\
\mathcal{C} &= k(R_A + R_B) + R_A R_B - (\Omega - \Omega_A)(\Omega - \Omega_B), \\
\mathcal{D} &= (k + R_A)(\Omega - \Omega_B) + (k + R_B)(\Omega - \Omega_A).
\end{aligned} \tag{B.11}$$

B.1.3 Three-state exchange lineshape

The analytical formula for two-state exchange lineshape is

$$S_{\text{three-state exch.}}(\Omega) = M_0 \frac{\mathcal{P}}{\mathcal{Q}}, \tag{B.12}$$

where

$$\begin{aligned}
\mathcal{P} &= p_A[\alpha_B \alpha_C + \alpha_B(k_{CA} + k_{CB} + k_{AC}) + \alpha_C(k_{BA} + k_{BC} + k_{AB})] \\
&\quad + p_B[\alpha_A \alpha_C + \alpha_A(k_{CA} + k_{CB} + k_{BC}) + \alpha_C(k_{AB} + k_{AC} + k_{BA})] \\
&\quad + p_C[\alpha_A \alpha_B + \alpha_A(k_{BA} + k_{BC} + k_{CB}) + \alpha_B(k_{AB} + k_{AC} + k_{CA})] \\
&\quad + \pi_A + \pi_B + \pi_C, \\
\mathcal{Q} &= \alpha_A \alpha_B \alpha_C + \alpha_A \alpha_B (k_{CA} + k_{CB}) + \alpha_A \alpha_C (k_{BA} + k_{BC}) + \alpha_B \alpha_C (k_{AB} + k_{AC}) \\
&\quad + \alpha_A \pi_A + \alpha_B \pi_B + \alpha_C \pi_C, \\
\pi_A &= k_{BA} k_{CA} + k_{BC} k_{CA} + k_{BA} k_{CB}, \\
\pi_B &= k_{AB} k_{CA} + k_{AB} k_{CB} + k_{AC} k_{CB}, \\
\pi_C &= k_{AC} k_{BA} + k_{AB} k_{BC} + k_{AC} k_{BC}.
\end{aligned}$$

Steady-state populations are calculated as

$$p_j^{\text{ss}} = \frac{\pi_j}{\pi_A + \pi_B + \pi_C} \tag{B.13}$$

with $j = A, B, C$. Equilibrium lineshapes are obtained when the input transition rate coefficients fulfill an additional condition $k_{AB} = \frac{k_{AC} k_{BA} k_{CB}}{k_{BC} k_{CA}}$, cf. Sec. 3.3.3.

The *Mathematica* implementation is

```

(* three-state exchange lineshape *)
threeStateExchFunc[δ_, δ0List_, R2List_, kList_, M0_] :=
Module[{αA, αB, αC, δA, δB, δC, δToΩConvFac, Ω, ΩA, ΩB, ΩC,
RA, RB, RC, kAB, kBA, kAC, kCA, kBC, kCB, pA, pB, pC, πA, πB, πC, P, Q},
δToΩConvFac = 2π*500.13;
{δA, δB, δC} = δ0List;
Ω = δ*δToΩConvFac;
ΩA = δA*δToΩConvFac;

```

```

ΩB = δB*δToΩConvFac;
ΩC = δC*δToΩConvFac;
{kAB, kBA, kAC, kCA, kBC, kCB} = kList;
πA = kBA kCA + kBC kCA + kBA kCB;
πB = kAB kCA + kAB kCB + kAC kCB;
πC = kAC kBA + kAB kBC + kAC kBC;
{RA, RB, RC} = R2List;
{pA, pB, pC} = {πA, πB, πC}/(πA + πB + πC);
αA = RA + I (Ω - ΩA); αB = RB + I (Ω - ΩB); αC = RC + I (Ω - ΩC);
P = pA (αB αC + αB (kCA + kCB + kAC)) + αC (kBA + kBC + kAB)
  + pB (αA αC + αA (kCA + kCB + kBC)) + αC (kAB + kAC + kBA)
  + pC (αA αB + αA (kBA + kBC + kCB)) + αB (kAB + kAC + kCA)
  + (πA + πB + πC);
Q = αA αB αC + αA αB (kCA + kCB) + αA αC (kBA + kBC) + αB αC (kAB + kAC)
  + αA πA + αB πB + αC πC;
MO Re[P/Q ]];
(* the use is illustrated in the following lines *)
deltaData = Range[0,2,.005];
absorptionSpectrum = threeStateExchFunc[deltaData,{(*δA*) .7, (*δB*)1, (*δC*)1.4},
{(*RA*)40, (*RB*)40, (*RC*)40},{(*kAB*)200, (*kBA*)80, (*kAC*)100, (*kCA*)150,
(*kBC*)70, (*kCB*)80}, (*MO*)10];

```

Using Eq. (B.9), the real part of three-state exchange lineshape can be expressed explicitly as

$$\begin{aligned}
\mathcal{A} &= p_A [R_2^B R_2^C + R_2^B (k_{CA} + k_{CB} + k_{AC}) + R_2^C (k_{BA} + k_{BC} + k_{AB}) - (\Omega - \Omega_B)(\Omega - \Omega_C)] \\
&\quad + p_B [R_2^A R_2^C + R_2^A (k_{CA} + k_{CB} + k_{BC}) + R_2^C (k_{AB} + k_{AC} + k_{BA}) - (\Omega - \Omega_A)(\Omega - \Omega_C)] \\
&\quad + p_C [R_2^A R_2^B + R_2^A (k_{BA} + k_{BC} + k_{CB}) + R_2^B (k_{AB} + k_{AC} + k_{CA}) - (\Omega - \Omega_A)(\Omega - \Omega_B)] \\
&\quad + \pi_A + \pi_B + \pi_C, \\
\mathcal{B} &= p_A [(R_2^C + k_{CA} + k_{CB} + k_{AC})(\Omega - \Omega_B) + (R_2^B + k_{BA} + k_{BC} + k_{AB})(\Omega - \Omega_C)] \\
&\quad + p_B [(R_2^C + k_{CA} + k_{CB} + k_{BC})(\Omega - \Omega_A) + (R_2^A + k_{AB} + k_{AC} + k_{BA})(\Omega - \Omega_C)] \\
&\quad + p_C [(R_2^B + k_{BA} + k_{BC} + k_{CB})(\Omega - \Omega_A) + (R_2^A + k_{AB} + k_{AC} + k_{CA})(\Omega - \Omega_B)], \\
\mathcal{C} &= R_2^A R_2^B R_2^C + (k_{CA} + k_{CB}) R_2^A R_2^B + (k_{BA} + k_{BC}) R_2^A R_2^C + (k_{AB} + k_{AC}) R_2^B R_2^C \\
&\quad - (R_2^C + k_{CA} + k_{CB})(\Omega - \Omega_A)(\Omega - \Omega_B) \\
&\quad - (R_2^B + k_{BA} + k_{BC})(\Omega - \Omega_A)(\Omega - \Omega_C) \\
&\quad - (R_2^A + k_{AB} + k_{AC})(\Omega - \Omega_B)(\Omega - \Omega_C) \\
&\quad + R_2^A \pi_A + R_2^B \pi_B + R_2^C \pi_C, \\
\mathcal{D} &= (R_2^B R_2^C + (k_{CA} + k_{CB}) R_2^B + (k_{BA} + k_{BC}) R_2^C + \pi_A)(\Omega - \Omega_A) \\
&\quad + (R_2^A R_2^C + (k_{CA} + k_{CB}) R_2^A + (k_{AB} + k_{AC}) R_2^C + \pi_B)(\Omega - \Omega_B) \\
&\quad + (R_2^A R_2^B + (k_{BA} + k_{BC}) R_2^A + (k_{AB} + k_{AC}) R_2^B + \pi_C)(\Omega - \Omega_C) \\
&\quad - (\Omega - \Omega_A)(\Omega - \Omega_B)(\Omega - \Omega_C)
\end{aligned} \tag{B.14}$$

with *Mathematica* code

```

(* three-state exchange lineshape with explicit real part *)
threeStateExchReFunc[δ_, δOList_, R2List_, kList_, MO_] :=
Module[{δA, δB, δC, δToΩConvFac, Ω, ΩA, ΩB, ΩC, RA,
  RB, RC, kAB, kBA, kAC, kCA, kBC, kCB, pA, pB, pC, πA, πB, πC, A, B, C, D},
δToΩConvFac = 2π*500.13;
{δA, δB, δC} = δOList;
Ω = δ*δToΩConvFac;

```



```

ΩA = δA*δToΩConvFac; ΩB = δB*δToΩConvFac; ΩC = δC*δToΩConvFac;
{kAB, kBA, kAC, kCA, kBC, kCB} = kList;
πA = kBA kCA + kBC kCA + kBA kCB;
πB = kAB kCA + kAB kCB + kAC kCB;
πC = kAC kBA + kAB kBC + kAC kBC;
{RA, RB, RC} = R2List;
{pA, pB, pC} = {πA, πB, πC}/(πA + πB + πC);
A = pA (RB RC + RB (kAC + kCA + kCB) + RC (kBA + kBC + kAB) - (Ω - ΩB) (Ω - ΩC))
  + pB (RA RC + RA (kCA + kCB + kBC) + RC (kAB + kAC + kBA) - (Ω - ΩA) (Ω - ΩC))
  + pC (RA RB + RA (kBA + kBC + kCB) + RB (kAB + kAC + kCA) - (Ω - ΩA) (Ω - ΩB))
  + πA + πB + πC;
B = pA ((RC + kCA + kCB + kAC) (Ω - ΩB) + (RB + kBA + kBC + kAB) (Ω - ΩC))
  + pB ((RC + kCA + kCB + kBC) (Ω - ΩA) + (RA + kAB + kAC + kBA) (Ω - ΩC))
  + pC ((RB + kBA + kBC + kCB) (Ω - ΩA) + (RA + kAB + kAC + kCA) (Ω - ΩB));
C = RA RB RC + (kCA + kCB) RA RB + (kBA + kBC) RA RC + (kAB + kAC) RB RC
  - (RC + kCA + kCB) (Ω - ΩA) (Ω - ΩB)
  - (RB + kBA + kBC) (Ω - ΩA) (Ω - ΩC)
  - (RA + kAB + kAC) (Ω - ΩB) (Ω - ΩC)
  + RA πA + RB πB + RC πC;
D = (RB RC + (kCA + kCB) RB + (kBA + kBC) RC + πA) (Ω - ΩA)
  + (RA RC + (kCA + kCB) RA + (kAB + kAC) RC + πB) (Ω - ΩB)
  + (RA RB + (kBA + kBC) RA + (kAB + kAC) RB + πC) (Ω - ΩC)
  - (Ω - ΩA) (Ω - ΩB) (Ω - ΩC);
MO (A C + B D)/(C^2 + D^2) ];
(* the use is illustrated in the following lines *)
deltaData = Range[0,2,.005];
absorptionSpectrum = threeStateExchReFunc[deltaData,{(*δA*) .7,(*δB*)1,(*δC*)1.4},
{(*RA*)40,(*RB*)40,(*RC*)40},{(*kAB*)200,(*kBA*)80,(*kAC*)100,(*kCA*)150,
(*kBC*)70,(*kCB*)80},{*M0*}10];

```

Exchange lineshapes for $N > 3$ can be readily derived in *Mathematica* software, which will be shown in Sec. B.2.

B.1.4 Spectral lineshape of two J-coupled spins

Two J-coupled nuclear spins $I = 1/2$ with offset Larmor frequencies Ω_A and Ω_B produce NMR spectrum with two doublets as shown in Fig. B.2 [76, 81]. If $\Omega_A = \Omega_B$, no splitting is observable. The coupling constant J (in Hz) describes strength of indirect dipole-dipole interaction (Sec. 3.1.4), two strongly interacting spins form so-called ‘AB system’ and weakly interacting spins form so-called ‘AX system’¹ ($|J| \ll |\Omega_A - \Omega_B|$). AB system shows so-called ‘roof effect’ when the inner peaks are more intensive than the outer peaks ($I_{\text{in}} > I_{\text{out}}$), see Fig. B.2, while peaks of AX system are equally intensive.

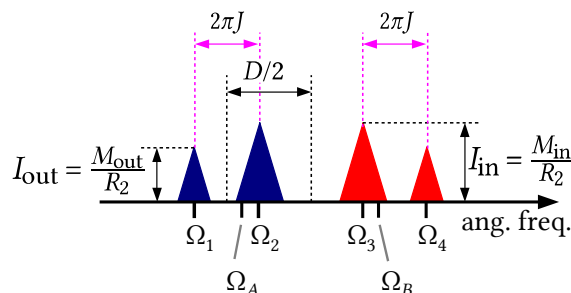


Figure B.2: Illustration of spectral lineshape two J-coupled spins.

¹Strength of J-coupling depends on the distance between spins, therefore A and B being close in the alphabet indicate strong interaction, A and X being far indicate weak interaction.

Spectral lineshape for J-coupled spins cannot be derived from Bloch equations, it requires quantum mechanical approach. The solution in the case of homonuclear AB system and $R_2^A = R_2^B = R_2$ can be found in [76] (page 615) and consists of four Lorentzian peaks,

$$S(\Omega) = \frac{M_{\text{out}}}{\alpha_1} + \frac{M_{\text{in}}}{\alpha_2} + \frac{M_{\text{in}}}{\alpha_3} + \frac{M_{\text{out}}}{\alpha_4}, \quad (\text{B.15})$$

where

$$\begin{aligned} \alpha_1 &= R_2 - i(\Omega - \Omega_1), \\ \alpha_2 &= R_2 - i(\Omega - \Omega_2), \\ \alpha_3 &= R_2 - i(\Omega - \Omega_3), \\ \alpha_4 &= R_2 - i(\Omega - \Omega_4), \\ \Omega_1 &= \frac{\Omega_A + \Omega_B}{2} + \frac{1}{2}(D + 2\pi J), \\ \Omega_2 &= \frac{\Omega_A + \Omega_B}{2} + \frac{1}{2}(D - 2\pi J), \\ \Omega_3 &= \frac{\Omega_A + \Omega_B}{2} - \frac{1}{2}(D - 2\pi J), \\ \Omega_4 &= \frac{\Omega_A + \Omega_B}{2} - \frac{1}{2}(D + 2\pi J), \\ D &= \sqrt{(\Omega_A - \Omega_B)^2 + (2\pi J)^2}. \end{aligned} \quad (\text{B.16a})$$

As shown in Fig. B.2, peaks and Ω_1 and Ω_2 as well as Ω_3 and Ω_4 are separated by $2\pi J$ in the angular frequency scale, which corresponds to splitting by \tilde{J} [ppm] = J [Hz]/ ν_0 [MHz] in the ppm scale, ν_0 is the spectrometer frequency. Furthermore, $I_{\text{in}} = M_{\text{in}}/R_2$ and $I_{\text{out}} = M_{\text{out}}/R_2$. Intensities of inner and outer peaks satisfy

$$\frac{M_{\text{in}}}{M_{\text{out}}} = \frac{I_{\text{in}}}{I_{\text{out}}} = \frac{D + 2\pi J}{D - 2\pi J}, \quad (\text{B.17a})$$

$$M_{\text{in}} + M_{\text{out}} = M_0, \quad (\text{B.17b})$$

where M_0 is equilibrium magnetization due to a single spin. Furthermore, a few other relationships can be derived from the above equations,

$$M_{\text{in}} = \frac{D + 2\pi J}{2D} M_0, \quad (\text{B.18a})$$

$$M_{\text{out}} = \frac{D - 2\pi J}{2D} M_0, \quad (\text{B.18b})$$

$$|\Omega_A - \Omega_B| = \sqrt{D^2 - (2\pi J)^2} = \sqrt{(\Omega_1 - \Omega_4)(\Omega_2 - \Omega_3)}. \quad (\text{B.18c})$$

Total area below the curve (in angular frequency scale) is $\pi R_2(2I_{\text{in}} + 2I_{\text{out}}) = 2\pi M_0$, which is the expected peak intensity due to two spins.

The *Mathematica* code is

(*J-coupled spins 1/2 lineshape*)

```
JcoupledSpinFunc[δ_, δ0List_, R2_, J_, M0_] :=
Module[{α1, α2, α3, α4, δA, δB, δToΩConvFac,
  Ω, ΩA, ΩB, Ω1, Ω2, Ω3, Ω4, D, Min, Mout},
δToΩConvFac = 2π*500.13;
{δA, δB} = δ0List;
Ω = δ*δToΩConvFac; ΩA = δA*δToΩConvFac; ΩB = δB*δToΩConvFac;
D = Sqrt[(ΩA - ΩB)^2];
```

```

Ω1 = 1/2 (ΩA + ΩB) + 1/2 (D + 2π J);
Ω2 = 1/2 (ΩA + ΩB) + 1/2 (D - 2π J);
Ω3 = 1/2 (ΩA + ΩB) - 1/2 (D - 2π J);
Ω4 = 1/2 (ΩA + ΩB) - 1/2 (D + 2π J);
α1 = R2 - I (Ω - Ω1); α2 = R2 - I (Ω - Ω2);
α3 = R2 - I (Ω - Ω3); α4 = R2 - I (Ω - Ω4);
Min = (D + 2π J)/(2 D) M0; Mout = (D - 2π J)/(2 D) M0;
Re[Mout/α1 + Min/α2 + Min/α3 + Mout/α4] ];
deltaData = Range[0,2,.002];
absorptionSpectrum = JcoupledSpinFunc[deltaData, {(*δA*) .9, (*δB*) 1.1}, (*R2*) 40,
(*J*) 12, (*M0*) 10];

```

B.2 Rapid derivation of N -state exchange lineshape in *Mathematica*

Here we show a “copy-paste” code in the *Mathematica* software to rapidly derive the four- and five-state exchange spectral lineshapes. However, it can be readily extended to an arbitrary number of states. We do not show the actual formulae generated by the code due to the excessive number of terms contained.

```

(* four-state exchange lineshape *)
Kmatrix = {{-kAB - kAC - kAD, kBA, kCA, kDA}, {kAB, -kBA - kBC - kBD, kCB, kDB},
{kAC, kBC, -kCA - kCB - kCD, kDC}, {kAD, kBD, kCD, -kDA - kDB - kDC}};
Amatrix = DiagonalMatrix[{αA, αB, αC, αD}];
(* three following command lines: keeping together the terms on diagonal (βj)
of the matrix B=A-K reduces the size of the final symbolic expression *)
Bmatrix = Amatrix - Kmatrix;
betaRule = MapThread[Rule[#1, #2] &,
{{βA, βB, βC, βD}, Diagonal[Bmatrix]}]; (* later restores the diagonal
values of B *)
Bmatrix = Bmatrix - DiagonalMatrix[Diagonal[Bmatrix]] +
DiagonalMatrix[{βA, βB, βC, βD}]; (* replaces diagonal values with βj *)
{fA, fB, fC, fD} = M0 Inverse[Bmatrix].{pA, pB, pC, pD};
fourStateLineshapeComplex = (Total[{fA, fB, fC, fD}] // Together) /. betaRule;
alphaRule = {αA -> RA + I (Ω - ΩA), αB -> RB + I (Ω - ΩB),
αC -> RC + I (Ω - ΩC), αD -> RD + I (Ω - ΩD)};
fourStateLineshapeComplex = fourStateLineshapeComplex /. alphaRule; (* substitute
for αj *)
(* in order to obtain the absorption part of the spectrum,
numerical values of the parameters must be inserted and then the real part
taken, see the example in the following lines *)
absorptionSpectrum = fourStateLineshapeComplex /. {Ω -> Range[4000],
ΩA -> 1000., ΩB -> 1500., ΩC -> 2500., ΩD -> 3000.,
M0 -> 1., pA -> 1/4, pB -> 1/4, pC -> 1/4, pD -> 1/4,
kAB -> 30., kBA -> 30., kAC -> 30., kCA -> 30., kAD -> 30., kDA -> 30.,
kBC -> 30., kCB -> 30., kBD -> 30., kDB -> 30., kCD -> 30., kDC -> 30.,
RA -> 30., RB -> 30., RC -> 30., RD -> 30.};
absorptionSpectrum = Re[absorptionSpectrum];

(* five-state exchange lineshape *)
Kmatrix = {{-kAB - kAC - kAD - kAE, kBA, kCA, kDA, kEA},
{kAB, -kBA - kBC - kBD - kBE, kCB, kDB, kEB},
{kAC, kBC, -kCA - kCB - kCD - kCE, kDC, kEC},
{kAD, kBD, kCD, -kDA - kDB - kDC - kDE, kED},

```

```

{kAE, kBE, kCE, kDE, -kEA - kEB - kEC - kED}};
Amatrix = DiagonalMatrix[{αA, αB, αC, αD, αE}];
(* three following command lines: keeping together the terms on diagonal (βj)
   of the matrix B=A-K reduces the size of the final symbolic expression *)
Bmatrix = Amatrix - Kmatrix;
betaRule = MapThread[Rule[#1, #2] &,
  {{βA, βB, βC, βD, βE}, Diagonal[Bmatrix]}]; (* later restores the
   diagonal values of B *)
Bmatrix = Bmatrix - DiagonalMatrix[Diagonal[Bmatrix]] +
  DiagonalMatrix[{βA, βB, βC, βD, βE}]; (* replaces diagonal values with βj *)
{fA, fB, fC, fD, fE} = MO Inverse[Bmatrix].{pA, pB, pC, pD, pE};
fiveStateLineshapeComplex = (Total[{fA, fB, fC, fD, fE}] // Together) /. betaRule;
alphaRule = {αA -> RA + I (Ω - ΩA), αB -> RB + I (Ω - ΩB),
  αC -> RC + I (Ω - ΩC), αD -> RD + I (Ω - ΩD), αE -> RE + I (Ω - ΩE)};
fiveStateLineshapeComplex = fiveStateLineshapeComplex /. alphaRule; (* substitute
   for αj *)
(* in order to obtain the absorption part of the spectrum,
   numerical values of the parameters must be inserted and then the real part
   taken, see the example in the following lines *)
absorptionSpectrum = fiveStateLineshapeComplex /. {Ω -> Range[4000],
  ΩA -> 1000., ΩB -> 1500., ΩC -> 2500., ΩD -> 3000., ΩE -> 3300.,
  MO -> 1., pA -> 1/4, pB -> 1/4, pC -> 1/4, pD -> 1/8, pE -> 1/8,
  kAB -> 30., kBA -> 30., kAC -> 30., kCA -> 30., kAD -> 30., kDA -> 30.,
  kAE -> 30., kEA -> 30., kBC -> 30., kCB -> 30., kBD -> 30., kDB -> 30.,
  kBE -> 30., kEB -> 30., kCD -> 30., kDC -> 30., kCE -> 30., kEC -> 30.,
  kDE -> 30., kED -> 30., RA -> 30., RB -> 30., RC -> 30.,
  RD -> 30., RE -> 30.};
absorptionSpectrum = Re[absorptionSpectrum];

```

B.3 Rapid derivation of three-state populations

Steady-state populations are derived with the following *Mathematica* code

```

(* three-state steady-state populations *)
Kmatrix = {{-kAB - kAC, kBA, kCA}, {kAB, -kBA - kBC, kCB},
  {kAC, kBC, -kCA - kCB}};
pvector = {pA, pB, pC};
equations = MapThread[Equal[#1, #2] &,
  {Kmatrix . pvector, ConstantArray[0, Length@pvector]}];
ssRule = Solve[(And @@ equations)~And~(Total@pvector == 1),
  pvector][[1]] // Factor

```

This code can be extended to higher number of states.

Three-state equilibrium populations are derived from the steady-state populations with additional equilibrium condition $k_{AB} = \frac{k_{AC}k_{BA}k_{CB}}{k_{BC}k_{CA}}$ (Eq. (3.39)). The corresponding code reads

```

(* three-state equilibrium populations *)
eqCondRule = {kAB -> kAC kBA kCB/(kBC kCA)}
ssRule /. eqCondRule // Factor

```

C Experimental details, models and data analysis

C.1 Properties of used chemicals

Quantities characterizing the compounds used in this study are given in Table C.1. They were used for calculation of concentrations during the experiments. Molar mass of oxoporphyrinogens was calculated in the *Chem Draw Pro* program. **OxP** and its derivatives were synthesized in National Institute for Materials Science (NIMS), Japan, by Jonathan Hill. Other compounds were purchased from Sigma-Aldrich.

Table C.1: Basic properties of chemicals used in this work.

compound	molar mass / g.mol ⁻¹	density / g.mL ⁻¹	notes
OxP	1125.57		powder
Bz₂OxP	1463.61		powder
Bz₄OxP	1801.68		powder
CSA	232.3	1.331	powder solubility in CDCl ₃ : 2.3 mg/mL [CSA] _{max} ≈ 0.01 M
DFA	96.03	1.526	liquid
TFA	114.0	1.489	liquid
CDCl ₃	120.4	1.5	μ ^{el} = 1.15 D
DMF	73.10	0.948	basic, pK _a ≈ -0.3 for conjugate acid, μ ^{el} =3.86 D

Full names of the investigated oxoporphyrinogens are *meso*-tetrakis(3,5-di-*tert*-butyl-4-oxo-2,5-cyclohexadienylidene)porphyrinogen (**OxP**), N₂₁,N₂₃-bis(4-bromobenzyl)-*meso*-tetrakis(3,5-di-*tert*-butyl-4-oxo-2,5-cyclohexadienylidene)porphyrinogen (**Bz₂OxP**) and N₂₁,N₂₂,N₂₃,N₂₄-tetrakis(4-bromobenzyl)-*meso*-tetrakis(3,5-di-*tert*-butyl-4-oxo-2,5-cyclohexadienylidene)porphyrinogen (**Bz₄OxP**). Numbering of the porphyrinogen macrocycle is shown in Fig. 1.2d.

C.2 UV/vis spectroscopy measurements

UV/vis absorption spectra were recorded using Jasco J-820, Shimadzu UV-3600, and Hitachi U-2910 spectrophotometers (1 cm or 1 mm quartz cells with screw caps were used). Baseline, determined by calibration measurement with empty cuvette, was subtracted from each spectrum. All UV/vis measurements were conducted at room temperature. Titration experiments (i.e. measurement of spectra at different guest concentrations when the guest is gradually added into the host solution) were performed by the addition of stock solution (acid dissolved in chloroform, in some cases in the presence of oxoporphyrinogen with the same concentration as in the sample to keep the total dye concentration constant) into the sample using a microsyringe. Deuterated spectroscopic chloroform was used as solvent. Spectral calculator [176] was used for the conversion of the measured UV/vis spectra into the corresponding RGB values.

Measurements of solvatochromic effects in **Bz₂OxP** and **Bz₄OxP** in DMF/CDCl₃ mixture were started in the pure solution of oxoporphyrinogen in CDCl₃. Then, appropriate amount of

the sample was gradually removed and replaced from the stock solution containing DMF (stock solution also contained the oxoporphyrinogen to prevent dilution), total sample volume was kept constant. Prior the sample preparation, the solvent was filtered through potassium carbonate to remove all residual acid molecules (which are sometimes created by solvent decomposition). Deuterated spectroscopic chloroform with silver foil was used to prevent formation of HCl. In these cases, titration is not appropriate because the complete picture of spectral changes should be obtained in the whole vol.f. range ($\varphi(\text{DMF}) = 1$ is inaccessible in titration starting in pure chloroform). Measurements of solvatochromic effects in **OxP** were investigated by titration of **OxP** in CDCl_3 with DMF similarly to the previous paragraph (solvatochromic spectral change is saturated at $\varphi(\text{DMF}) = 0.2$, which is accessible in titration).

C.3 RGB coordinates of UV/vis spectra

Conversion of UV/vis spectra to corresponding RGB coordinates was done using Excel spreadsheet developed by Bruce Lindbloom (Spectral Calculator) [176]. Input data in this spreadsheet are in the form of transmittance T , which was calculated from the measured absorbance A using $T = 10^{-A}$. Output format was set to Adobe RGB (1998). RGB values of spectra of individual absorbing components obtained by titration of oxoporphyrinogens with DFA are listed in Table C.2.

	R	G	B
OxP	177	92	170
OxP²⁺	212	152	166
OxP⁴⁺	151	186	160
Bz₂OxP	217	129	171
Bz₂OxP⁺	204	137	197
Bz₂OxP²⁺	157	169	206
Bz₂OxP⁴⁺	147	196	192
Bz₄OxP	244	165	173
Bz₄OxP H-bond	225	135	184
Bz₄OxP⁺	195	168	194
Bz₄OxP²⁺	121	195	213

Table C.2: RGB values of protonated oxoporphyrinogen species (listed in Fig. 6.3, 6.4 and 6.5) calculated from measured absorbance using [176].

C.4 NMR spectroscopy measurements

NMR spectroscopy measurements were performed using Bruker Avance III HD 500 and JEOL AL300BX spectrometers. Tetramethylsilane was used as the internal standard, chloroform with silver foil was used as solvent to prevent the formation of HCl. Titration experiments (i.e. measurement of spectra at different guest concentrations when the guest is gradually added into the host solution) were performed by the addition of stock solution (acid dissolved in chloroform, in some cases in the presence of oxoporphyrinogen with the same concentration as in the sample) into the sample using a microsyringe. All measurements were performed at 25 °C unless otherwise specified. Variable temperature measurements were performed at the Bruker spectrometer, refrigerator was used for cooling up to ca. -40 °C, liquid nitrogen was used to reach lower temperatures.

C.5 NMR analysis of host-guest binding

In order to obtain useful information presented in Chapter 6, the measured spectra were analyzed in two steps:

Step 1 – Lineshape fitting:

Fitting of all relevant peaks in raw data (after proper phase correction and subtraction of baseline with polynomial curve fitted outside peaks) using appropriate formulae: Lorentzian, J-coupled or exchange lineshapes. This type of fitting is illustrated in Fig. C.1. The spectra were not fitted in the whole ppm range, but rather only in the close vicinity of the analyzed peaks. Partially overlapping peaks can be fitted without problems. Fitting of the exchange lineshapes requires careful treatment of δ_j and R_2^j parameters, which is discussed in next sections.

Fitting of *tert*-butyl resonances of the host, methyl group resonances of the guest (averaged signal due to fast exchange between **G** and **HG** in both CSA and DFA), and resonance due to water (averaged signal due to fast exchange between **W** and **HW**) also provided accurate number of equivalents of guest and water (with respect to host) and consequently their total concentrations using $[G]_t = \text{equiv.}(\mathbf{G}) \times [\mathbf{H}]_t$ and $[W]_t = \text{equiv.}(\mathbf{W}) \times [\mathbf{H}]_t$. Total concentration of host was calculated from the weight of the sample in powder form and the corresponding volume of solvent (the effect of dilution during the titration was also incorporated). The lineshape fitting procedure was conducted in *Excel*.

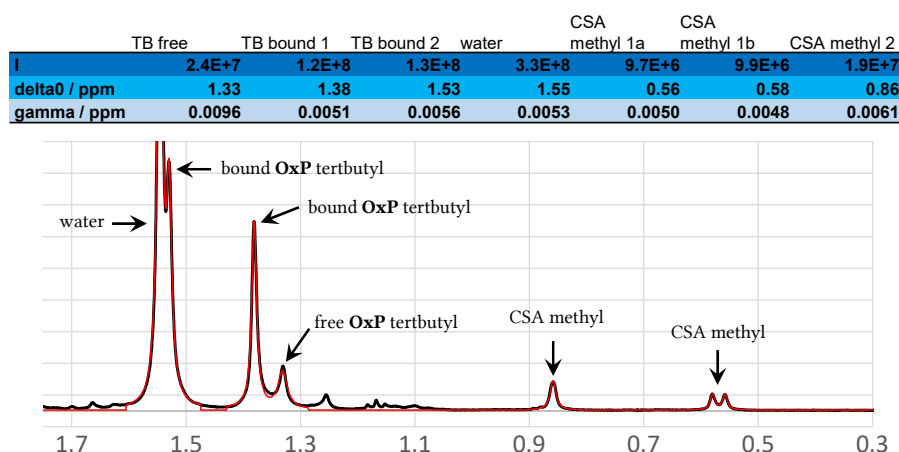


Figure C.1: Illustration of fitting the measured spectra (**OxP** + (*R*)-CSA). Experimental data (black) are successfully reproduced by the Lorentzian fitting curves (red).

Step 2 – Fitting of binding isotherms:

Simultaneous least squares fitting of appropriate binding isotherms to the peak parameters obtained in Step 1 as every titration experiment shows several spectral changes (peak shifts or newly formed isolated peaks). Figure C.2 shows isotherms calculated from shift of averaged peaks in fast exchange that were fitted using Eq. (3.62) (isotherms 2,3,4) and isotherms constructed as ratio of peak areas that were fitted by Eq. (3.61) (isotherms 1,5,6). Proper binding model should be used in order to describe the system correctly. The lineshape fitting procedure was conducted in *Excel* using Solver add-in. Contributions of sum of squares due to particular isotherms to the total sum of squares are weighted to be in the same order as shown in Fig. C.3. This figure also shows all parameters used to describe the binding isotherms within the 1:4 host-guest binding model. The parameters with blue background were fitted, those with yellow background were fixed and obtained directly from the spectra. The least-square fitting procedure was conducted

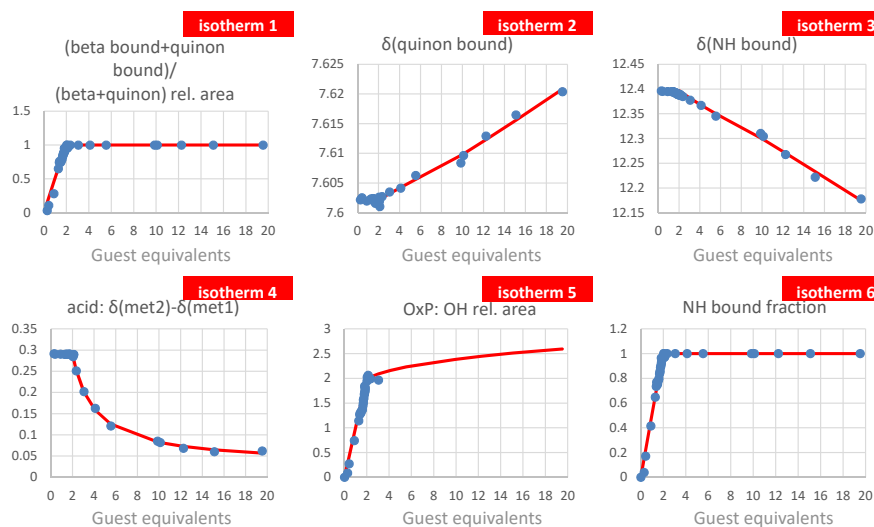


Figure C.2: Illustration of fitted binding isotherms.

in *Excel* using Solver add-in and VBA programming environment to implement calculations of binding isotherms. The fitted values are listed in Table 6.1. In order to obtain the corresponding

	isotherm 1	isotherm 2	isotherm 3	isotherm 4	isotherm 5	isotherm 6
sums of squares	6.9E-2	7.0E-6	9.4E-4	3.8E-4	4.0E+1	2.7E-2
weights	250	2000000	20000	50000	1	200
weighted SSq	17.24	14.06	18.77	18.85	40.48	5.38
TOTAL SSq	114.8					

OxP + R-CSA NMR	isotherm 2	isotherm 3	isotherm 4	
K1	4.8E+5	7.60	12.4	$\delta G(\text{CSA met2}) - \delta G(\text{CSA met1})$ 0.021
K2	1.1E+8	7.60	12.4	$\delta HG(\text{CSA met2}) - \delta HG(\text{CSA met1})$ 0.27
K3	131	7.61	12.2	$\delta HG2(\text{CSA met2}) - \delta HG2(\text{CSA met1})$ 0.29
K4	2.8	8.41	3.82	$\delta HG3(\text{CSA met2}) - \delta HG3(\text{CSA met1})$ 0.24
				$\delta HG4(\text{CSA met2}) - \delta HG4(\text{CSA met1})$ 1.46

Figure C.3: Illustration of simultaneous weighted fitting of binding isotherms (OxP + (R)-CSA) in *Excel*. Parameters with blue background were fitted, the other were obtained directly from the measured spectra.

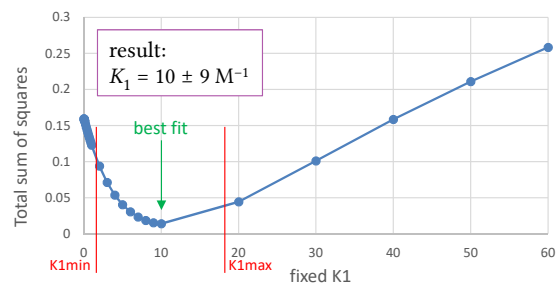


Figure C.4: Plot of total sum of squares for repeated fitting procedure with different fixed K_1 values (Bz_4OxP + DFA).

error, the fitting procedure was repeated several times with different fixed values of K_j . Acceptable fits were determined by visual inspection of fitted curves and subsequently confidence intervals were constructed from minimum and maximum values of these fitted parameters. This

procedure is illustrated in Fig. C.4 for the titration of **Bz₄OxP** with DFA.

C.6 Fitting of exchange lineshapes

This appendix describes describes fitting exchanging of *tert*-butyl-H signals with symmetric two-state exchange lineshape (Eq. (B.11)). Dependence on guest concentration or on temperature of the transition rate coefficient k_{obs} was determined. The fitting procedure was realized in *Mathematica* (using in-house scripts with FindFit command for the minimization procedure). Chemical shifts δ_A and δ_B and transverse relaxation rates R_2^A and R_2^B (including the inhomogeneous broadening effect, see Sec. 3.1.3) have a big influence on the resulting exchange coefficients. However, they could not be fixed at one particular value throughout the whole experiment since they might change with concentration or temperature.

Parameters δ_A and δ_B were obtained from the positions of peak maxima in the slow regime (at low guest concentration or at low temperature). During the titration experiment the parameters δ_A and δ_B might change with changing polarity of the solvent (e.g. upon addition of acid) or when new species form in the sample (which must not change the number of states). In this situation, the difference $|\delta_A - \delta_B|$ often remains nearly constant and only the center $(\delta_A + \delta_B)/2$ shifts with increasing guest concentration. On the other hand, variable temperature experiments imply significant change in δ_A and δ_B parameters including their difference, which makes the fitting procedure more difficult. Due to these reasons, change of δ_A and δ_B during the whole experiment is also given.

Relaxation rates R_2^A and R_2^B (which include homogeneous broadening) were also obtained at the slow regime from peak width 2γ (R_2 [s⁻¹] = $2\pi\nu_0$ [MHz] γ [ppm]). In some cases, $R_2^A = R_2^B$ was assumed. However, they might also vary with concentration or with temperature. During

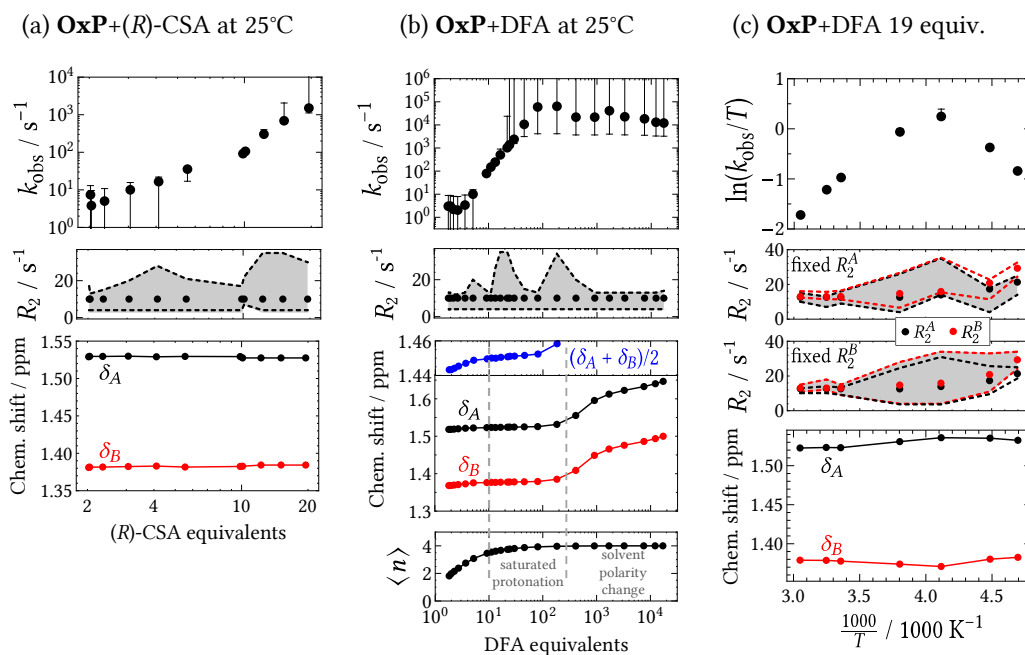


Figure C.5: Fitting of **OxP** *tert*-butyl signals with symmetric two-state exchange lineshape. (a) Titration of **OxP** with (R)-CSA (full spectra in Fig. E.3). (b) Titration of **OxP** with DFA (full spectra in Fig. E.2). (c) Variable temperature measurements at 18 equiv. of DFA (full spectra in Fig. E.9). The quantity $\langle n \rangle$ stands for mean mrotonation number.

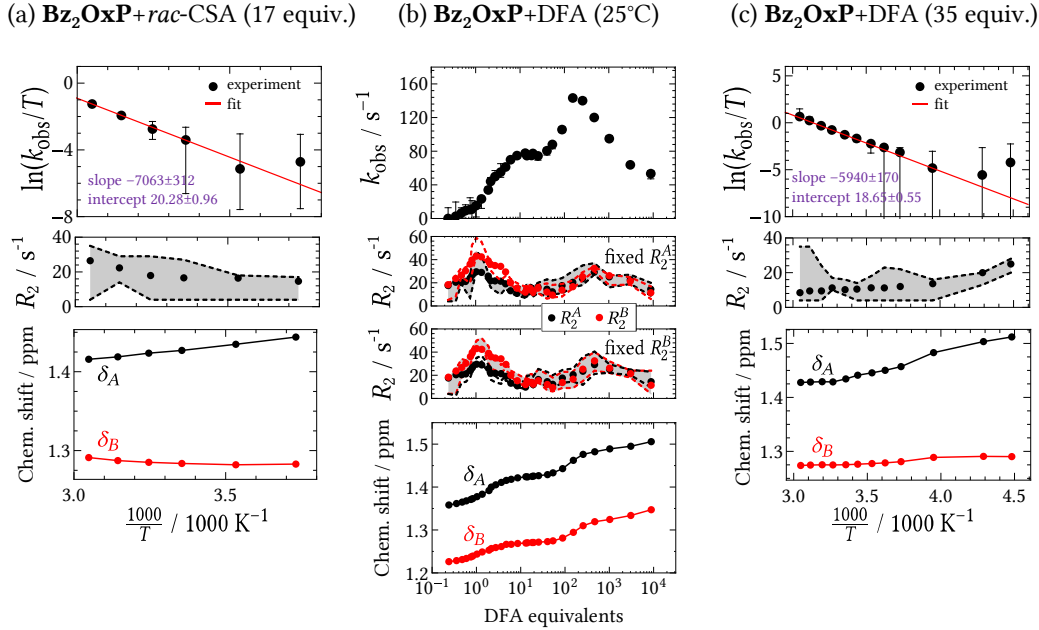


Figure C.6: Fitting of Bz_2OxP tert-butyl signals with symmetric two-state exchange lineshape. (a) Variable temperature measurements at 17 equiv. of (rac)-CSA (full spectra in Fig. E.15). (b) Titration of Bz_2OxP with DFA (full spectra in Fig. E.4). (c) Variable temperature measurements at 35 equiv. of DFA (full spectra in Fig. E.11).

the fitting procedure, one of the relaxation rates were fixed. The parameters adjusted during the fitting procedure are: k_{obs} , one of transverse relaxation rates R_2^A or R_2^B (only if they are not equal) and total magnetization in relative units M_0 (here, M_0 plays the role of arbitrary proportionality constant). To obtain the confidence intervals, fitting was repeated for every measured spectrum for different fixed R_2 values, which were varied in a reasonable range of $R_2 \in [10 \text{ s}^{-1}, 40 \text{ s}^{-1}]$. Quality of the fits to the experimental data was determined upon visual inspection, fixed R_2 parameters providing a good fit are denoted by gray zone in Fig. C.5 and C.6. Error bars of fitted k_{obs} in these figures correspond to maximum or minimum R_2 inside the gray zone. Mean values of k_{obs} correspond to best fit (with minimal sum of squares) or to the fixed R_2 obtained from peak width at slow exchange.

Analysis of OxP + acid systems is shown in Fig. C.5. Fitting of titration of OxP with (R)-CSA in Fig. C.5a assumed $R_2^A = R_2^B$, δ_A and δ_B were almost constant during the whole experiment. Fitting of titration of OxP with DFA in Fig. C.5b also assumed $R_2^A = R_2^B$, $|\delta_A - \delta_B|$ was kept constant during the whole experiment. However, the center $(\delta_A + \delta_B)/2$ slightly shifted due to formation of HG_4 complex at 10–300 equiv. of DFA until saturation (blue curve, number of spin states in HG_4 does not change compared to HG_2), furthermore, the center is shifted by a greater extent above 300 equiv. due to solvent polarity change. Fitting of variable temperature measurement of OxP with 18 equiv. of DFA in Fig. C.5c showed different values of R_2^A or R_2^B . Therefore, in the first part of the analysis, R_2^A was fixed and R_2^B was fitted together with k_{obs} and M_0 , which is shown in the second panel. Successful fits correspond to R_2^A within the gray zone, the fitted values of R_2^B are enclosed by red dashed line. Mean values of R_2^A (black circles) correspond to mean values of R_2^B (red circles). In the second part of the analysis, the roles of R_2^A and R_2^B were swapped as shown in the third panel. Confidence intervals in the Eyring plot in the first panel correspond to intersection of k_{obs} confidence intervals obtained in the second and third panel. Values of δ_A and δ_B varied during the experiment.

Analysis of **Bz₂OxP** + acid systems is shown in Fig. C.6. Fitting of variable temperature measurement of **Bz₂OxP** with 17 equiv. of (*rac*)-CSA in Fig. C.6a assumed $R_2^A = R_2^B$, δ_A and δ_B significantly varied during the whole experiment. Linear Eyring plot (see Sec. 2.1.4) in first panel yielded (using weighted linear least squares) the energy barrier parameters listed in Table 7.1, the following *Mathematica* code has been used

```
(* Eyring plot ln(k/T)=a 1/T + b *)
ΔH‡ = -R a/1000 /. R -> 8.3145 (* in kJ/mol *)
ΔS‡ = -R (b - Log[kBoltz/h])/1000 /. {R -> 8.3145,
kBoltz -> 1.3806*10^-23, h -> 6.6261*10^-34} (* in kJ/mol/K *)
```

Fitting of titration of **Bz₂OxP** with DFA in Fig. C.6b showed different values of R_2^A or R_2^B . Fitting was performed for range of fixed R_2^A values and then for fixed R_2^B values as shown in second and third panel, respectively. However, the “hillocks” in R_2^A and R_2^B values below 20 equiv. of DFA (corresponding to $\langle n \rangle < 1$) in Fig. C.6b indicate that two-state fitting model is not appropriate, which causes the failed linear dependence in Fig. 7.9. In this concentration range, population

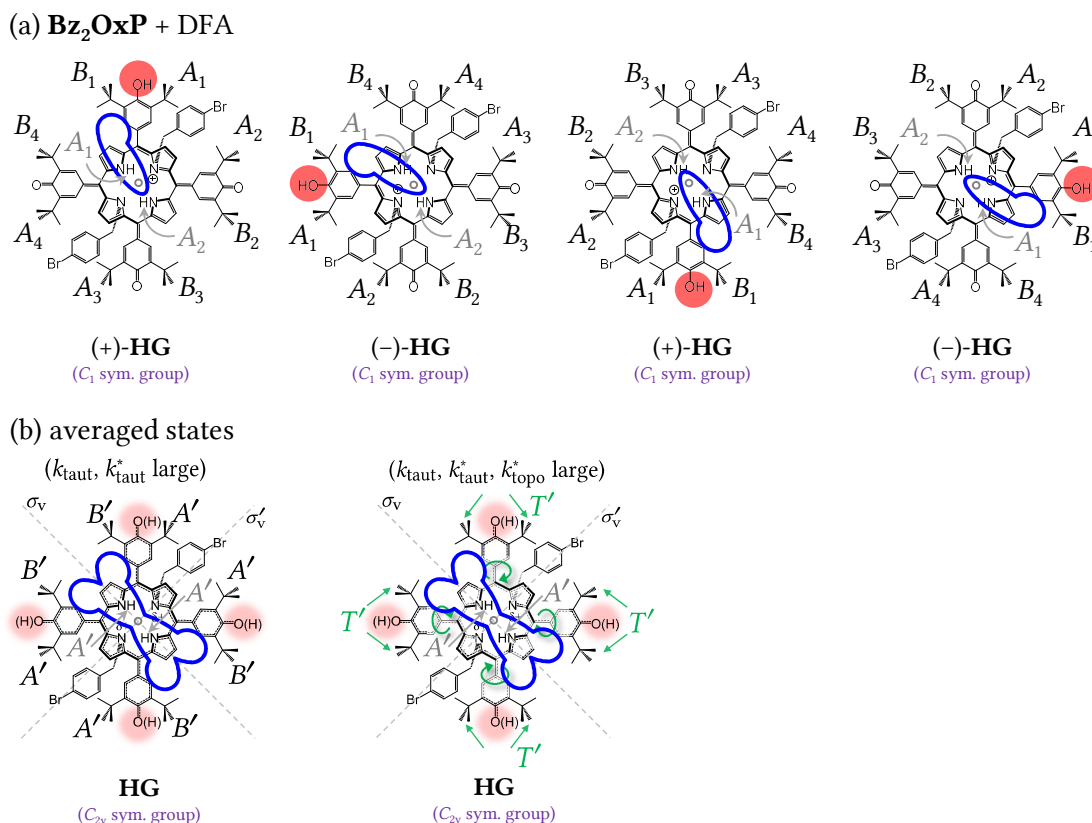


Figure C.7: Structures of **Bz₂OxP** + DFA complex, corresponding states of *tert*-butyl-H spins (black labels; also valid for *ortho*-H and β -H spins) and states of NH spins (gray labels). (a) Spin states in **Bz₂OxP**⁺·DFA complex, (b) Fast prototropic tautomerization processes cause averaging of chiral field in **Bz₂OxP** + DFA system (consists of averaged states from (a)). Slow topomerization process (left) induces two states (i.e. A' and B') of *tert*-butyl-H and *ortho*-H spins, fast topomerization process (right) induces one averaged state (i.e. T'). In the latter case, topomerization does not influence states of β -H spins. Geometric shapes with proper symmetry represent chiral field created by the acid anion. Dashed straight lines denote symmetry elements σ_v (vertical mirror plane). Fast exchange implies (Eq. (3.62)) in *tert*-butyl-H peaks: $\delta_{A'} = \frac{1}{4}(\delta_{A_1} + \delta_{A_2} + \delta_{A_3} + \delta_{A_4})$, $\delta_{B'} = \frac{1}{4}(\delta_{B_1} + \delta_{B_2} + \delta_{B_3} + \delta_{B_4})$ and $\delta_{T'} = \frac{1}{2}(\delta_{A'} + \delta_{B'})$. Fast exchange implies in NH peaks: $\delta_{A'} = \frac{1}{2}(\delta_{A_1} + \delta_{A_2})$.

of free host is not negligible and broadens the signals. Three-state exchange fitting is needed to obtain corrected k_{obs} values. Parameters δ_A and δ_B significantly varied but their difference remained almost constant. Structures and corresponding states of the **Bz₂OxP** + DFA system are shown in Fig. C.7. Fitting of variable temperature measurement of **Bz₂OxP** with 35 equiv. of DFA in Fig. C.6c assumed $R_2^A = R_2^B$, δ_A and δ_B significantly varied during the whole experiment. Linear Eyring plot (see Sec. 2.1.4) in first panel yielded (using weighted linear least squares) the energy barrier parameters listed in Table 7.1.

C.7 Fitting of experimental data in Bz₂OxP + (R)-CSA system

The following text describing data analysis from Sec. 7.2.3 is taken from our paper [2], where files with implementation in *Excel* and *Mathematica* can also be found. The fitting procedure comprises of two consecutive steps:

Step 1 – Lineshape fitting:

Fitting of tertiary butyl (TB) resonances of the host, methyl (MET) group resonances of the guest (averaged signal due to fast exchange between **G** and **HG**), and resonance due to water (averaged signal due to fast exchange between **W** and **HW**) with Lorentzian lineshapes was performed in order to provide accurate number of equivalents of guest and water (with respect to host) and consequently their total concentrations. This fitting procedure was realized in *Excel* (with Solver add-in using iterative generalized reduced gradient (GRG) nonlinear method for minimization). Total concentration of host was calculated from the weight of the sample in powder form and the corresponding volume of solvent (the effect of dilution during the titration was also incorporated). Values of $[\mathbf{G}]_t$, $[\mathbf{W}]_t$ and $[\mathbf{H}]_t$ were interpolated and served for calculation of continuous values of all transition rate coefficients and concentrations shown in Fig. 7.12 by solid lines.

Symmetric two-state exchange lineshape (Eq. (B.11)) and half-symmetric three-state exchange (Eq. (3.52) and (3.37)) lineshapes were fitted to the NH peaks of the host to obtain the concentration dependence of transition rate coefficients. Parameters $\delta_{A'_I}$, $\delta_{A'_{II}}$ and $\delta_{C'}$ were obtained from the positions of peak maxima in the slow regime. In the intermediate and fast regime the Larmor frequencies shifted due to variation in solvent polarity, so the difference $|\delta_{A'_I} - \delta_{A'_{II}}|$ was fixed, while the center $(\delta_{A'_I} + \delta_{A'_{II}})/2$ was adjusted during the fitting, see gray dashed lines in Fig. 7.12a,b. Relaxation rates were assumed to be equal, i.e. $R_2^{A'_I} = R_2^{A'_{II}} = R_2^{C'}$. During the fitting procedure, the relaxation rates were fixed. The parameters adjusted during the fitting procedure are: $k_{A'}$ for two-state exchange, $k_{A'}$ and $k_{A'C'}$ for three-state exchange and total magnetization in relative units M_0 . This fitting procedure was realized in *Mathematica* (using in-house scripts with FindFit command for the minimization procedure). The excellent match of fits is illustrated in Fig. C.8a,c. The fitting procedure was repeated for every measured spectrum (at particular $[\mathbf{G}]_t$) for different R_2 values, which were varied in a reasonable range of $R_2 \in [10 \text{ s}^{-1}, 40 \text{ s}^{-1}]$, gray zone in Fig. C.8b,d (bottom). Upon visual inspection, the fitted lineshapes provided an excellent match to the experimental data in almost the whole range of the R_2 parameter (gray zone in Fig. C.8b,d (bottom)). Maximum and minimum values of the fitted transition rate coefficients $k_{A'}$, $k_{A'C'}$ and $k_{C'A'}$ form the error bars in Fig. C.8b,d (top). Mean $k_{A'}$, $k_{A'C'}$ and $k_{C'A'}$ values correspond to $R_2 = 30 \text{ s}^{-1}$. Figure 7.12c in the main manuscript shows that the transition rate coefficients in the intermediate exchange regime are determined with the highest accuracy since the spectral lineshape at this point is most sensitive to the change of the transition rate coefficients (see also section S7.4 in [2]).

Step 2 – Binding model fitting:

Simultaneous fitting of $k_{A'}$ concentration dependence together (using Eq. (C.8a)) with shift of

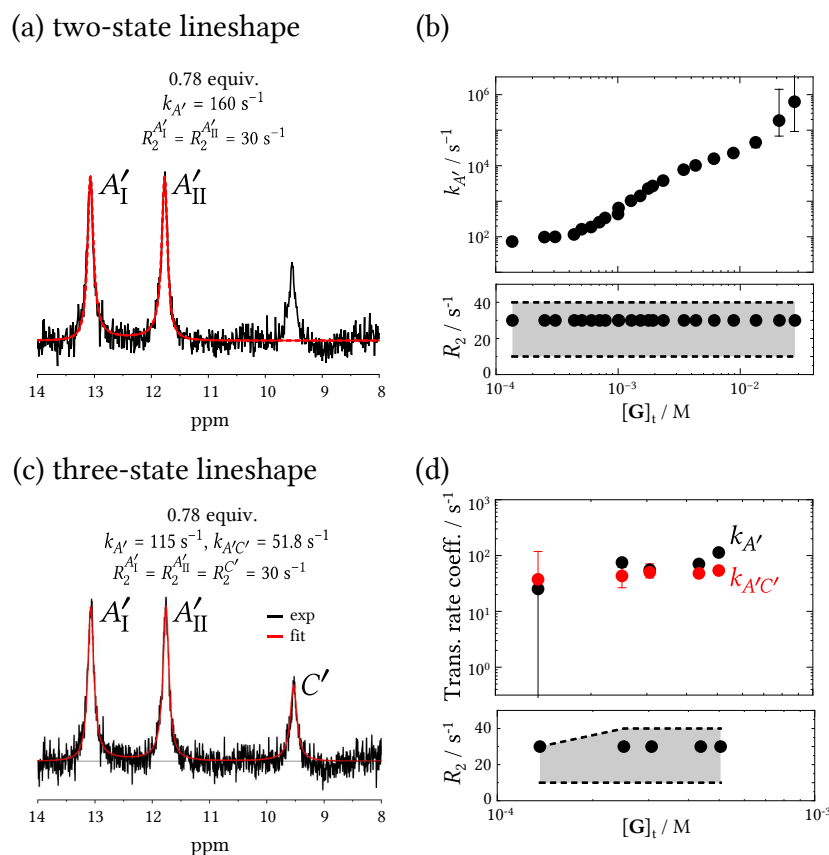


Figure C.8: Fitting of two- and three-state exchange spectral lineshape and determination of transition rate coefficients in Step 1. (a) Example of two-state fitting (spectrum at 0.78 equiv. of guest). Parameters $\delta_{A'_I}, \delta_{A'_II}$ and $R_2^{A'_I} = R_2^{A'_II}$ were fixed during the fitting of one particular spectrum. Values of $k_{A'}$ and M_0 are fitted parameters (value of M_0 differs from the real total magnetization since it depends on technical details of the experimental setup and data processing). (b) Resulting values of $k_{A'}$ from two-state fitting (top) and the range of the R_2 parameter, where the fit was successful (gray zone, bottom). Full circles (top and bottom) correspond to $R_2 = 30 \text{ s}^{-1}$, error bars denote maximum and minimum $k_{A'}$ values from repeated fitting procedure at different R_2 values in the range $R_2 \in [10 \text{ s}^{-1}, 40 \text{ s}^{-1}]$ (gray zone, bottom). (c) Example of three-state fitting. Parameters $\delta_{A'_I}, \delta_{A'_II}, \delta_{C'}$ and $R_2^{A'_I} = R_2^{A'_II} = R_2^{C'}$ were fixed during the fitting of one particular spectrum. Values of $k_{A'}, k_{A'C'}$ and M_0 are fitted parameters. (d) Resulting values of $k_{A'}$ and $k_{A'C'}$ from three-state fitting (top) and the range of the R_2 parameter (gray zone, bottom). Full circles (top and bottom) correspond to $R_2 = 30 \text{ s}^{-1}$, error bars denote maximum and minimum $k_{A'C'}, k_{C'A'}$ values from the repeated fitting procedure at different R_2 values in the range $R_2 \in [10 \text{ s}^{-1}, 40 \text{ s}^{-1}]$ (gray zone, bottom).

fast-exchanging peaks of \mathbf{H}/\mathbf{HW} (denoted as C' in Fig. 7.12b), two host TB resonances and two guest MET resonances (Fig. C.9) was performed. This fit provides equilibrium binding constant $K_{\mathbf{HG}}$, reaction rate coefficient κ_{faut} and Larmor frequencies of the peaks involved. The following

binding isotherms were fitted simultaneously

$$\begin{aligned}
 k_{A'} &= 2\kappa_{\text{taut}}[\mathbf{G}], \\
 \delta_{C'} &= \delta_C \frac{[\mathbf{H}]}{[\mathbf{H}] + [\mathbf{HW}]} + \delta_D \frac{[\mathbf{HW}]}{[\mathbf{H}] + [\mathbf{HW}]}, \\
 \delta(\mathbf{G} \text{ MET1 observed}) &= \delta(\mathbf{G} \text{ MET1 free}) \frac{[\mathbf{G}]}{[\mathbf{G}]_t} + \delta(\mathbf{G} \text{ MET1 bound}) \frac{[\mathbf{HG}]}{[\mathbf{G}]_t}, \\
 \delta(\mathbf{G} \text{ MET2 observed}) &= \delta(\mathbf{G} \text{ MET2 free}) \frac{[\mathbf{G}]}{[\mathbf{G}]_t} + \delta(\mathbf{G} \text{ MET2 bound}) \frac{[\mathbf{HG}]}{[\mathbf{G}]_t}, \\
 \delta(\mathbf{H} \text{ TB1 observed}) &= \delta(\mathbf{H} \text{ TB1 free}) \frac{[\mathbf{H}]}{[\mathbf{H}]_t} + \delta(\mathbf{H} \text{ TB1 bound}) \frac{[\mathbf{HG}]}{[\mathbf{H}]_t}, \\
 \delta(\mathbf{H} \text{ TB2 observed}) &= \delta(\mathbf{H} \text{ TB2 free}) \frac{[\mathbf{H}]}{[\mathbf{H}]_t} + \delta(\mathbf{H} \text{ TB2 bound}) \frac{[\mathbf{HG}]}{[\mathbf{H}]_t}.
 \end{aligned}$$

Concentrations of the chemical species were obtained from numerical solution of the competitive host-ligand binding model, summarized in Eq. (A.2) and (A.3). The fitting procedure was realized in *Mathematica* (using in-house scripts). In order to obtain error of all fitted parameters, the fitting procedure was repeated several times with different fixed values of parameters $K_{\mathbf{HG}}$, κ_A , δ_C or δ_D . Acceptable fits were determined by visual inspection of fitted curves and subsequently confidence intervals were constructed from minimum and maximum values of these fitted parameters. The parameter κ_{taut} was calculated as a mean, hence its error was determined as standard deviation. Furthermore, errors of other parameters, denoted here as σ_{\bullet} , were obtained from propagation of uncertainty. In particular, error of κ_{CA} (where $\kappa_{\text{ass}} = K_{\mathbf{HG}}\kappa_{\text{diss}}/4$) was calculated as $\sigma_{\kappa_{\text{diss}}} = \sqrt{(\frac{\kappa_{\text{diss}}}{4})^2 \sigma_{K_{\mathbf{HG}}}^2 + (\frac{K_{\mathbf{HG}}}{4})^2 \sigma_{\kappa_{\text{diss}}}^2}$, errors of Gibbs energy barriers ΔG_j^\ddagger (where $\Delta G_j^\ddagger = -RT \ln \frac{h\kappa_j}{k_B T}$) were calculated as $\sigma_{\Delta G_j^\ddagger} = \sqrt{(\frac{RT}{\kappa_j})^2 \sigma_{\kappa_j}^2}$ and errors of standard reaction Gibbs energies ΔG_j° (where $\Delta G_j^\circ = -RT \ln K_j$) were calculated as $\sigma_{\Delta G_j^\circ} = \sqrt{(\frac{RT}{K_j})^2 \sigma_{K_j}^2}$. All parameters used during the fitting procedures in Step 1 and Step 2 are listed in Table C.3

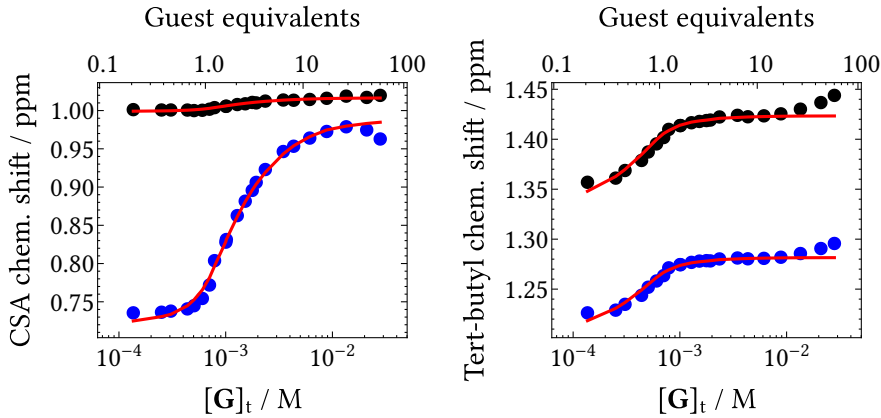


Figure C.9: Result of simultaneous fitting of chemical shifts of guest's MET resonances (left panel) and host's TB resonances (right panel) during the Step 2.

Table C.3: Overview of fitted and fixed parameters in the fitting procedures in Step 1 and Step 2.

Parameter	Value	Comment
$k_{A'C'}$	$47 \pm 6 \text{ s}^{-1}$	mean from raw data fit (Step 1)
$k_{A'}$	$10\text{--}10^5 \text{ s}^{-1}$	fitted (Step 1), changing with concentration
K_{HG}	$(7.4 \pm 3.0) \times 10^4 \text{ M}^{-1}$	fitted (Step 2), analysis from Sec. 6.9 yields $(8.0 \pm 5.0) \times 10^4 \text{ M}^{-1}$
K_{HW}	$240 \pm 35 \text{ M}^{-1}$	reported value [40], fixed during Step 2
κ_{taut}	$(10 \pm 1) \times 10^5 \text{ M}^{-1}\text{s}^{-1}$	fitted (Step 2), binding isotherm $k_{A'} = 2\kappa_{\text{taut}}[\text{G}]$
κ_{diss}	$47 \pm 6 \text{ s}^{-1}$	$\kappa_{\text{diss}} = k_{A'C'}$, mean from raw data fit (Step 1), fixed during Step 2
κ_{ass}	$(9 \pm 4) \times 10^5 \text{ M}^{-1}\text{s}^{-1}$	calculated, $\kappa_{\text{ass}} = K_{\text{HG}}\kappa_{\text{diss}}/4$,
$\delta_{A'_I}$	13.08 ppm	at low $[\text{G}]_t$; fixed (Step 1), obtained from peak maxima,
	< 13.08 ppm	at high $[\text{G}]_t$; fitted with fixed $\delta_{A'_I} - \delta_{A'_{II}}$ (Step 1)
$\delta_{A'_{II}}$	11.76 ppm	at low $[\text{G}]_t$; fixed (Step 1), obtained from peak maxima,
	< 11.76 ppm	at high $[\text{G}]_t$; fitted with fixed $\delta_{A'_I} - \delta_{A'_{II}}$ (Step 1)
$\delta_{C'}$	9.19–9.65 ppm	fixed, changing with concentration, obtained from the peak maxima (Step 1)
δ_C	$7.8 \pm 0.2 \text{ ppm}$	fitted (Step 2), reported value [40] $7.86 \pm 0.40 \text{ ppm}$
δ_D	$10.4 \pm 0.2 \text{ ppm}$	fitted (Step 2), reported value [40] $10.09 \pm 0.03 \text{ ppm}$
$\delta(\text{G MET1 free})$	$1.02 \pm 0.01 \text{ ppm}$	fitted (Step 2)
$\delta(\text{G MET1 bound})$	$1.00 \pm 0.01 \text{ ppm}$	fitted (Step 2)
$\delta(\text{G MET2 free})$	$0.99 \pm 0.01 \text{ ppm}$	fitted (Step 2)
$\delta(\text{G MET2 bound})$	$0.71 \pm 0.01 \text{ ppm}$	fitted (Step 2)
$\delta(\text{H TB1 free})$	$1.33 \pm 0.01 \text{ ppm}$	fitted (Step 2)
$\delta(\text{H TB1 bound})$	$1.42 \pm 0.01 \text{ ppm}$	fitted (Step 2)
$\delta(\text{H TB2 free})$	$1.20 \pm 0.01 \text{ ppm}$	fitted (Step 2)
$\delta(\text{H TB2 bound})$	$1.28 \pm 0.01 \text{ ppm}$	fitted (Step 2)

C.8 Kinetic models of **Bz₂OxP** + (*R*)-CSA system

Our model for the system of di-bromobenzylated oxoporphyrinogen (**H**) in the presence of (*R*)-camphorsulfonic acid (**G**) and water (**W**) combines the competitive host-ligand binding scheme for chemical kinetics (described in Sec. 2.2.3) and the half-symmetric three-state exchange model for the observed spin kinetics. The spin kinetics corresponds to the central NH nuclear spins located at the host.

There are multiple levels of description of the molecular kinetics in our system. The most elementary description is a simplified chemical kinetics scheme in Fig. C.10a, assuming the host molecule in three possible chemical states, **H**, **HG** and **HW**. The simplified chemical kinetics is governed by the following equations

$$\frac{d[\text{HG}]}{dt} = -\kappa_{\text{diss}}[\text{HG}] + 4\kappa_{\text{ass}}[\text{G}][\text{H}], \quad (\text{C.1a})$$

$$\frac{d[\text{H}]}{dt} = -4\kappa_{\text{ass}}[\text{G}][\text{H}] - \kappa_{CD}[\text{W}][\text{H}] + \kappa_{\text{diss}}[\text{HG}] + \kappa_{DC}[\text{HW}], \quad (\text{C.1b})$$

$$\frac{d[\text{HW}]}{dt} = -\kappa_{DC}[\text{HW}] + \kappa_{CD}[\text{W}][\text{H}]. \quad (\text{C.1c})$$

The factor of four in the term $4\kappa_{\text{ass}}[\text{G}][\text{H}]$ is caused by the presence of four possible protonation sites, increasing the rate of **HG** formation. The reaction rate coefficients are related to the equilibrium constants through the following formulae (obtained for the time derivatives in Eq. (C.1a))

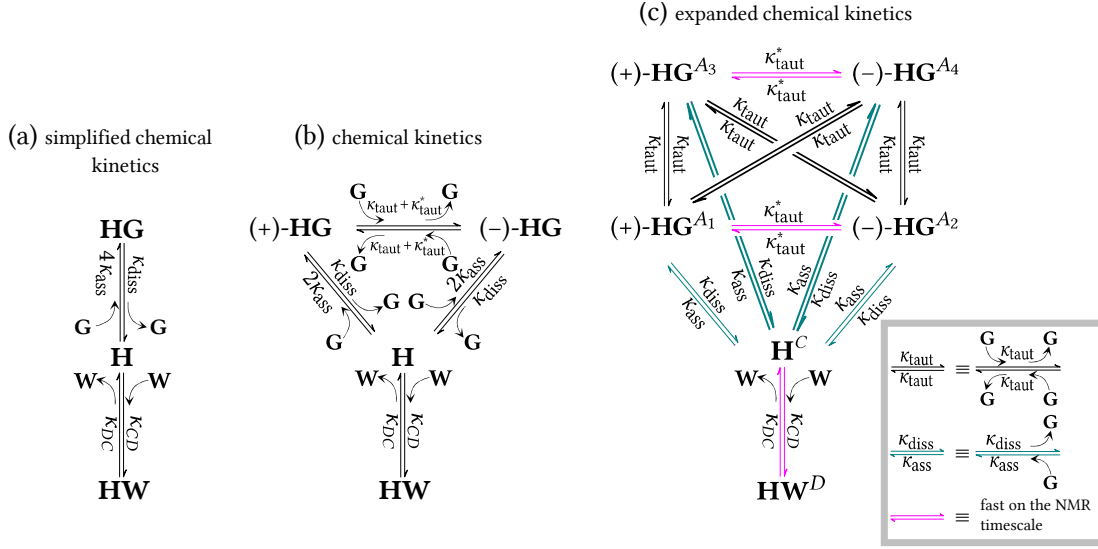


Figure C.10: Chemical kinetics schemes for the multi-state system of di-bromobenzylated oxoporphyrinogen (host **H**) in the presence of two potential ligands, (*R*)-camphorsulfonic acid (ligand **G**) and water (ligand **W**). (a) Simplified chemical kinetics scheme corresponding to 1:1 **H**:**G** binding with competitive 1:1 **H**:**W** binding described by Eq. (C.1). (b) Chemical kinetics scheme described by Eq. (C.4). (c) Expanded chemical kinetics scheme described by Eq. (C.5). All relevant molecular processes and their reaction rate coefficients are shown. Processes denoted by magenta arrows are fast on the NMR timescale.

and (C.1c) equal to zero)

$$K_{\text{HG}} = 4K_{\text{HG}}^{\text{micro}} = 4 \frac{k_{\text{ass}}}{k_{\text{diss}}}, \quad (\text{C.2a})$$

$$K_{\text{HW}} = \frac{k_{\text{CD}}}{k_{\text{DC}}}. \quad (\text{C.2b})$$

The stepwise binding constant, defined as $K_{\text{HG}} = [\text{HG}]/([\text{H}][\text{G}])$, describes protonation of any of the four possible protonation sites (hence the factor of four in Eq. (C.2a)). On the other hand, the microscopic binding constant $K_{\text{HG}}^{\text{micro}} = k_{\text{ass}}/k_{\text{diss}}$ describes binding to one particular binding site [48]. Thus, this constant corresponds to the “microscopic” standard reaction Gibbs energy $\Delta G_{\text{HG}}^{\text{micro}}$, which can be obtained from quantum chemistry calculations [177] as $\Delta G_{\text{HG}}^{\text{micro}} = \mu_{\text{HG}}^{\circ} - \mu_{\text{H}}^{\circ} - \mu_{\text{G}}^{\circ}$ (μ_{\bullet}° represents the standard chemical potential of species indicated in subscript). The quantity $\Delta G_{\text{HG}}^{\text{micro}}$ is higher than the standard reaction Gibbs energy $\Delta G_{\text{HG}}^{\circ} = -RT \ln K_{\text{HG}}$ (the quantity $\Delta G_{\text{HG}}^{\circ}$ is always defined by the latter equation from the stepwise equilibrium constant, but it should be correctly interpreted [46, 73]),

$$\Delta G_{\text{HG}}^{\text{micro}} = -RT \ln K_{\text{HG}}^{\text{micro}} = -RT \ln \frac{K_{\text{HG}}}{4} = \Delta G_{\text{HG}}^{\circ} + RT \ln 4. \quad (\text{C.3})$$

The numerical value is $\Delta G_{\text{HG}}^{\text{micro}} = \Delta G_{\text{HG}}^{\circ} + RT \ln 4 = -27.76 + 3.43 \text{ kJ}\cdot\text{mol}^{-1} = -24.33 \text{ kJ}\cdot\text{mol}^{-1}$ at $T = 298 \text{ K}$.

In fact, the protonated species **HG** are not identical as they form two enantiomers (+)-(**HG**) and (-)-(**HG**). The chemical kinetics is described by the scheme in Fig. C.10b, which is an expan-

sion of the scheme in Fig. C.10a. This scheme is governed by the following equations

$$\frac{d[(+)\text{-(HG)}]}{dt} = -(\kappa_{\text{taut}} + \kappa_{\text{taut}}^*)[\mathbf{G}][(+)\text{-(HG)}] - \kappa_{\text{diss}}[(+)\text{-(HG)}] + (\kappa_{\text{taut}} + \kappa_{\text{taut}}^*)[\mathbf{G}][(-)\text{-(HG)}] + 2\kappa_{\text{ass}}[\mathbf{G}][\mathbf{H}], \quad (\text{C.4a})$$

$$\frac{d[(-)\text{-(HG)}]}{dt} = -(\kappa_{\text{taut}} + \kappa_{\text{taut}}^*)[\mathbf{G}][(-)\text{-(HG)}] - \kappa_{\text{diss}}[(-)\text{-(HG)}] + (\kappa_{\text{taut}} + \kappa_{\text{taut}}^*)[\mathbf{G}][(+)\text{-(HG)}] + 2\kappa_{\text{ass}}[\mathbf{G}][\mathbf{H}], \quad (\text{C.4b})$$

$$\frac{d[\mathbf{H}]}{dt} = -4\kappa_{\text{ass}}[\mathbf{G}][\mathbf{H}] - \kappa_{\text{CD}}[\mathbf{W}][\mathbf{H}] + \kappa_{\text{diss}}([(+)\text{-(HG)}] + [(-)\text{-(HG)}]) + \kappa_{\text{DC}}[\mathbf{HW}], \quad (\text{C.4c})$$

$$\frac{d[\mathbf{HW}]}{dt} = -\kappa_{\text{DC}}[\mathbf{HW}] + \kappa_{\text{CD}}[\mathbf{W}][\mathbf{H}]. \quad (\text{C.4d})$$

These equations are compatible with the simplified chemical kinetics (Eq. (C.1)) by adding Eq. (C.4a) and (C.4b), and substituting $[(+)\text{-(HG)}] + [(-)\text{-(HG)}] = [\mathbf{HG}]$.

In order to establish the connection between chemical kinetics and spin kinetics, the scheme in Fig. C.10b must be expanded to incorporate all relevant microstates, see Fig. C.10c. The states $(+)\text{-(HG)}^{A_1}$, $(+)\text{-(HG)}^{A_3}$, $(-)\text{-(HG)}^{A_2}$ and $(-)\text{-(HG)}^{A_4}$ correspond to protonation of different C=O sites (see Fig. 7.10 for molecular structures). The expanded chemical kinetics is described by the following equations

$$\begin{aligned} \frac{d[(+)\text{-(HG)}^{A_1}]}{dt} &= -(2\kappa_{\text{taut}} + \kappa_{\text{taut}}^*)[\mathbf{G}][(+)\text{-(HG)}^{A_1}] - \kappa_{\text{diss}}[(+)\text{-(HG)}^{A_1}] \\ &\quad + \kappa_{\text{taut}}[\mathbf{G}][(+)\text{-(HG)}^{A_3}] + \kappa_{\text{taut}}^*[\mathbf{G}][(-)\text{-(HG)}^{A_2}] + \kappa_{\text{taut}}[\mathbf{G}][(-)\text{-(HG)}^{A_4}] + \kappa_{\text{ass}}[\mathbf{G}][\mathbf{H}^{\text{C}}], \end{aligned} \quad (\text{C.5a})$$

$$\begin{aligned} \frac{d[(+)\text{-(HG)}^{A_3}]}{dt} &= -(2\kappa_{\text{taut}} + \kappa_{\text{taut}}^*)[\mathbf{G}][(+)\text{-(HG)}^{A_3}] - \kappa_{\text{diss}}[(+)\text{-(HG)}^{A_3}] \\ &\quad + \kappa_{\text{taut}}[\mathbf{G}][(+)\text{-(HG)}^{A_1}] + \kappa_{\text{taut}}[\mathbf{G}][(-)\text{-(HG)}^{A_2}] + \kappa_{\text{taut}}^*[\mathbf{G}][(-)\text{-(HG)}^{A_4}] + \kappa_{\text{ass}}[\mathbf{G}][\mathbf{H}^{\text{C}}], \end{aligned} \quad (\text{C.5b})$$

$$\begin{aligned} \frac{d[(-)\text{-(HG)}^{A_2}]}{dt} &= -(2\kappa_{\text{taut}} + \kappa_{\text{taut}}^*)[\mathbf{G}][(-)\text{-(HG)}^{A_2}] - \kappa_{\text{diss}}[(-)\text{-(HG)}^{A_2}] \\ &\quad + \kappa_{\text{taut}}^*[\mathbf{G}][(+)\text{-(HG)}^{A_1}] + \kappa_{\text{taut}}[\mathbf{G}][(+)\text{-(HG)}^{A_3}] + \kappa_{\text{taut}}[\mathbf{G}][(-)\text{-(HG)}^{A_4}] + \kappa_{\text{ass}}[\mathbf{G}][\mathbf{H}^{\text{C}}], \end{aligned} \quad (\text{C.5c})$$

$$\begin{aligned} \frac{d[(-)\text{-(HG)}^{A_4}]}{dt} &= -(2\kappa_{\text{taut}} + \kappa_{\text{taut}}^*)[\mathbf{G}][(-)\text{-(HG)}^{A_4}] - \kappa_{\text{diss}}[(-)\text{-(HG)}^{A_4}] \\ &\quad + \kappa_{\text{taut}}[\mathbf{G}][(+)\text{-(HG)}^{A_1}] + \kappa_{\text{taut}}^*[\mathbf{G}][(+)\text{-(HG)}^{A_3}] + \kappa_{\text{taut}}[\mathbf{G}][(-)\text{-(HG)}^{A_2}] + \kappa_{\text{ass}}[\mathbf{G}][\mathbf{H}^{\text{C}}], \end{aligned} \quad (\text{C.5d})$$

$$\begin{aligned} \frac{d[\mathbf{H}^{\text{C}}]}{dt} &= -4\kappa_{\text{ass}}[\mathbf{G}][\mathbf{H}^{\text{C}}] - \kappa_{\text{CD}}[\mathbf{W}][\mathbf{H}^{\text{C}}] \\ &\quad + \kappa_{\text{diss}}([(+)\text{-(HG)}^{A_1}] + [(+)\text{-(HG)}^{A_3}] + [(-)\text{-(HG)}^{A_2}] + [(-)\text{-(HG)}^{A_4}]) + \kappa_{\text{DC}}[\mathbf{HW}^{\text{D}}], \end{aligned} \quad (\text{C.5e})$$

$$\frac{d[\mathbf{HW}^{\text{D}}]}{dt} = -\kappa_{\text{DC}}[\mathbf{HW}^{\text{D}}] + \kappa_{\text{CD}}[\mathbf{W}][\mathbf{H}^{\text{C}}]. \quad (\text{C.5f})$$

The above equations are compatible with the chemical kinetics scheme governed by Eq. (C.4) as shown by summing Eq. (C.5a) with (C.5b) and Eq. (C.5c) with (C.5d), and substituting $[(+)\text{-(HG)}^{A_1}] + [(+)\text{-(HG)}^{A_3}] = [(+)\text{-(HG)}]$ and $[(-)\text{-(HG)}^{A_2}] + [(-)\text{-(HG)}^{A_4}] = [(-)\text{-(HG)}]$.

This expanded chemical kinetics scheme accounts for all relevant molecular processes and microstates (i.e. there is absence of integer prefactors or sums of reaction rate coefficients accounting for degeneracy) and explains all the prefactors or sums of reaction rate coefficients used in previous schemes in Fig. C.10a,b. The corresponding energy barriers can be calculated employing the Eyring equation (Eq. 2.20). In order to simplify the model, the guest-mediated prototropic tautomerization between states $(+)\text{-(HG)}^{A_1} \leftrightarrow (+)\text{-(HG)}^{A_3}$, $(+)\text{-(HG)}^{A_1} \leftrightarrow (-)\text{-(HG)}^{A_4}$, $(-)\text{-(HG)}^{A_2} \leftrightarrow (+)\text{-(HG)}^{A_3}$ and $(-)\text{-(HG)}^{A_2} \leftrightarrow (-)\text{-(HG)}^{A_4}$ is characterized by single effective reaction rate coefficient κ_A , although the real energy barriers might differ. On the other hand, prototropic

tautomerization between states (+)-(HG)^{A1} ↔ (-)-(HG)^{A2} and (+)-(HG)^{A3} ↔ (-)-(HG)^{A4} is characterized by the reaction rate coefficient κ_A^* .

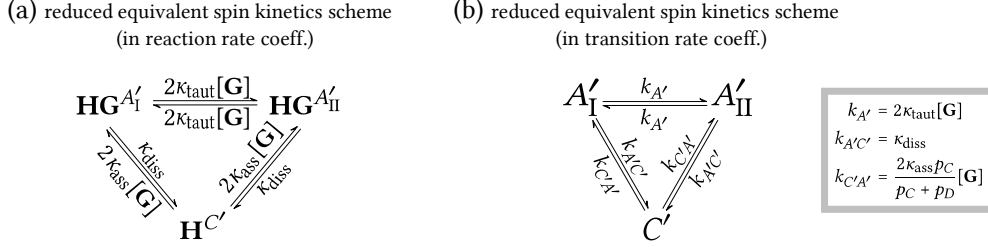


Figure C.11: Reduced equivalent spin kinetics schemes for the multi-state system of di-bromobenzylated oxoporphyrinogen (host **H**) in the presence of two potential ligands, (*R*)-camphorsulfonic acid (ligand **G**) and water (ligand **W**). The schemes refer to central NH protons of the host molecule. (a) Spin kinetics in terms of reaction rate coefficients as obtained from contraction of the scheme in Fig. C.10c. It is described by Eq. (C.6). (b) Corresponding spin kinetics in terms of transition rate coefficients described by Eq. (C.7). It has the form of half-symmetric three-state exchange. Comparison with (a) gives the relationship between transition and reaction rate coefficients in Eq. (C.8). This model was used for lineshape fitting, see Sec. C.7 for details.

At this point, fast exchange between states (+)-(HG)^{A1} ↔ (-)-(HG)^{A2}, (+)-(HG)^{A3} ↔ (-)-(HG)^{A4} and **H**^C ↔ **H**^D should be taken into account. The equivalent reduced spin kinetics scheme is constructed from the expanded chemical kinetics scheme in Fig. C.10c according to the procedure explained in Sec. 3.3.8. In this case, Eq. (C.5a) and (C.5c), Eq. (C.5b) and (C.5d), and Eq. (C.5e) and (C.5f), respectively, are added up, expansion $[\text{H}^C] = \frac{[\text{H}^C]}{[\text{H}^C] + [\text{H}^D]}([\text{H}^C] + [\text{H}^D])$ is made and substitutions $[(+)\text{-(HG)}^{A1}] + [(-)\text{-(HG)}^{A2}] = [\text{HG}^{A'1}]$, $[(+)\text{-(HG)}^{A3}] + [(-)\text{-(HG)}^{A4}] = [\text{HG}^{A'2}]$ and $[\text{H}^C] + [\text{H}^D] = [\text{H}^C]$ are used. This procedure yields governing equations for three-state spin kinetics corresponding to the spin kinetics scheme in Fig. C.11a, in particular

$$\frac{d[\text{HG}^{A'1}]}{dt} = -2\kappa_{\text{taut}}[\text{G}][\text{HG}^{A'1}] - \kappa_{\text{diss}}[\text{HG}^{A'1}] + 2\kappa_{\text{taut}}[\text{G}][\text{HG}^{A'2}] + 2\kappa_{\text{ass}}[\text{G}][\text{H}^C], \quad (\text{C.6a})$$

$$\frac{d[\text{HG}^{A'2}]}{dt} = -2\kappa_{\text{taut}}[\text{G}][\text{HG}^{A'2}] - \kappa_{\text{diss}}[\text{HG}^{A'2}] + 2\kappa_{\text{taut}}[\text{G}][\text{HG}^{A'1}] + 2\kappa_{\text{ass}}[\text{G}][\text{H}^C], \quad (\text{C.6b})$$

$$\frac{d[\text{H}^C]}{dt} = -4\kappa_{\text{ass}} \frac{[\text{G}][\text{H}^C]}{[\text{H}^C] + [\text{H}^D]}[\text{H}^C] + \kappa_{\text{diss}}([\text{HG}^{A'1}] + [\text{HG}^{A'2}]). \quad (\text{C.6c})$$

These equations can be further compared with a spin kinetics scheme formulated in terms of transition rate coefficients as shown in Fig. C.11b, which is governed by the following set of differential equations

$$\frac{dp_{A'1}}{dt} = -(k_{A'} + k_{A'C'})p_{A'1} + k_{A'}p_{A'2} + k_{C'A'}p_{C'}, \quad (\text{C.7a})$$

$$\frac{dp_{A'2}}{dt} = -(k_{A'} + k_{A'C'})p_{A'2} + k_{A'}p_{A'1} + k_{C'A'}p_{C'}, \quad (\text{C.7b})$$

$$\frac{dp_{C'}}{dt} = -2k_{C'A'}p_{C'} + k_{A'C'}p_{A'1} + k_{A'C'}p_{A'2}. \quad (\text{C.7c})$$

Relationships between the transition and reaction rate coefficients are established by comparison of Eq. (C.6) and (C.7) using the appropriate relations for spin state populations, i.e. $p_{A'1} = \frac{[\text{HG}^{A'1}]}{[\text{H}]_t}$, $p_{A'2} = \frac{[\text{HG}^{A'2}]}{[\text{H}]_t}$ and $p_{C'} = \frac{[\text{H}^C]}{[\text{H}]_t}$. As a result, we also obtain the dependencies of transition rate

coefficients on guest concentration,

$$k_{A'} = 2\kappa_A[\mathbf{G}], \quad (\text{C.8a})$$

$$k_{A'C'} = \kappa_{\text{diss}}, \quad (\text{C.8b})$$

$$k_{C'A'} = \frac{2\kappa_{\text{ass}}p_C}{p_C + p_D}[\mathbf{G}]. \quad (\text{C.8c})$$

C.9 Prototropic tautomerization processes in **Bz₂OxP**

Prototropic tautomerization consists of several steps, which are indistinguishable in NMR (due to fast exchange regime). Here, we suggest the corresponding reaction schemes for tautomerization processes. Initial state (-)-(HG) can undergo the tautomerization process in three different ways as illustrated in Fig. C.12 (final state (+)-(HG), characterized by κ_A^*), Fig. C.13 (final state (-)-(HG), characterized by κ_A) and Fig. C.14 (final state (+)-(HG), characterized by κ_A). All three reaction paths are reversible and consist of analogous steps. In step (1), an incoming guest (red) forms a hydrogen bond with host's carbonyl group thus forming hydrogen-bonded structure (-)-HG·G. The next step (2) shows the suggested transition state (HG₂)[‡], which is created by formal redistribution of electron density (denoted by blue arrows in (-)-HG·G structure and by dashed bonds in (HG₂)[‡] structure). Formally, partial charge + δ is located at the central alkylated amines and at the protonated carbonyls although in reality, the charge is delocalized. During this step, guest anions change positions and are located close to partially charged atoms of the host in (HG₂)[‡]. This allows for (blue and red) guest anions exchange, as shown in step (3). Hence, prototropic tautomerization can, in principle, proceed in two branches, i.e. the initial guest counteranion (blue) either remains at the porphyrinogen center or it is replaced by the incoming acid anion (red). Both branches lead to hydrogen bonded species (+)-HG·G or (-)-HG·G in step (4) (redistribution of electron density is denoted by blue arrows in (HG₂)[‡] structure). Finally, monoprotonated host-guest complex (+)-(HG) or (-)-(HG) is formed in step (5). Although the tautomerization processes consist of several steps, in our chemical kinetics scheme (Fig. 7.13c), they are described by effective barriers ΔG_A^\ddagger (corresponding to κ_A) and $\Delta G_A^{\ddagger*}$ (corresponding to κ_A^*). Step (1) is a second-order reaction, hence the effective forward reaction rate ($\kappa_A[\mathbf{G}][\mathbf{HG}]$ or $\kappa_A^*[\mathbf{G}][\mathbf{HG}]$) depends linearly on $[\mathbf{G}]$ and $[\mathbf{HG}]$. Similarly, step (5) is also a second-order reaction, hence the effective backward reaction rate ($\kappa_A[\mathbf{G}][\mathbf{HG}]$ or $\kappa_A^*[\mathbf{G}][\mathbf{HG}]$) depends linearly on $[\mathbf{G}]$ and $[\mathbf{HG}]$. Note that we assume equal effective barriers ΔG_A^\ddagger for protonation transfers “left-to-right” (Fig. C.13) and “left-to-down” and (Fig. C.14).

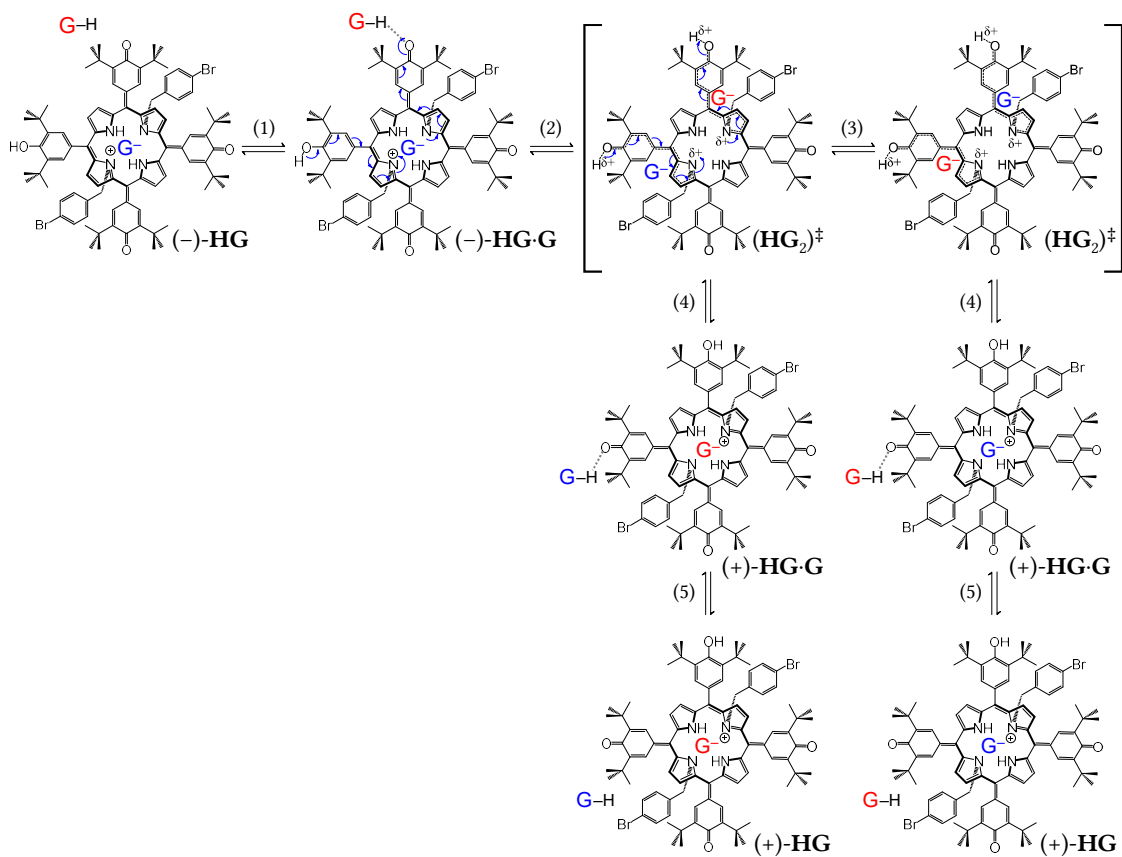


Figure C.12: Suggested transition structures for prototropic tautomerization from (-)-(HG) to (+)-(HG) characterized by κ_A^* . Dotted line in (+)-HG·G and (-)-HG·G represents hydrogen bond, dashed lines in $(HG_2)^\ddagger$ represent partially formed bonds in transition state.

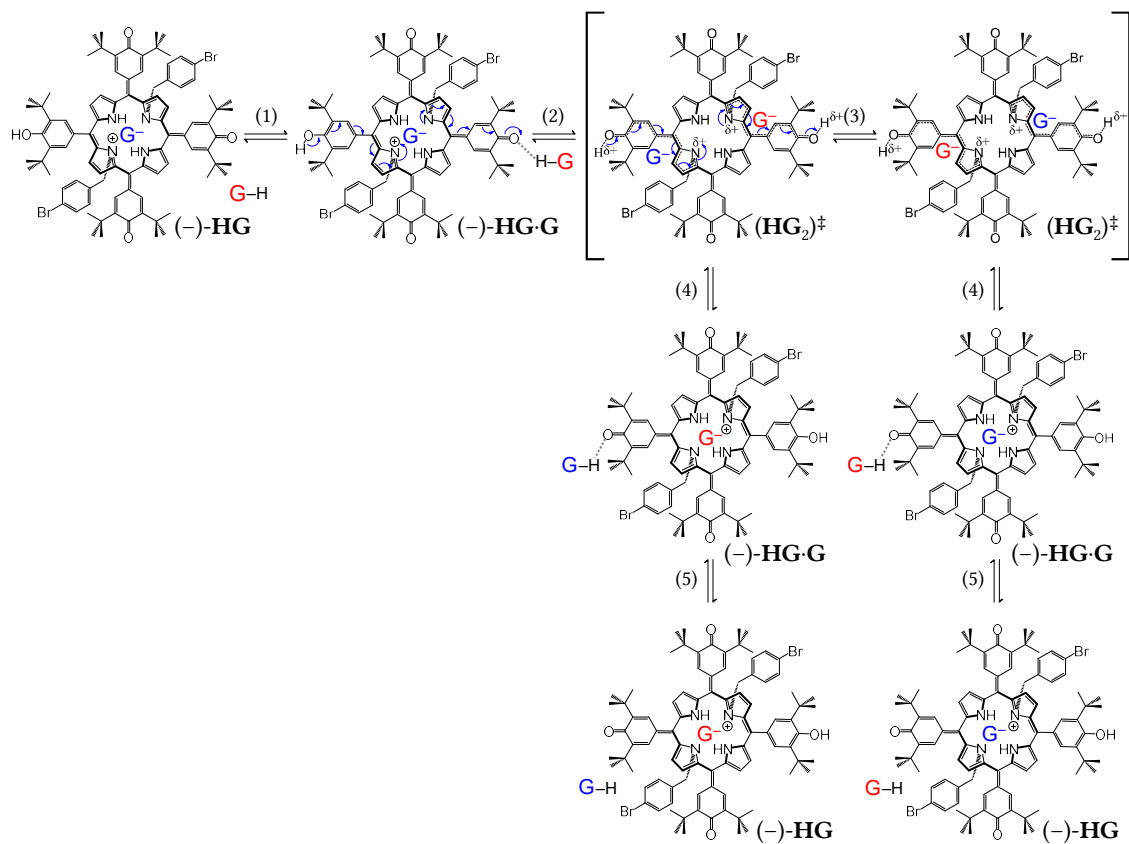


Figure C.13: Suggested transition structures for prototropic tautomerization from $(-)\text{-HG}$ to $(-)\text{-HG-G}$ characterized by κ_A . Dotted line in $(-)\text{-HG-G}$ represents hydrogen bond, dashed lines in $(\text{HG}_2)^\ddagger$ represent partially formed bonds in transition state.

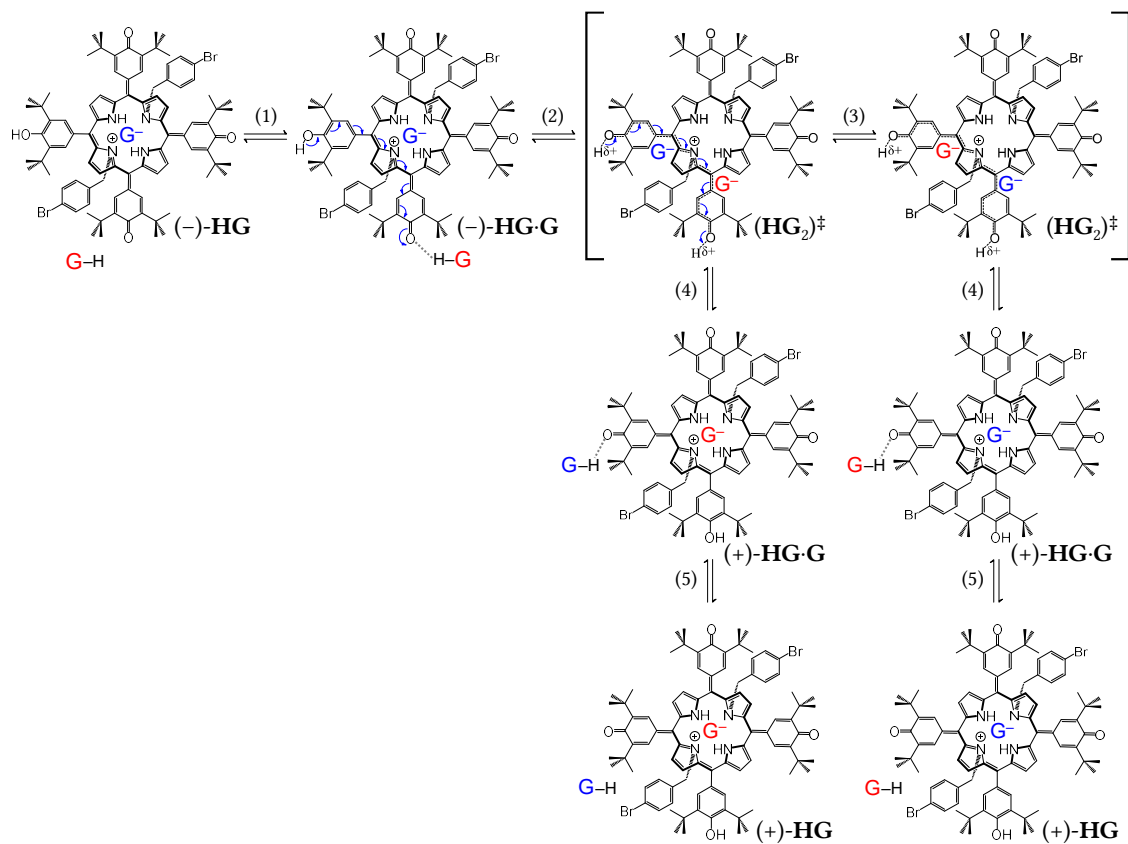


Figure C.14: Suggested transition structures for prototropic tautomerization from (-)-(HG) to (+)-(HG) characterized by κ_A . Dotted line in (+)-HG·G and (-)-HG·G represents hydrogen bond, dashed lines in (HG₂)[‡] represent partially formed bonds in transition state.

D SVD analysis

SVD of UV/vis spectra (in the matrix \mathbf{A}) was conducted in *Mathematica* using:

```
{U, S, V} = SingularValueDecomposition[A];
```

Graphs showing quality of DSIS and SVD analysis were adopted from our publication [2].

D.1 Bz₂OxP + (R)-CSA titration

Total host concentration changed during the titration due to dilution by stock solution (see Appendix C.2). Therefore, the measured spectra shown in Fig. D.1a were rescaled to the initial

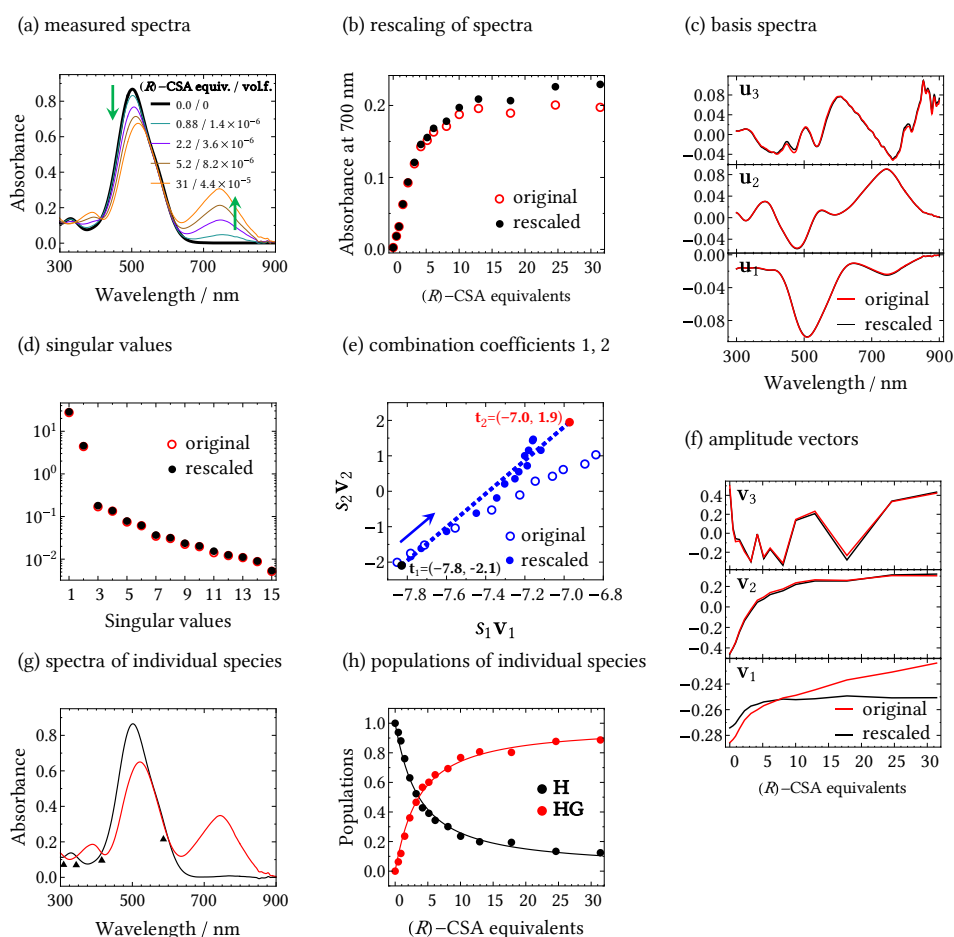


Figure D.1: UV/vis titration of **Bz₂OxP** (7×10^{-6} M, dilution during measurement compensated by rescaling to initial dye concentration, CDCl₃) with (R)-CSA and corresponding analyses. (a) Rescaled titration spectra. (b) Absorbance at 700 nm before and after rescaling. SVD analysis before and after rescaling yielded (c) basis spectra, (d) singular values, (e) mutual plot of combination coefficients 1,2 with coordinates \mathbf{t}_1 and \mathbf{t}_2 of the absorbing species, and (f) amplitude vectors. The fitting procedure, described in Sec. 5.2.3, yielded (g) spectra of free (identified as bold line in panel (a)) and protonated host (obtained from SVD analysis), black triangles denote intersections which produce isosbestic points as explained in Sec. 4.4; and (h) populations \mathbf{P}_{exp} (points) and \mathbf{P}_{mod} (lines) of these species.

concentration (see footnote 7 on page 58), plot of the absorbance at 700 nm in Fig. D.1b illustrates the extent of rescaling. SVD analysis following Sec. 5.2.3 was performed on the original as well as rescaled spectra. Figures D.1c–f prove little effect of spectra rescaling on SVD except for the first amplitude vector \mathbf{v}_1 (note that \mathbf{v}_1 is most sensitive to concentration fluctuations also in other UV/vis titration experiments) and consequently the first combination coefficient $s_1\mathbf{v}_1$. Mutual plot of combination coefficients (e) forms a straight line segment confirming the presence of two absorbing species – free and protonated $\mathbf{Bz}_2\mathbf{OxP}$ host. Spectrum of free host \mathbf{z}_1 is identified as the first titration measurement \mathbf{a}_1 (black bold line in (a)) yielding the \mathbf{t}_1 coordinates in combination coefficients space (black point in (e)) using $\mathbf{t}_1 = \mathbf{U}_2^\dagger \mathbf{a}_1$ (from Eq. (5.21) and (5.14)). However, spectrum of protonated host \mathbf{z}_2 was not measured in isolation (the spectral change was not saturated) and its \mathbf{t}_2 coordinates were obtained by minimization of the function in Eq. (5.26) using the 1:1 host-guest binding model (see Sec. 2.2.1) for the model populations \mathbf{P}_{mod} . Fitting parameters were the vector \mathbf{t}_2 (i.e. T_{12} and T_{22} elements of the transformation matrix \mathbf{T}) and K_1 . Resulting value is $K_{\text{HG}} = (5.0 \pm 1.0) \times 10^4 \text{ M}^{-1}$, uncertainties were obtained by repeated fitting with fixed K_1 value.

D.2 OxP + TFA titration

The presence of three absorbing components and two consecutive spectral changes is confirmed by singular values in Fig. D.2a and mutual plot of combination coefficients in Fig. D.2c–f showing

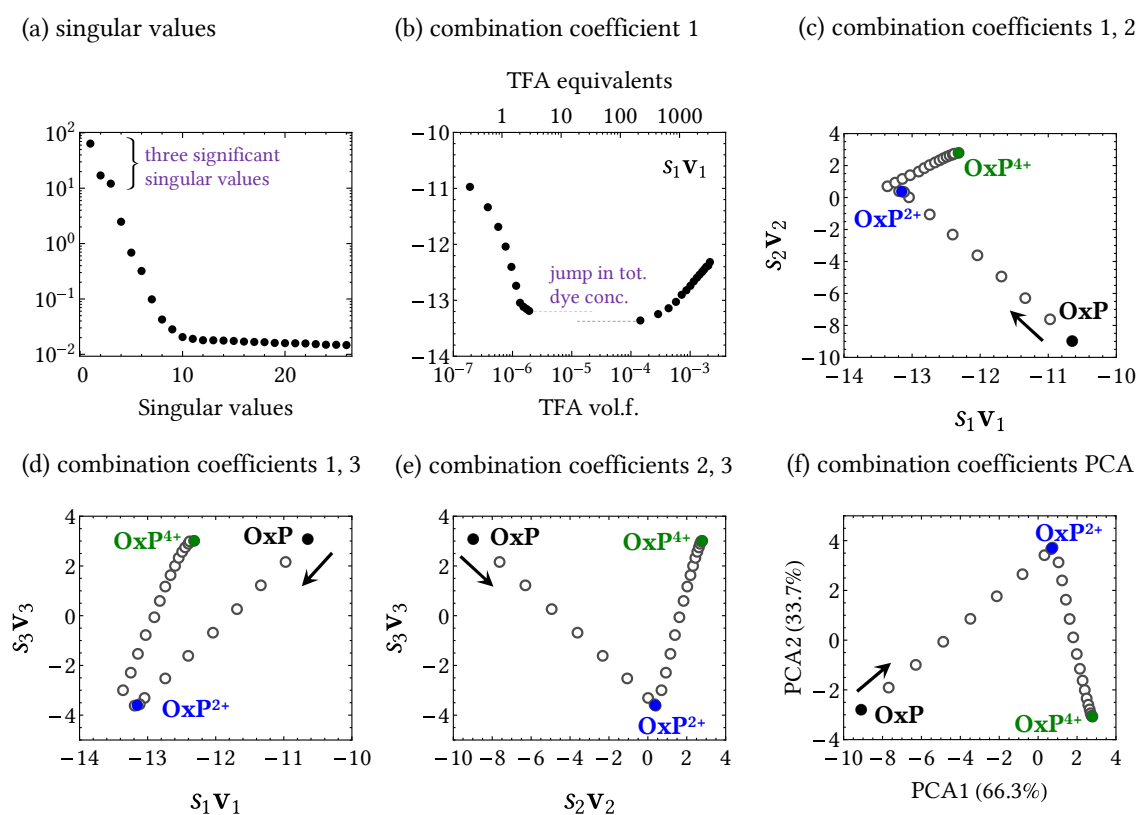


Figure D.2: SVD analysis of UV/vis titration of \mathbf{OxP} (10^{-5} M , CH_2Cl_2 , data first published in [31]) with TFA, (a) singular values, (b) first combination coefficients shows ‘jump’ due to total dye concentration change, (c,d,e) mutual plot of combination coefficients. (f) PCA on the combination coefficients enables 2D plot of 3D data lying on a plane.

two straight line segments. Coordinates of **OxP**²⁺ in Fig. D.2c,d do not lie exactly at the intersection of the two lines as would be expected. This is probably because the first combination coefficient $s_1\mathbf{v}_1$ in Fig. D.2b shows a “jump” since it is sensitive to the total dye concentration, which slightly changed (by a few percent as estimated from the ratio of $s_1\mathbf{v}_1$ values before and after the jump) when different stock solution was used to observe the second spectral change.

Since the combination coefficients from three species lie on a 2D plane in 3D combination coefficients space, they can be conveniently visualized when the plane is properly rotated. As shown in Sec. 5.3, the PCA method is suited for such a rotation. The PCA of combination coefficients almost coincides with the mutual plot of the second and third combination coefficients (the plane has been originally oriented almost perpendicularly to the $s_1\mathbf{v}_1$ dimension), where the effect of total dye concentration change is filtrated out (**OxP**²⁺ lies at the intersection of straight lines). See Appendix D.5 for the detailed implementation of PCA.

Populations were obtained by DSIS (see Sec. 5.2.5) by minimization of Eq. 5.28 with constraints $\sum \mathbf{p}_j = 1$ and $0 \leq \mathbf{p}_j \leq 1$. These populations were fitted using the 1:4 host-guest binding model from Appendix A.4. Furthermore, fitting of the host-guest model was performed in the 3D SVD approximation as described in Sec. 5.2.3, see [1] for details. Figure D.3 shows the high quality of these fits.

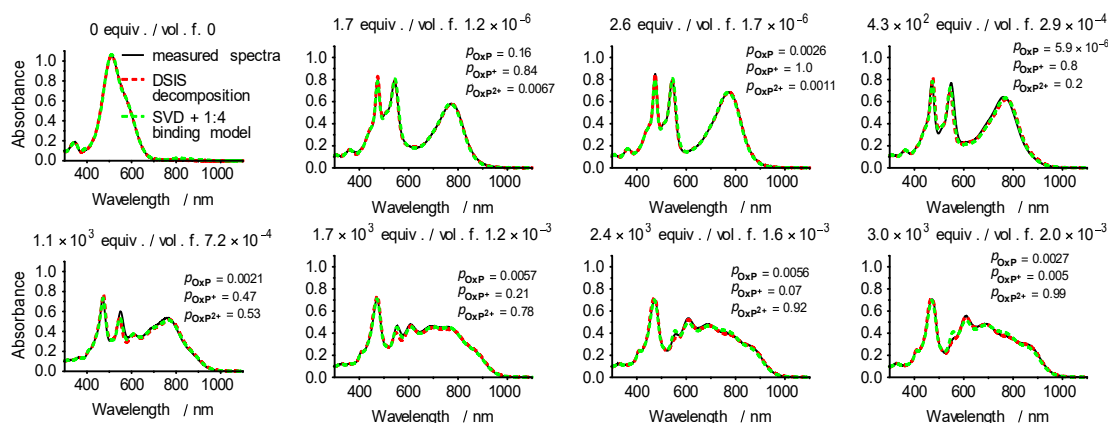


Figure D.3: Quality check of DSIS and 3D SVD approximation of UV/vis titration of **OxP** with TFA.

D.3 Bz_2OxP + DFA titration

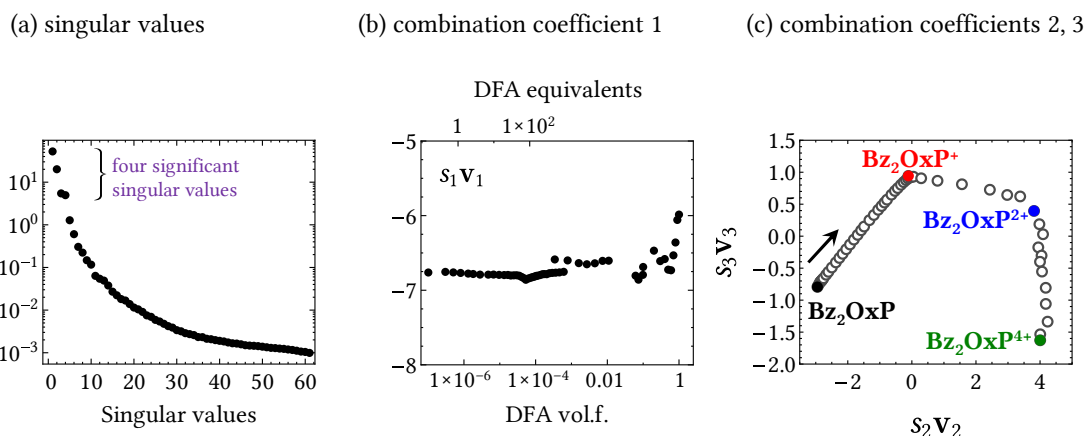


Figure D.4: SVD analysis of UV/vis titration of Bz_2OxP ($7 \times 10^{-6}\text{M}$, CDCl_3 , diluted during measurement, rescaled to initial concentration) with DFA, (a) singular values, (b) first combination coefficients shows “jumps” due to total dye concentration change, (c) mutual plot of combination coefficients 2,3.

The presence of four absorbing components and three consecutive spectral changes is confirmed by singular values in Fig. D.2a and mutual plot of combination coefficients in Fig. D.2c showing three line segments. Because of experimental difficulties (unwanted photoreaction at high acid contents) the titration was repeated several times at different concentration ranges. Finally, the spectra were combined together and rescaled to the initial dye concentration (see footnote 7 on page 58). Imperfections in rescaling are visible as small “jumps” in the first combination coefficient in Fig. D.2b. Mutual plot of combination coefficients 2 and 3 is not sensitive to these concentration fluctuations.

Populations were obtained by DSIS (see Sec. 5.2.5) by minimization of Eq. 5.28 with con-

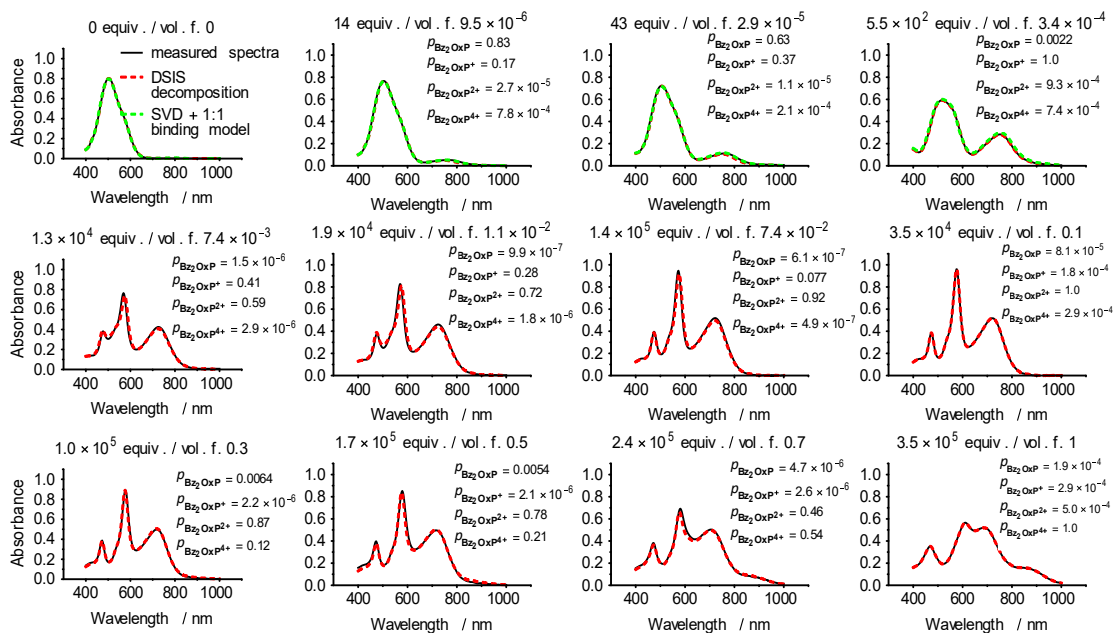


Figure D.5: Quality check of DSIS and 4D SVD approximation of UV/vis titration of Bz_2OxP with DFA.

straints $\sum \mathbf{p}_j = 1$ and $0 \leq \mathbf{p}_j \leq 1$. These populations were fitted using the 1:1 host-guest binding model from Appendix A.1. Furthermore, fitting of the host-guest model was performed in the 4D SVD approximation as described in Sec. 5.2.3, see [1] for details. Figure D.5 shows the high quality of these fits.

D.4 **Bz₄OxP** + DFA titration

Data excluding the first solvatochromic shift were analyzed by SVD. The presence of three absorbing components and two consecutive spectral changes is confirmed by singular values in Fig. D.6a and mutual plot of combination coefficients in Fig. D.6c showing two straight line segments. Rescaling of the experimental spectra produced on jumps (only one outlier) as deduced from the first combination coefficient in Fig. D.6b.

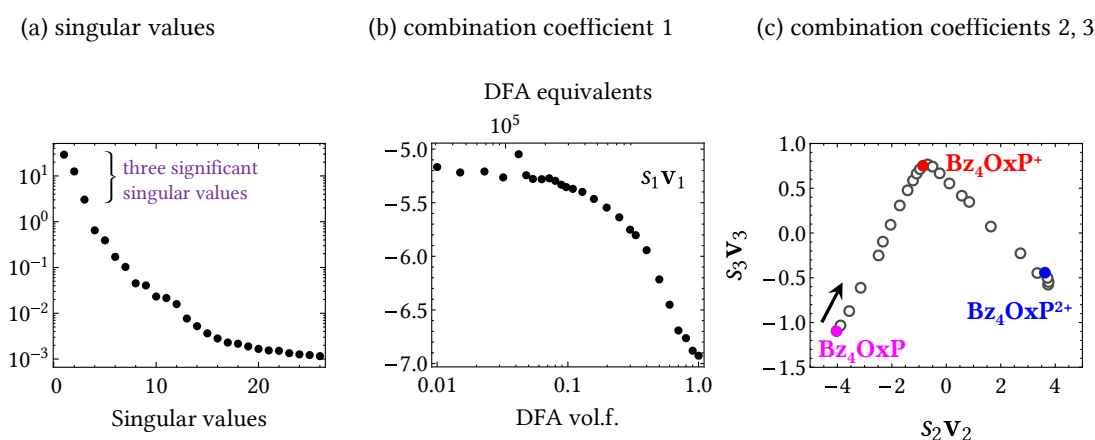


Figure D.6: SVD analysis of UV/vis titration of **Bz₄OxP** (5×10^{-6} M, CDCl₃) with DFA, diluted during measurement, measured spectra rescaled to initial concentration. (a) Singular values, (b) first combination coefficients shows only one outlier, so the rescaling was very good, (c) mutual plot of combination coefficients 2,3.

Populations were obtained by DSIS (see Sec. 5.2.5) by minimization of Eq. 5.28 with constraints $\sum \mathbf{p}_j = 1$ and $0 \leq \mathbf{p}_j \leq 1$. Figure D.7 shows the high quality of this decomposition.

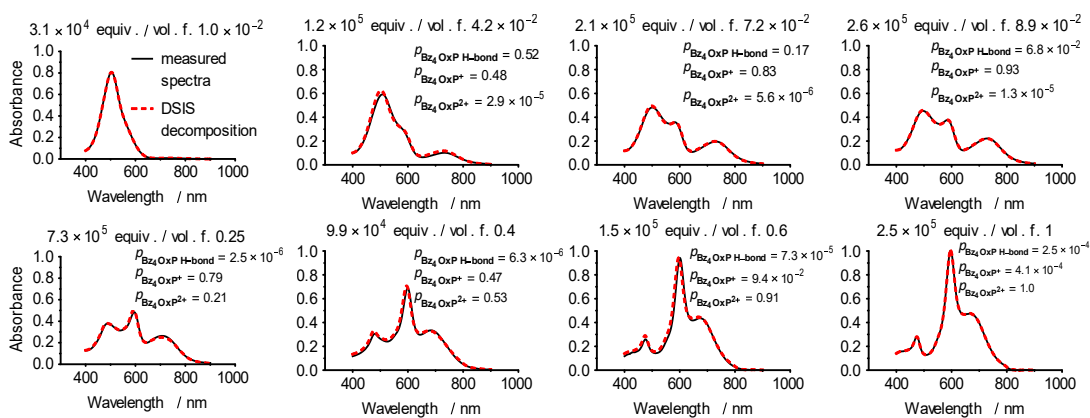


Figure D.7: Quality check of DSIS of UV/vis titration of **Bz₄OxP** with DFA.

D.5 OxP + DMF titration

Plot of singular values in Fig. D.8a indicates three significant components (capturing 99.975% of the signal), which indicates presence of three absorbing species. However, the structure of combination coefficients plot in Fig. D.8b is nontrivial. The **OxP** and **OxP H-bond** species lie on a straight line but before saturation of the spectral change, there is a “digression” to the **porphodimethene form**. In our publication [1], we concluded the existence of two hydrogen-bonded species (very similar to each other and thus consistent with three significant singular values), however, here we propose this simpler solution with only three absorbing species on the basis of the combination coefficients plot (not present in [1]).

Figure D.8b shows PCA of the combination coefficients. Since PCA identifies directions of maximal variance in a dataset (see Sec. 5.3), the first two PCs project the combination coefficients from 3D to 2D plane so that they are maximally “stretched out”. This projection accounts for 99.95% of the signal variance, so the combination coefficients indeed lie on a 2D plane in their 3D space, further supporting the presence of three absorbing components (N_{spc} components form an $(N_{\text{spc}} - 1)$ -dimensional object in the combination coefficients space, although small overlap with other dimensions occurs due to the noise). Coordinates of **OxP** and the **porphodimethene form** in combination coefficients plot were determined from the measured spectra at the end-points of the spectral changes, where they appear in isolation (black and blue points in Fig. D.8b). The coordinates of **OxP H-bond** were determined as intersection of the lines delineating the first and third spectral changes in the PCA plot (green lines in Fig. D.8b) while setting the PC3 coordinate of all three species to zero as shown in Fig. D.8c.

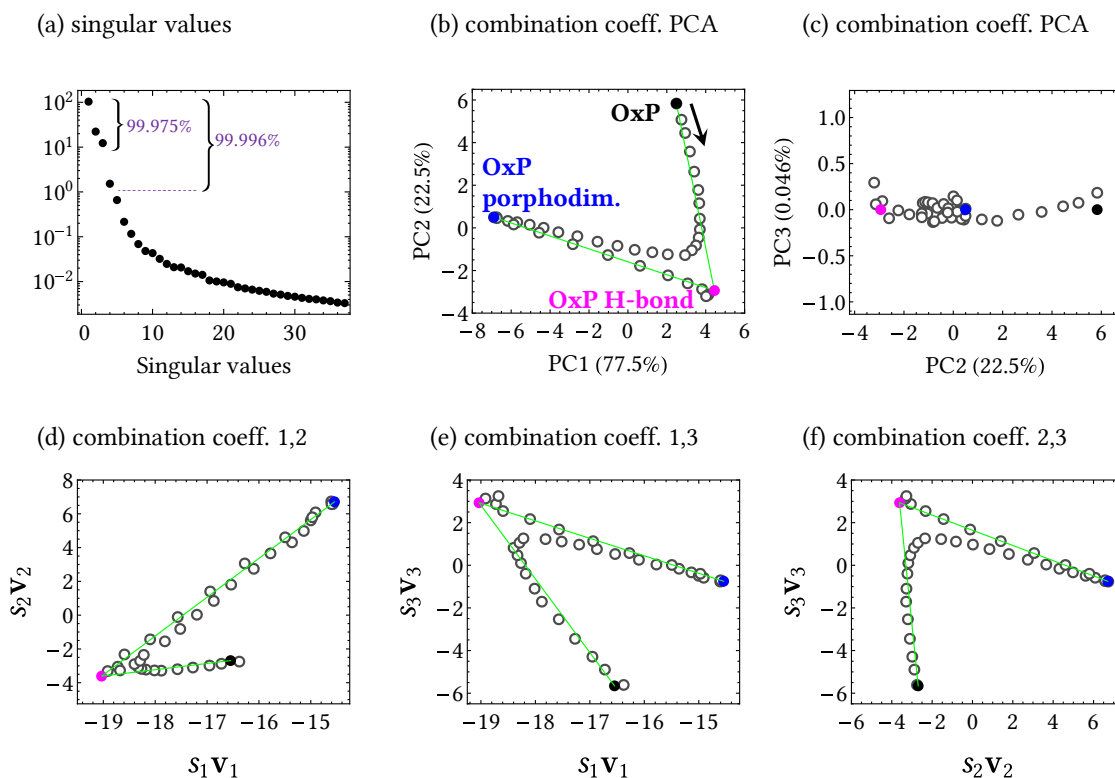


Figure D.8: SVD analysis of UV/vis titration of **OxP** with DMF, (a) singular values show three significant species, (b,c) PCA plot of combination coefficients, (d,e,f) mutual plot of combination coefficients 2, 3. Colored points show coordinates of the absorbing components. Coordinates in PCA3 are set to zero in (c).

For calculation of PCA, the combination coefficients were arranged as columns x_j of the row-wise data matrix \mathbf{B} according to Eq. (5.29). This calculation was implemented in *Mathematica*. Since this software does not include command for direct calculation of PCA, it was obtained using SVD (Sec. 5.3.2) using the following code

```
B = Transpose@{s1 v1, s2 v2, s3 v3 }; (* row-wise data matrix *)
Bbar = B - Outer[Times, ConstantArray[1, Length@B],
Mean[B]]; (* mean-centering *)
{U, S, V} = SingularValueDecomposition[Bbar];
V1 = V[[All, 1]]; V2 = V[[All, 2]]; V3 = V[[All, 3]];
PC1 = Bbar.V1; (* = U1 s1 *)
PC2 = Bbar.V2; (* = U2 s2 *)
PC3 = Bbar.V3; (* = U3 s3 *)
```

Consequently, PCs are transformed to coordinates t_j (using inversion of Eq. 5.31) with

$$\underbrace{\begin{pmatrix} \dots & t_1 & \dots \\ \dots & t_2 & \dots \\ \dots & t_3 & \dots \end{pmatrix}}_{\mathbf{T}^T} = \begin{pmatrix} \vdots & \vdots & \vdots \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \text{PC1(comp.1)} & \text{PC2(comp.1)} & 0 \\ \text{PC1(comp.2)} & \text{PC2(comp.2)} & 0 \\ \text{PC1(comp.3)} & \text{PC2(comp.3)} & 0 \end{pmatrix} + \bar{\mathbf{B}}^T, \quad (\text{D.1})$$

where α_j are PCA loadings. As the data during PCA become mean-centered, the mean $\bar{\mathbf{B}}$ should be added in order to obtain true coordinates. The coordinates obtained through PCA nicely fit into the three original combination coefficients planes, Fig. D.8d–f. It would be very tedious if not impossible to find the proper positions of the **OxP H-bond** coordinate directly in these plots without the use of PCA since bad placement results in negative spectra or forbidden values of populations.

When all coordinates (columns of the transformation matrix \mathbf{T}) are known, the individual spectra are reconstructed using $\mathbf{Z} = \mathbf{U}_{N_{\text{spc}}} \mathbf{T}$ and populations using $\mathbf{P}_{\text{exp}}^T = \mathbf{T}^{-1} \mathbf{S}_{N_{\text{spc}}} \mathbf{V}_{N_{\text{spc}}} = \mathbf{Z}^\dagger \mathbf{A}$. The reconstruction of spectra and populations can be implemented as follows

```
{U, S, V} = SingularValueDecomposition[A];
T = Transpose@{t1, t2, t3};
Z = U[[All, 1;;3]].T;
PT = PseudoInverse[Z].A;
(* alternatively: PT=Inverse[T].(S.Transpose[V])[[1;;3,All]];*)
```

D.6 Bz₂OxP + DMF at different vol.f.

Singular values in Fig. D.9a indicate the presence maximum three absorbing components (99.99% of the signal), however, interpretation of mutual plot of combination coefficients in Fig. D.9b requires three components and two straight line segments. The analysis was conducted using PCA in similar manner as in the previous section. Coordinates of **Bz₂OxP** in (b) were obtained directly from the measured spectra. The second spectral change is unsaturated, thus, only interval for reasonable location for PCA coordinates of **form 2** (transparent green line segment) was determined, it lies on the straight line corresponding to the second spectral change (orange line). PCA coordinates of **form 1** lie at the intersection of the two spectral changes (orange lines in Fig. D.9b). PC3 component of the coordinates was set to zero and the coordinates in Fig. D.9d–f were calculated. Then individual spectra and their populations were obtained analogously to the previous appendix.

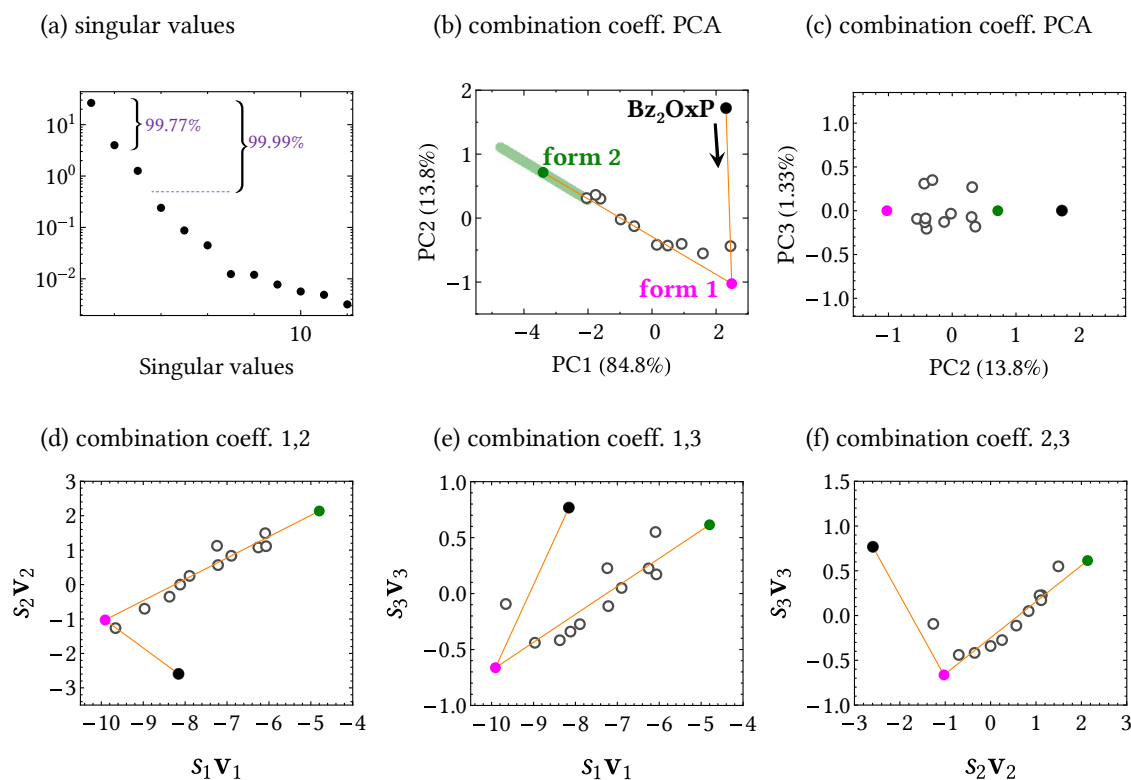


Figure D.9: SVD analysis of UV/vis titration of **Bz₂OxP** with DMF, (a) singular values show three significant species, (b,c) PCA plot of combination coefficients, (d,e,f) mutual plot of combination coefficients 2, 3. Colored points show coordinates of the absorbing components. Coordinates in PCA3 are set to zero in (c).

D.7 Bz_4OxP + DMF at different vol.f.

Since the measured spectral change is very small, it cannot be decided if the solvatochromism is caused by formation of new species or by shift of absorption bands (see Sec. 4.5). Here, the hypothetical formation of new species is investigated. The SVD analysis is fully analogous to the previous appendix. The hypothetical individual spectra in Fig. D.10g resemble the hydrogen bonded species of Bz_2OxP in Fig. 6.16b. However, no absorption bands arise, which suggests that no new species are formed.

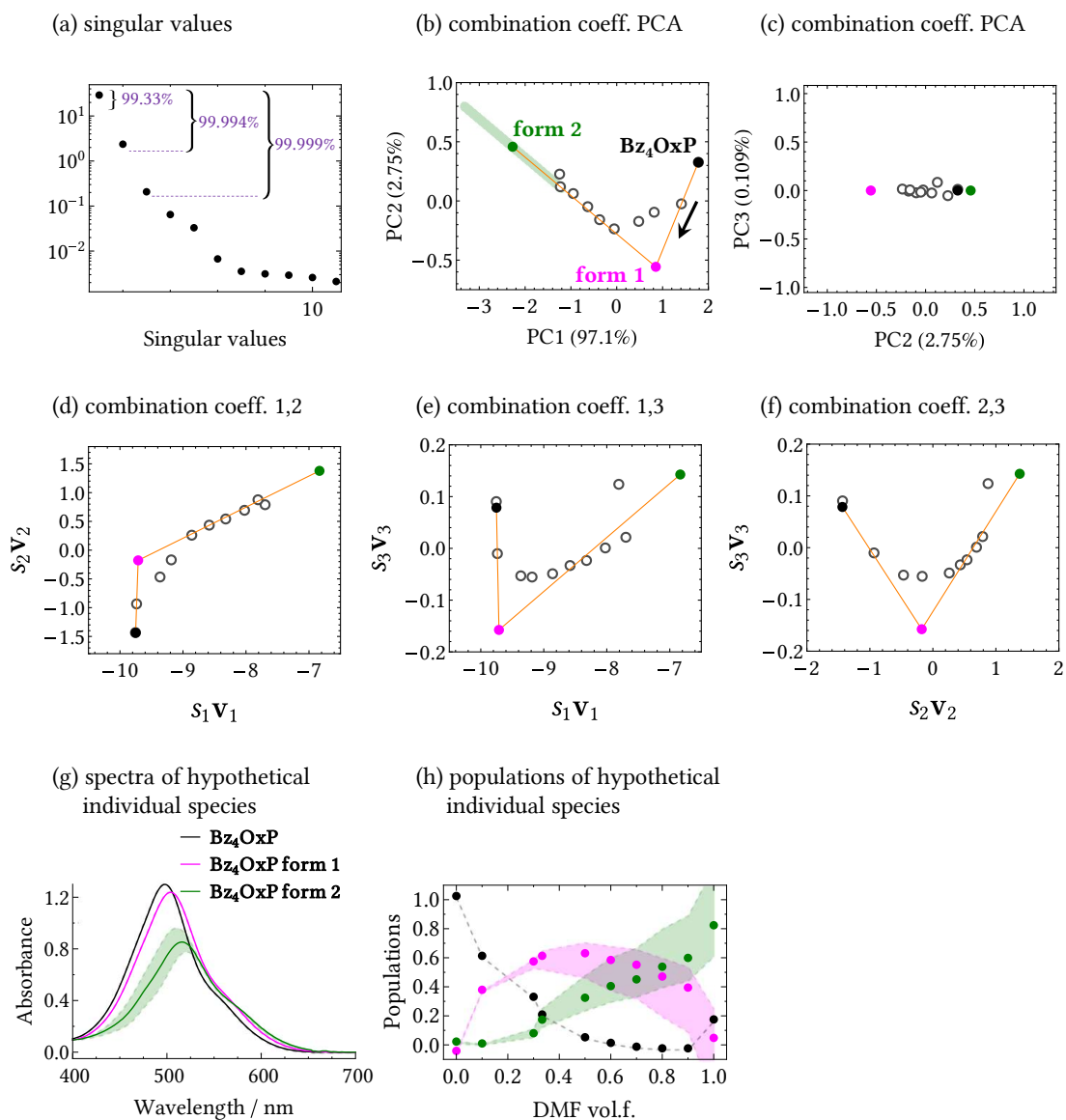


Figure D.10: SVD analysis of UV/vis titration of Bz_4OxP with DMF, (a) singular values, (b,c) PCA plot of combination coefficients, (d,e,f) mutual plot of combination coefficients 2, 3. Colored points show coordinates of the absorbing components. Coordinates in PCA3 are set to zero in (c).

E NMR spectra

NMR spectra analyzed in this work are shown here in the full ppm range. Several figures were adopted from our publication [1] with modifications. Experimental details about NMR measurements can be found in Appendix C.4. Note that NMR assignment of **Bz₂OxP** in our study [1] and in the master's thesis [45] was partially incorrect, the correct assignment can be found here or in [40].

E.1 Spectra of acids

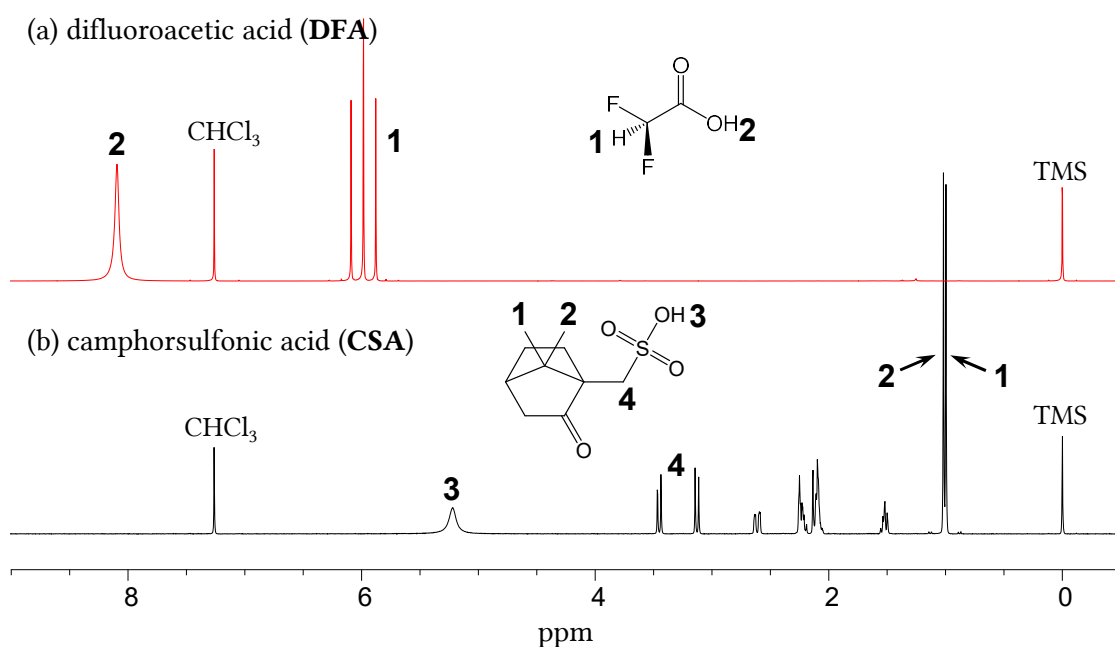


Figure E.1: NMR spectra of (a) difluoroacetic acid and (b) camphorsulfonic acid, both in CDCl_3 at 298 K.

E.2 OxP + DFA titration

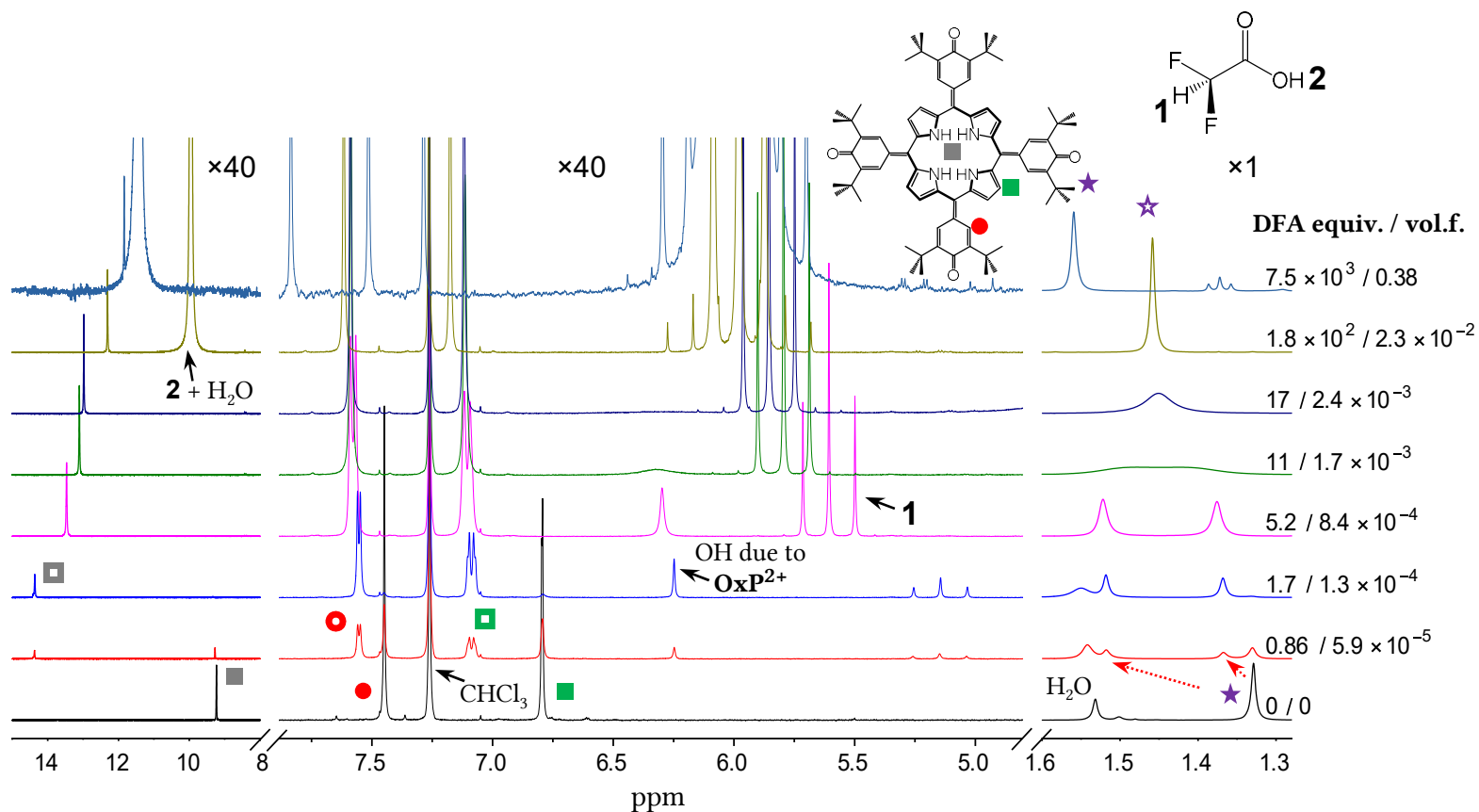


Figure E.2: Titration of OxP with DFA (in CDCl_3 at 298 K, host concentration 0.002 M decreased upon addition of stock solution). Free and protonated/complexed species are denoted by full and empty symbols, respectively. Red dotted arrows denote splitting of *tert*-butyl signal due to symmetry breaking after protonation and anion binding. Some regions of the spectra are magnified for clarity. The factor of magnification is given above each region (see factors above each region).

E.3 OxP + (R)-CSA titration

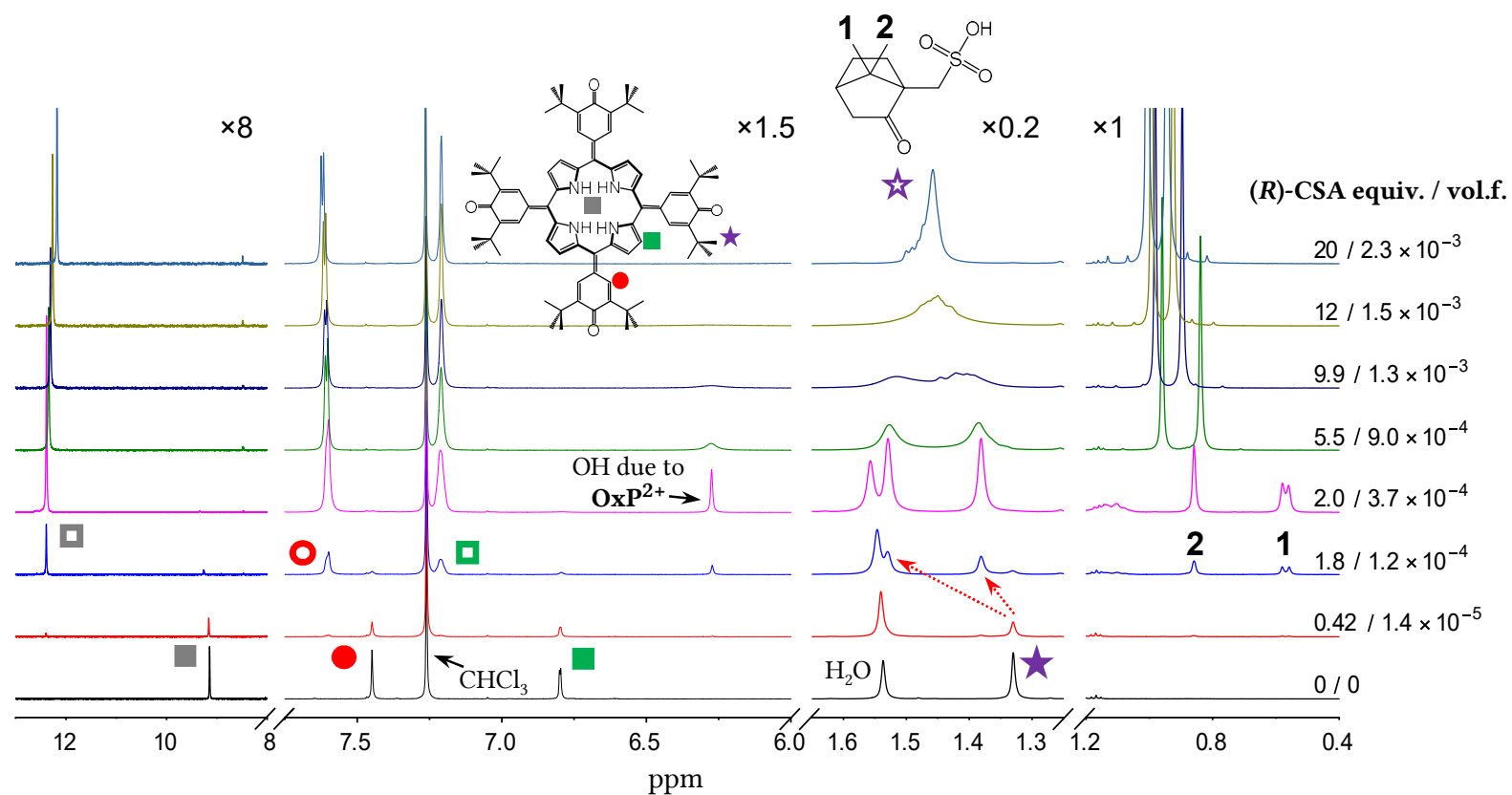


Figure E.3: Titration of **OxP** with (R)-CSA (in CDCl_3 at 298 K, host concentration 9×10^{-4} M decreased upon addition of stock solution). Free and protonated/complexed species are denoted by full and empty symbols, respectively. Red dotted arrows denote splitting of *tert*-butyl signal due to symmetry breaking after protonation and anion binding. Certain regions of spectra have been magnified for clarity (see factors above each region). Splitting of methyl signals due to (R)-CSA at 0.57 ppm below 2 equiv. is caused by formation of (+)-**H(R)-G₂** and (R)-**H(R)-G₂** diastereomers.

E.4 Bz₂OxP + DFA titration

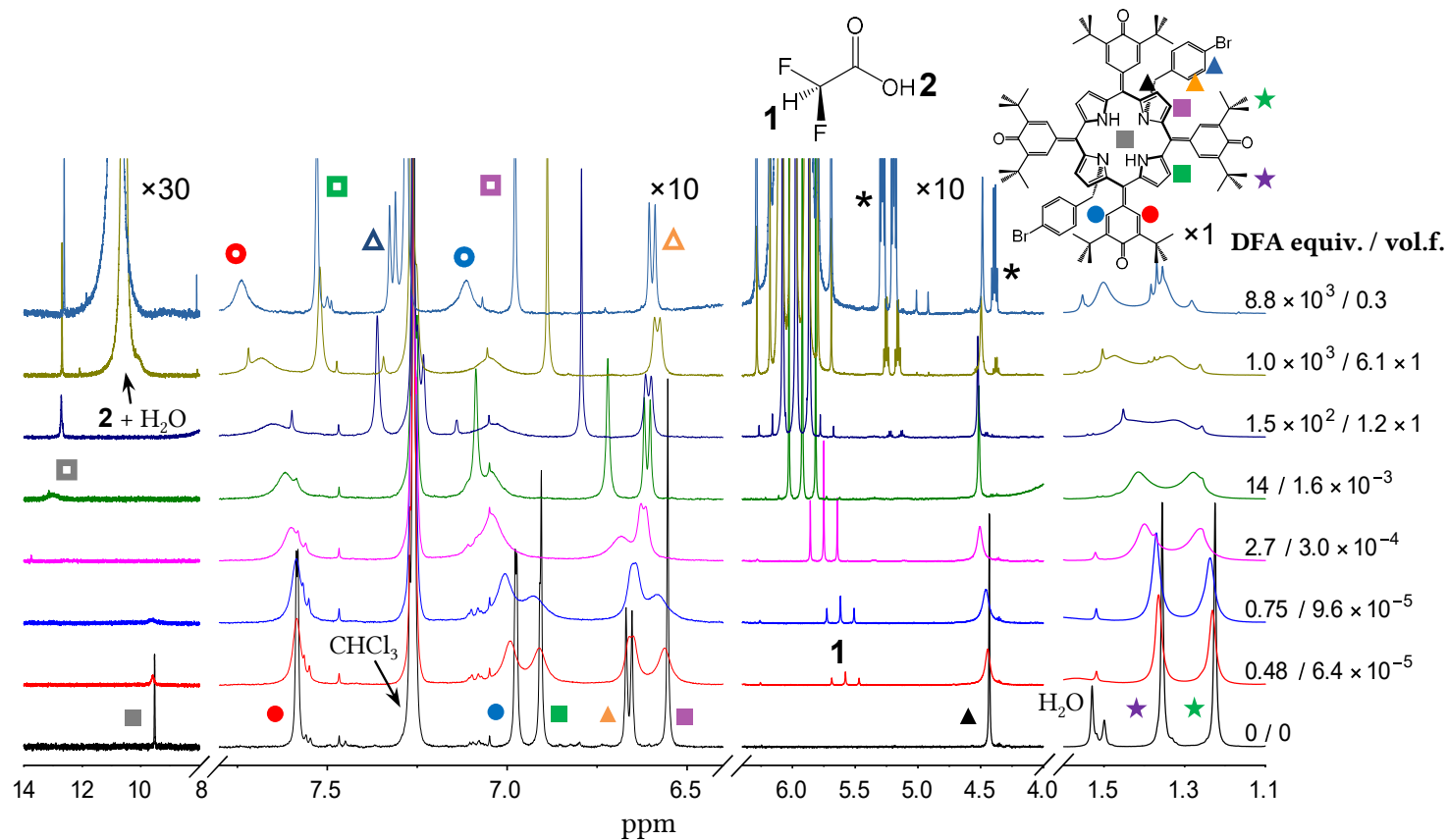


Figure E.4: Titration of **Bz₂OxP** with DFA (in CDCl₃ at 298 K, host concentration 7×10^{-4} M decreased upon addition of stock solution). Free and protonated/complexed species are denoted by full and empty symbols, respectively. Asterisk (*) denotes impurities contained in acid. Some regions of the spectra are magnified for clarity (see factors above each region). The factor of magnification is given above each region.

E.5 Bz₂OxP + (R)-CSA titration

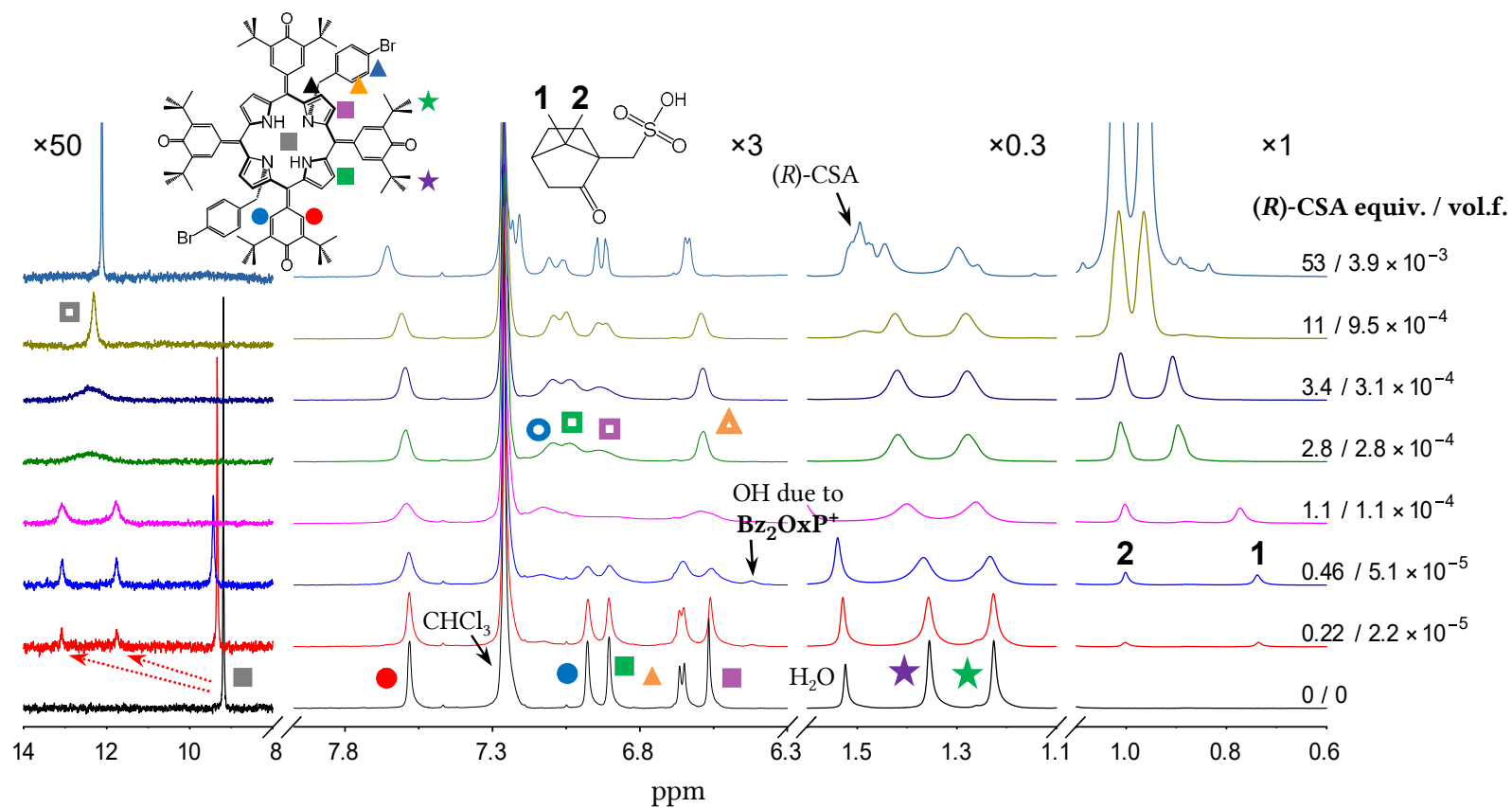


Figure E.5: Titration of **Bz₂OxP** with (R)-CSA (in CDCl₃ at 298 K, host concentration 6×10^{-4} M decreased upon addition of stock solution). Free and protonated/complexed species are denoted by full and empty symbols, respectively. Red dotted arrows denote splitting of *tert*-butyl signal due to symmetry breaking after protonation and anion binding. Certain regions of spectra have been magnified for clarity (see factors above each region).

E.6 Bz₂OxP + (*rac*)-CSA titration

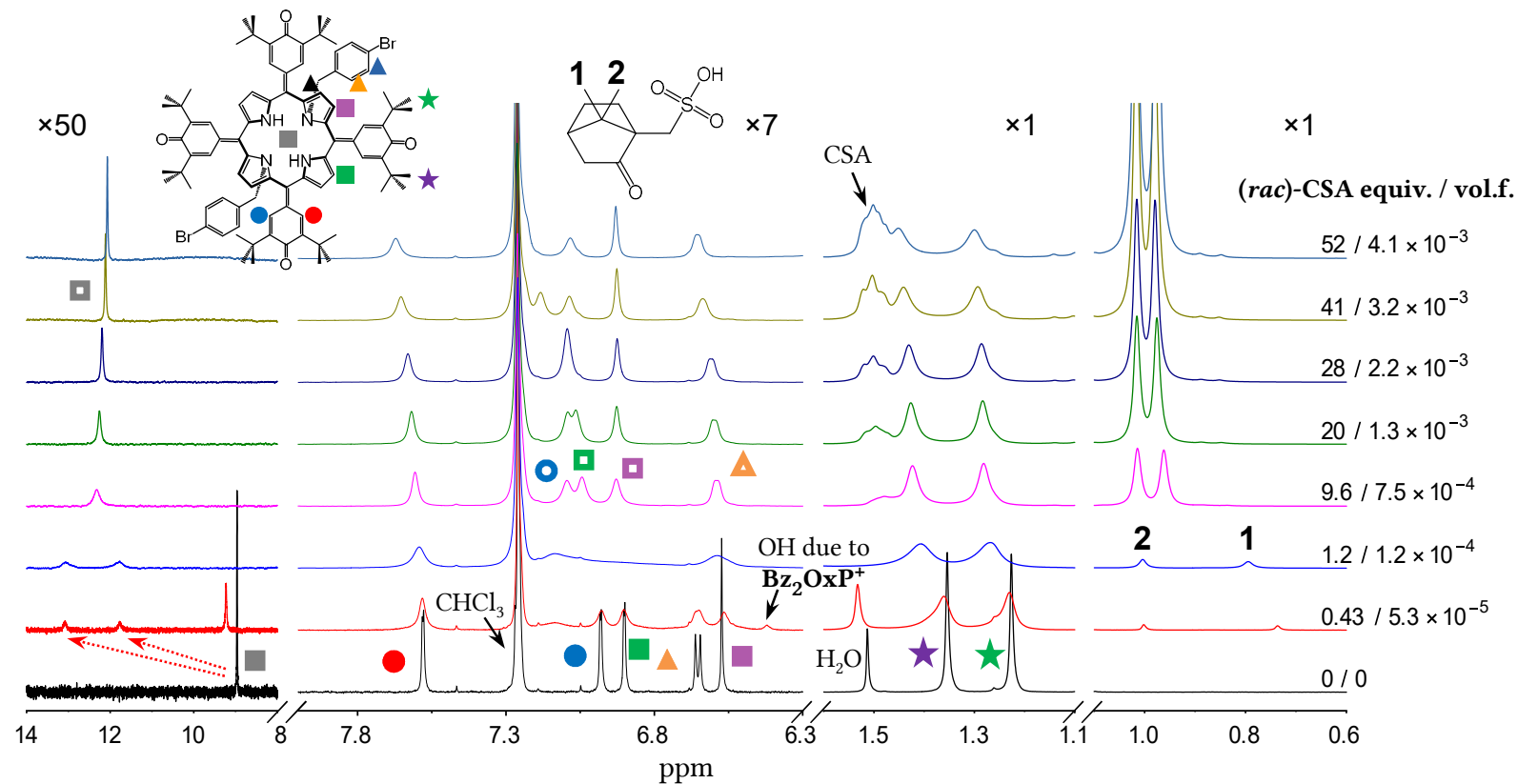


Figure E.6: Titration of **Bz₂OxP** with (*rac*)-CSA (in CDCl₃ at 298 K, host concentration 6×10^{-4} M decreased upon addition of stock solution). Free and protonated/complexed species are denoted by full and empty symbols, respectively. Red dotted arrows denote splitting of *tert*-butyl signal due to symmetry breaking after protonation and anion binding. Certain regions of spectra have been magnified for clarity (see factors above each region). Structure of the protonated form is confirmed by COSY in [45].

E.7 Bz₄OxP + DFA titration

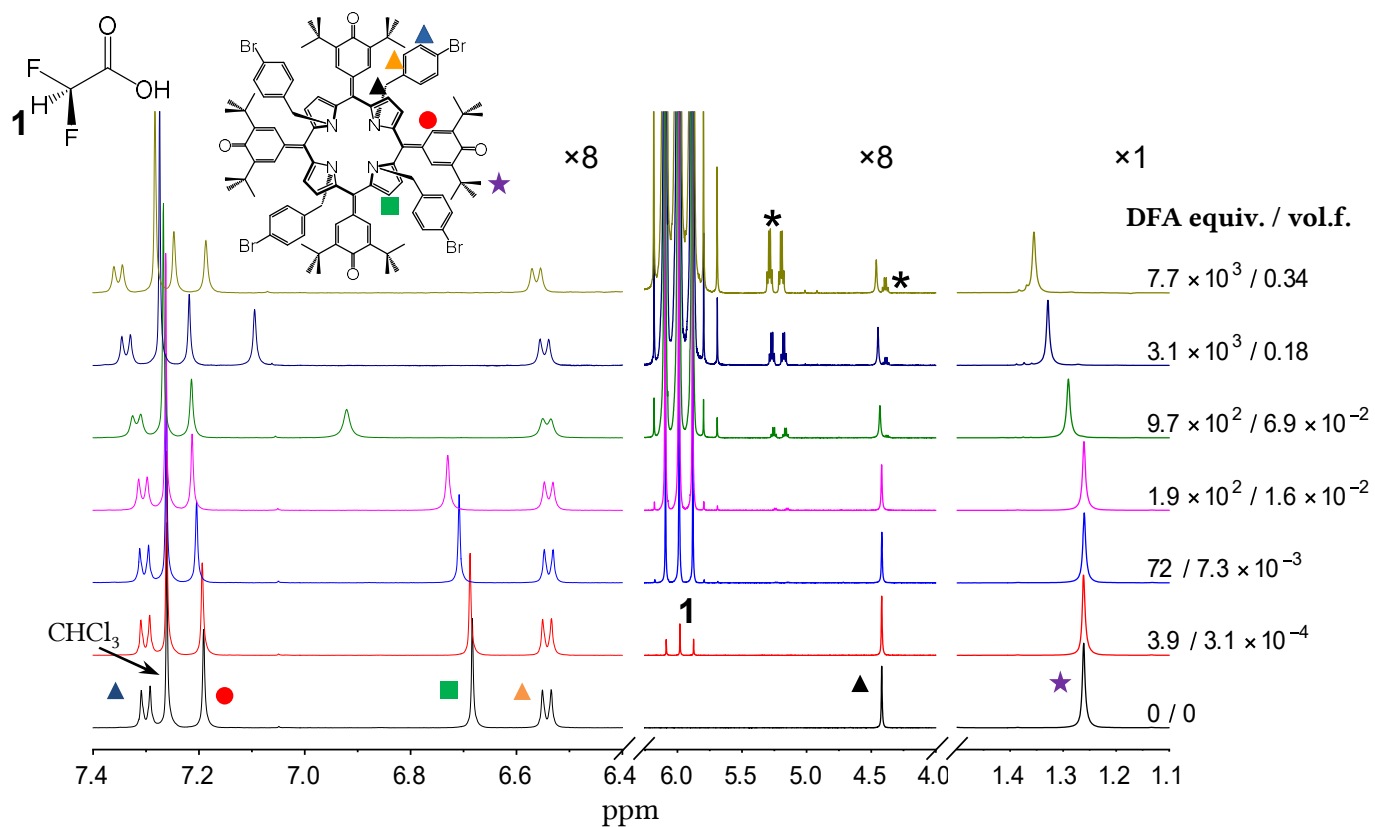


Figure E.7: Titration of **Bz₄OxP** with DFA (in CDCl₃ at 298 K, host concentration 8×10^{-4} M decreased upon addition of stock solution). Asterisk (*) denotes impurities contained in acid. Some regions of the spectra are magnified for clarity (see factors above each region).

E.8 OxP + DFA 1.1 equiv. variable temperature

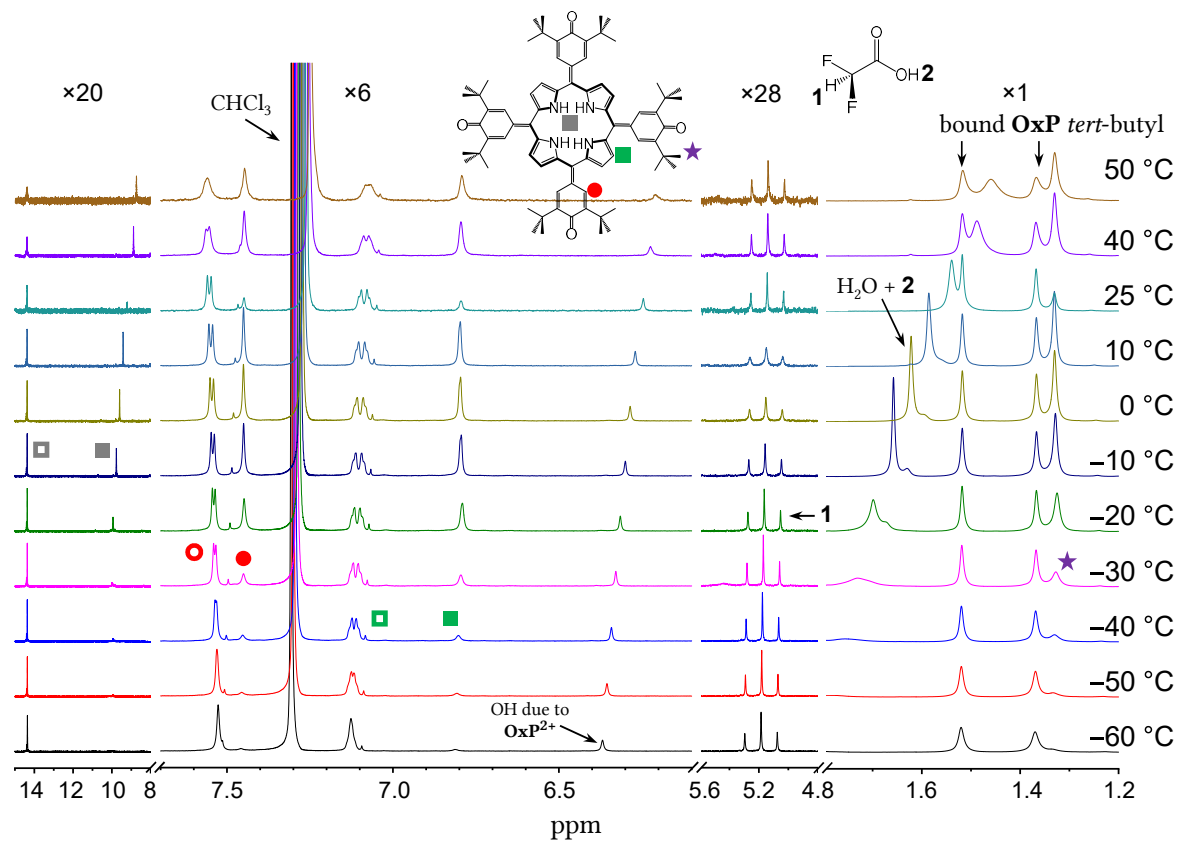


Figure E.8: Variable temperature measurement of OxP with 1.1 equiv. DFA (in CDCl_3 , host concentration 8.9×10^{-4} M). Full and empty symbols denote free and bound species, respectively. The spectrum at 25 °C shows decrease of free host peak intensity (probably due to insufficiently dissolved host since this spectrum was measured first).

E.9 OxP + DFA 18 equiv. variable temperature

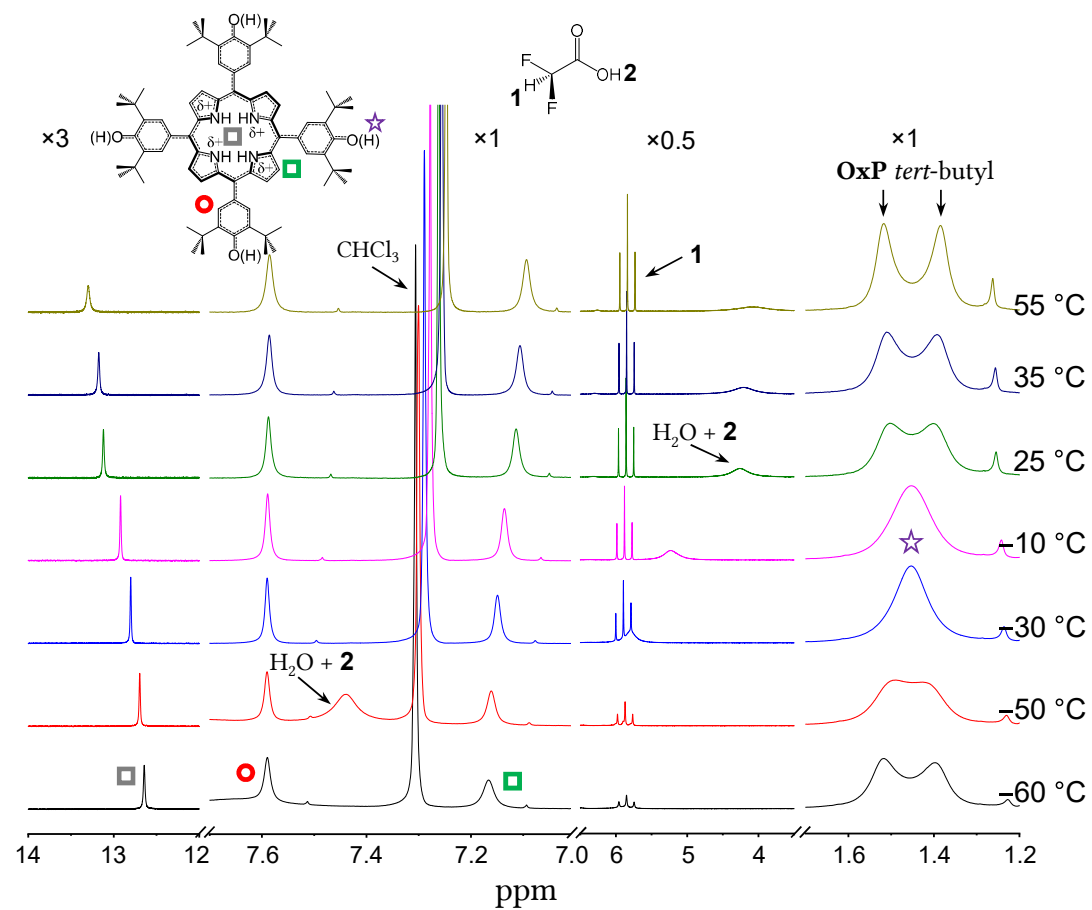


Figure E.9: Variable temperature measurement of OxP with 18 equiv. DFA (in CDCl_3 , host concentration 8.4×10^{-4} M).

E.10 Bz₂OxP variable temperature

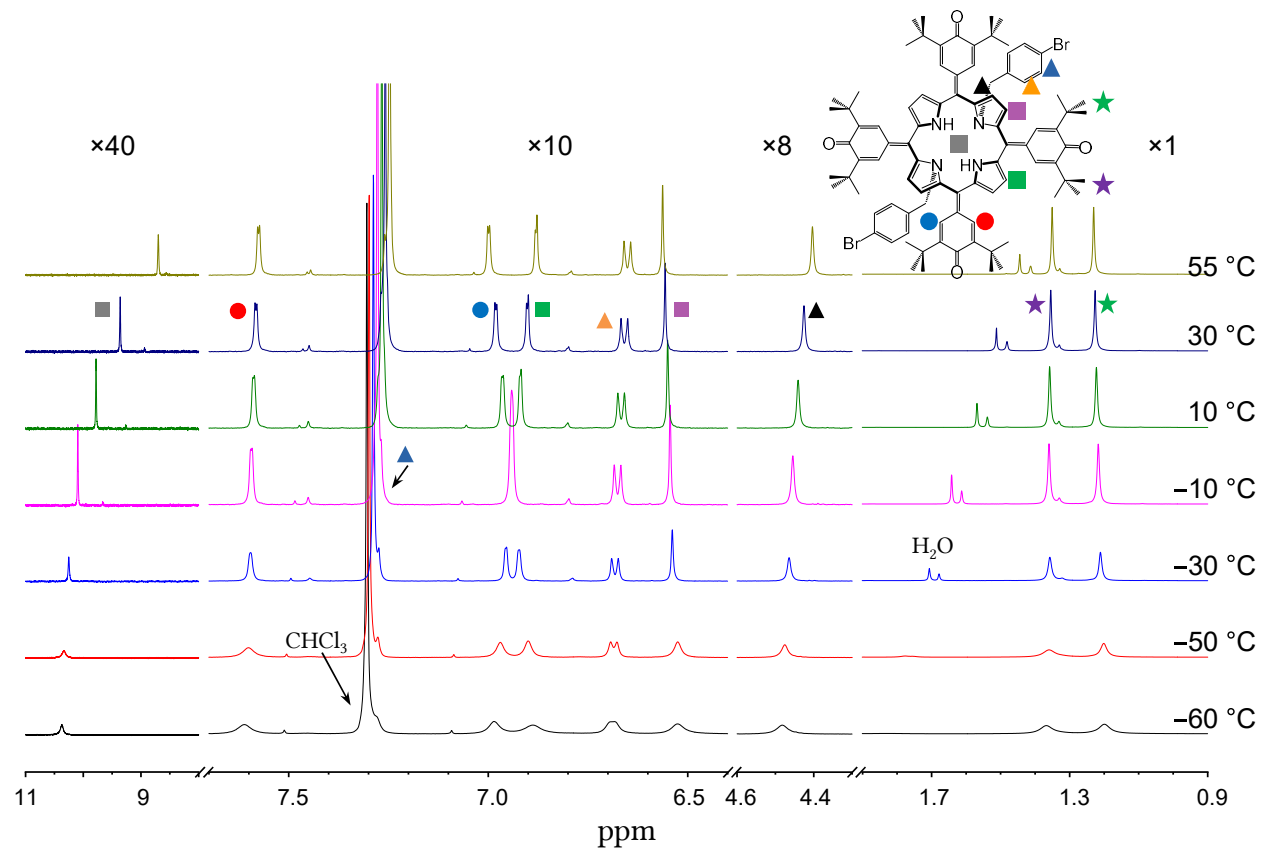


Figure E.10: Variable temperature measurement of pure Bz₂OxP (in CDCl₃).

E.11 Bz₂OxP + DFA 35 equiv. variable temperature

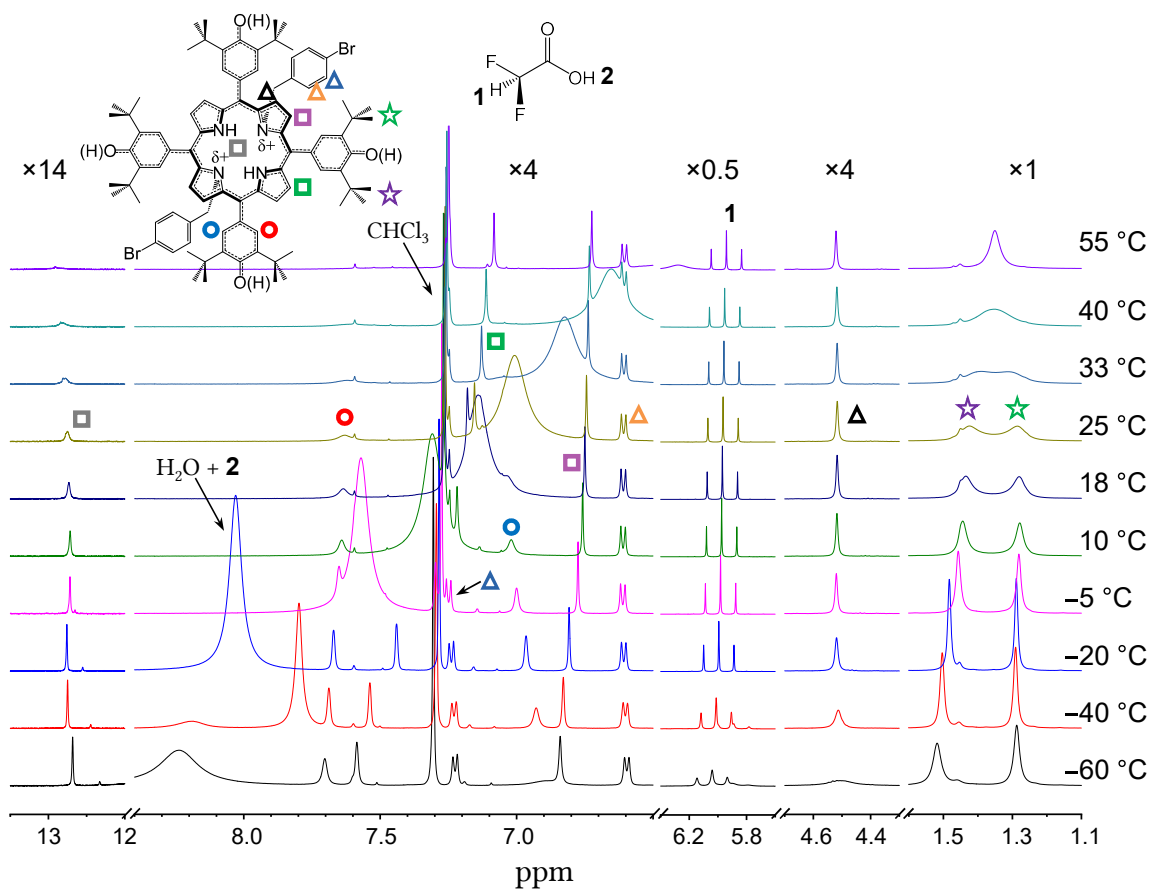


Figure E.11: Variable temperature measurement of Bz₂OxP with 35 equiv. DFA (in CDCl₃, host concentration 7 × 10⁻⁴ M).

E.12 Bz_2OxP + (*R*)-CSA 0.59 equiv. variable temperature

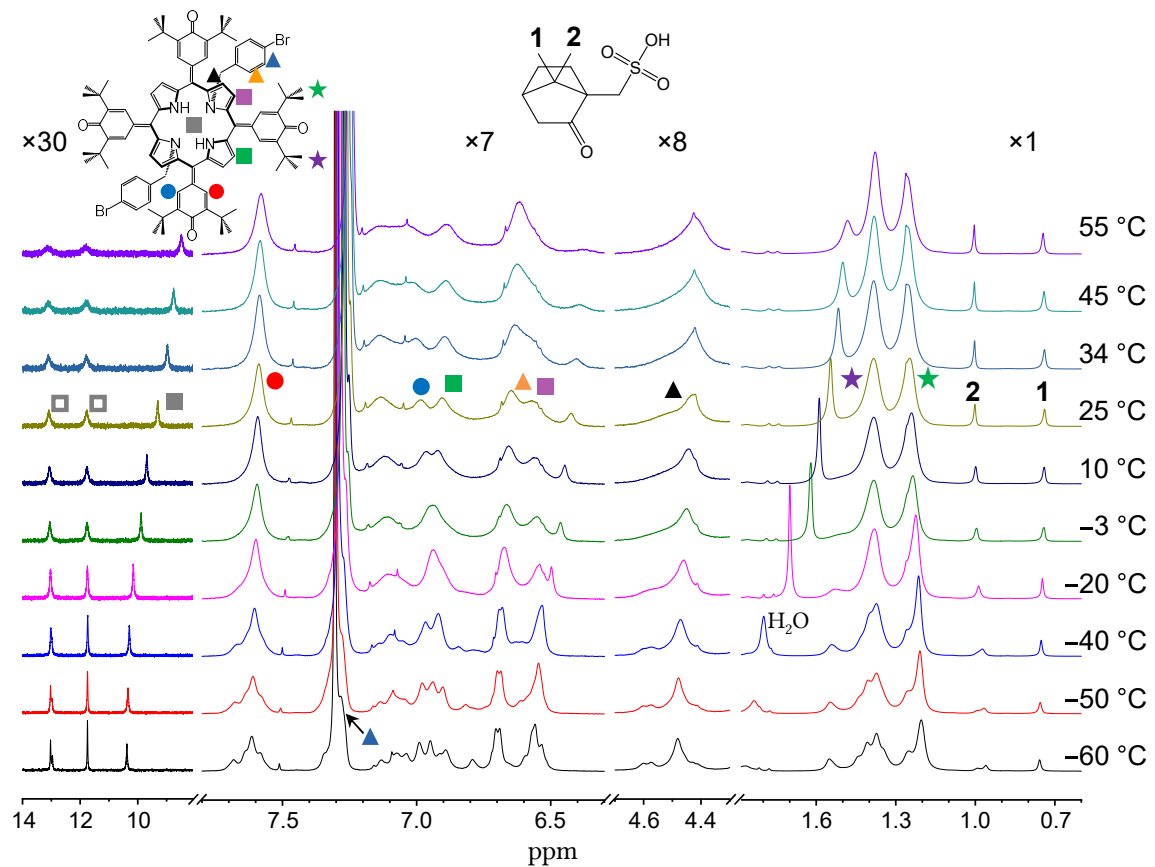


Figure E.12: Variable temperature measurement of Bz_2OxP with 0.59 equiv. (*R*)-CSA (in CDCl_3).

E.13 Bz₂OxP + (R)-CSA 14 equiv. variable temperature

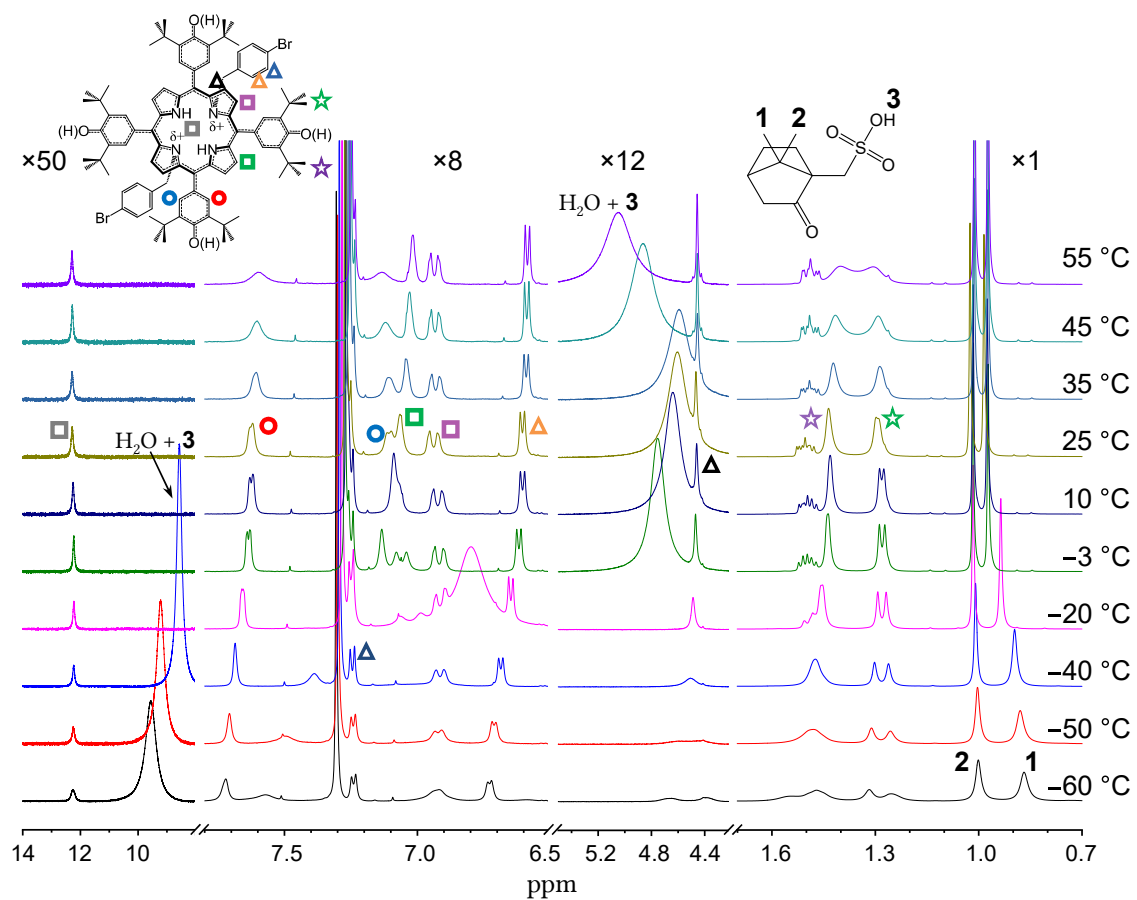


Figure E.13: Variable temperature measurement of Bz₂OxP with 14 equiv. (R)-CSA (in CDCl₃, host concentration 8.4×10^{-4} M). Structure of the protonated form is confirmed by COSY in [45].

E.14 Bz_2OxP + (*rac*)-CSA 0.5 equiv. variable temperature

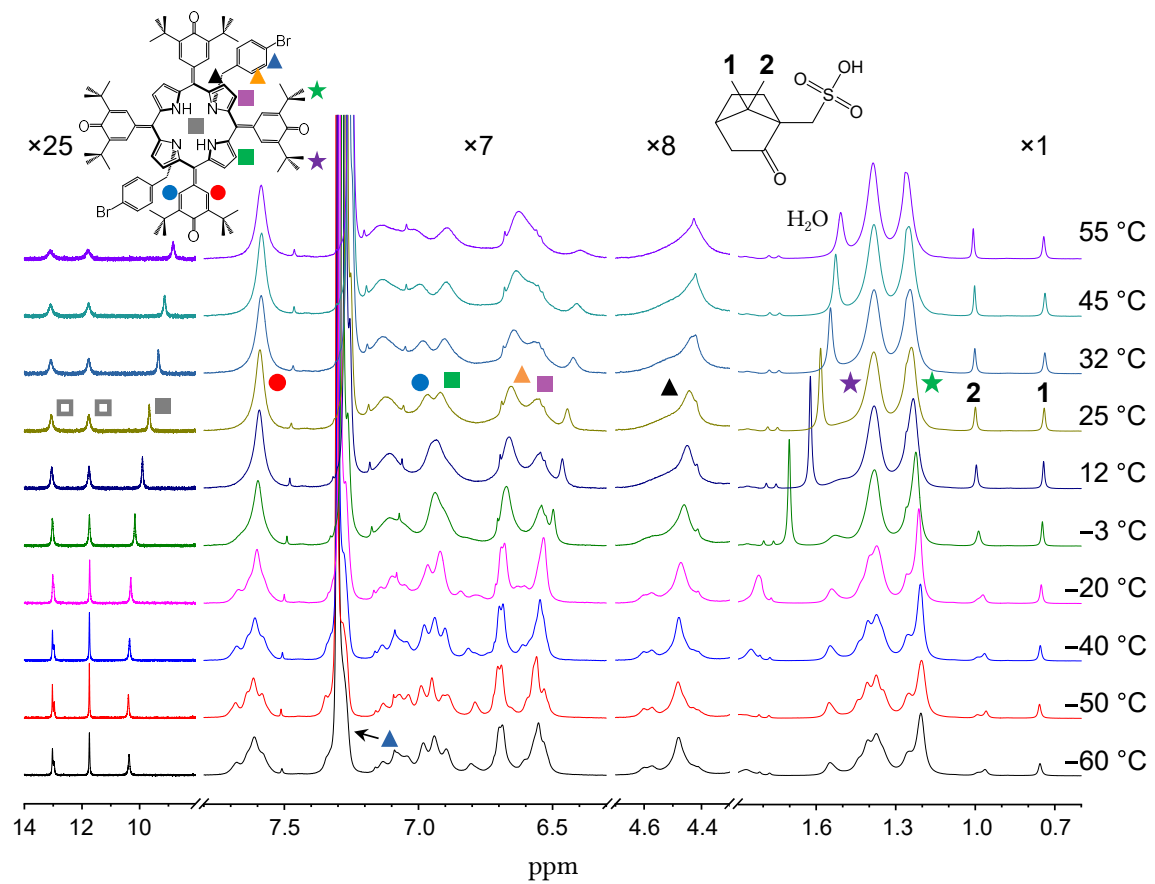


Figure E.14: Variable temperature measurement of Bz_2OxP with 0.5 equiv. (*rac*)-CSA (in CDCl_3).

E.15 Bz₂OxP + (*rac*)-CSA 17 equiv. variable temperature

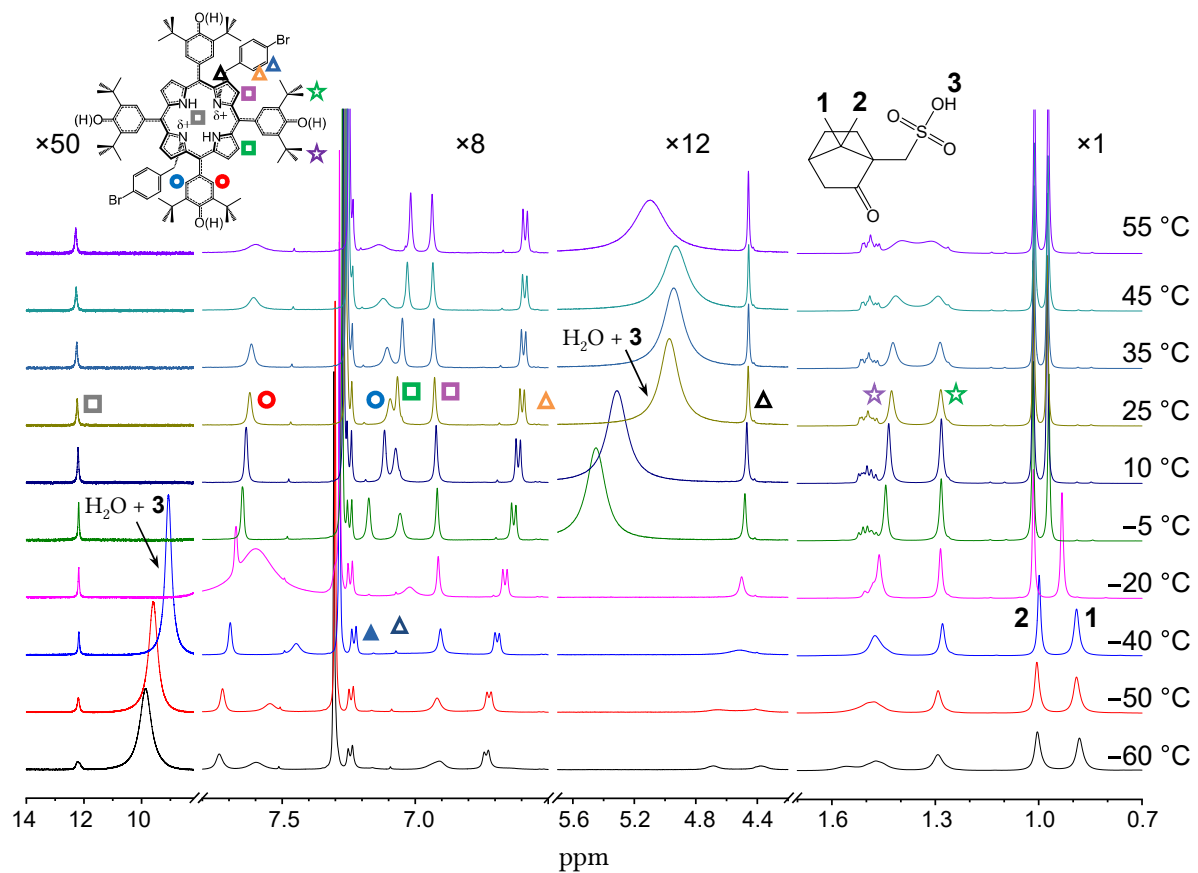


Figure E.15: Variable temperature measurement of Bz₂OxP with 17 equiv. (*rac*)-CSA (in CDCl₃).

E.16 OxP + DMF titration

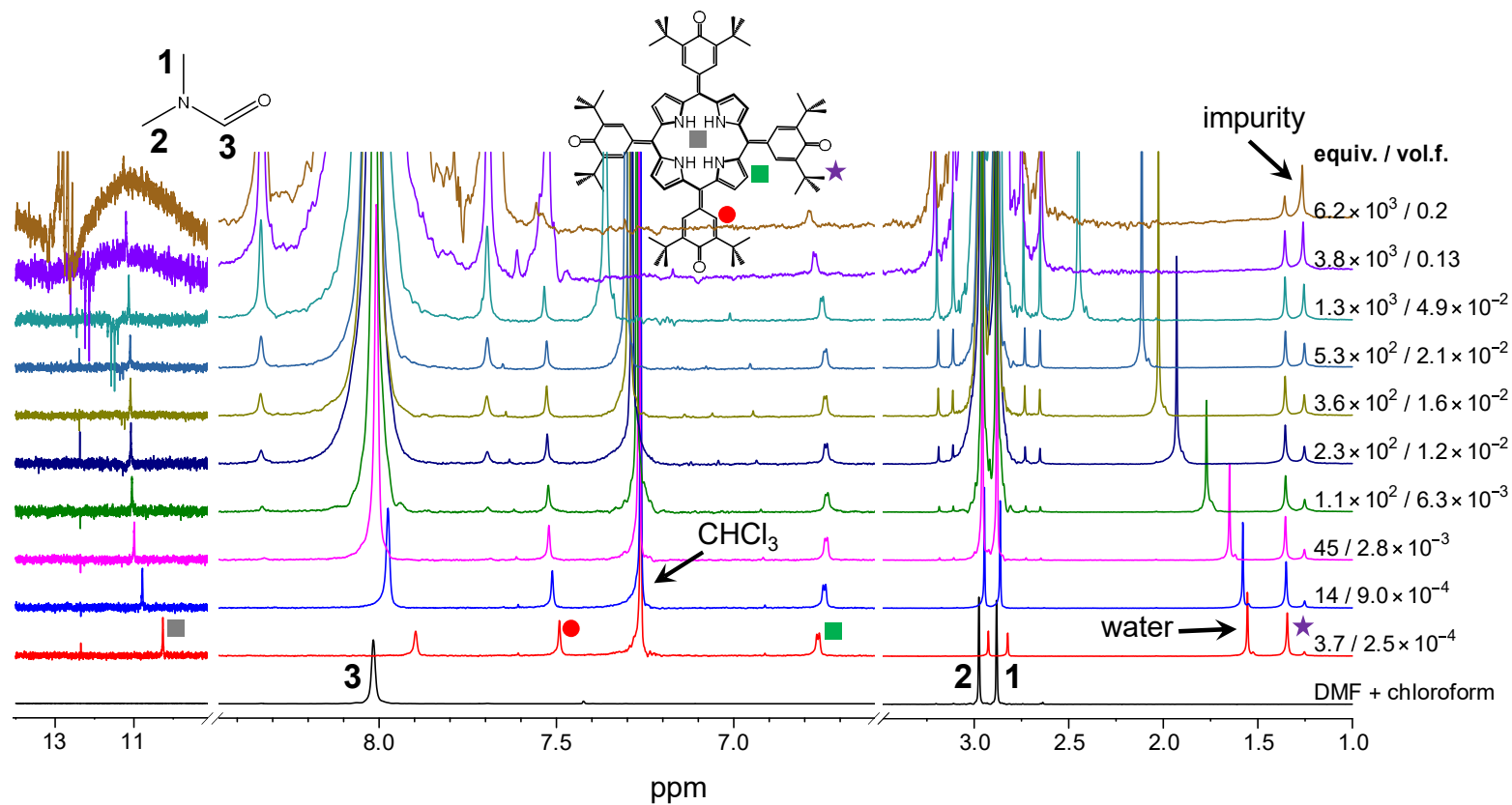


Figure E.16: NMR titration of **OxP** with DMF (in CDCl_3 at 298 K, host concentration 0.001 M decreased to 5×10^{-4} M upon addition of stock solution). Spectra were rescaled to constant **OxP** concentration. Spectrum of DMF in CDCl_3 , which was used as stock solution, is included. Peak of free water (1.56 ppm) is shifting probably due to the solvent polarity change.

E.17 Detection of OH signal in Bz₂OxP

Hydrogen atoms from the host OH groups exchange with hydrogen atoms of water present in the sample. This exchange is mediated by the acid OH groups. Hence, the presence of an OH signal in the host NMR spectrum can be confirmed by addition of small amount of D₂O into the sample. Then, the host OH signal should vanish since the deuterium is not detectable in ¹H NMR measurement. Figure E.17a confirms the presence of OH resonance (red arrow) in **OxP** ·DFA complex. Also, Fig. E.17b suggests assignment of peak at 6.5 ppm as OH (red arrow), although this assignment is not as unequivocal as in the previous case.

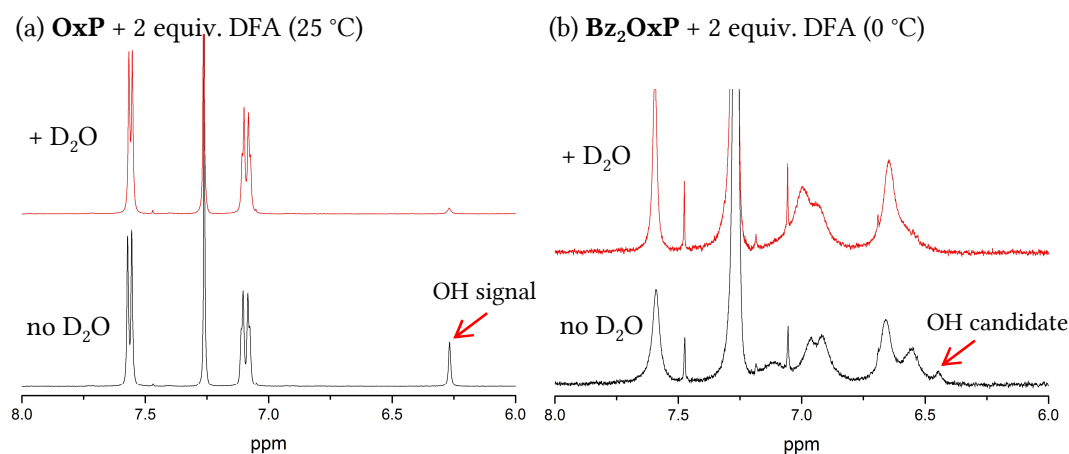


Figure E.17: Confirmation of presence of exchangeable phenolic OH peak (a) bottom: **OxP** + 2 equiv. of DFA (25 °C), top: after addition of D₂O, (b) bottom: **Bz₂OxP** + 0.5 equiv. of CSA (0 °C), top: after addition of D₂O. The OH peak denoted by red arrow disappears in both cases.

E.18 Bz₄OxP assignment - HMBC

Two-dimensional technique Heteronuclear Multiple Bond Correlation (HMBC) detected correlations via chemical bonds between ¹H and ¹³C nuclei. The measurement shows correlations (i) between *tert*-butyl carbon and hemiquinonoid hydrogen (2 ↔ ●, red circle), (ii) between CH₂ carbon and inner benzyl hydrogen (3 ↔ ▲, blue circle). Peaks 1, 2 and 3 were assigned according to their chemical shift and peak area, the HMBC correlation was used to assign ● and ▲.

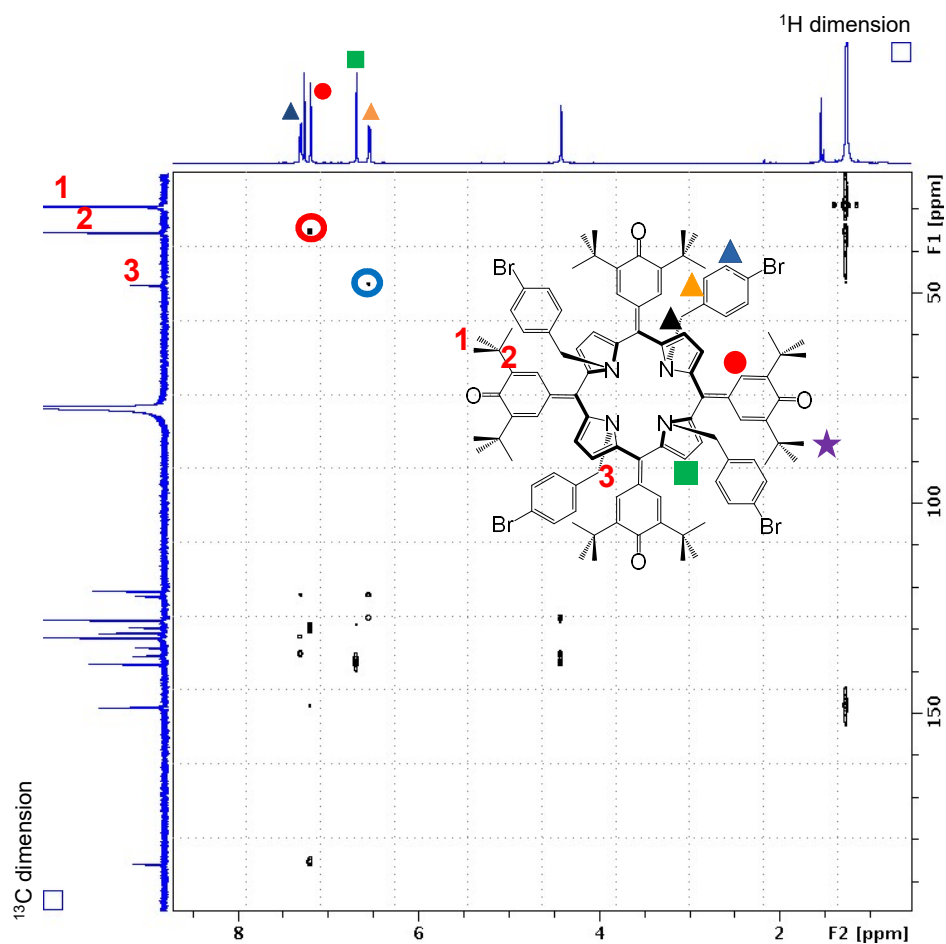


Figure E.18: NMR HMBC measurement of Bz₄OxP. Geometric shapes denote assignment of hydrogen nuclei, numbers denote assignment of carbon nuclei.

List of publications

Publications of the author connected to this thesis

- [1] V. Březina, S. Ishihara, J. Lang, L. Hanyková, K. Ariga, J. P. Hill, and J. Labuta, “Structural Modulation of Chromic Response: Effects of Binding-Site Blocking in a Conjugated Calix[4]pyrrole Chromophore”, *ChemistryOpen* **7**, 323–335 (2018).
- [2] V. Březina, L. Hanyková, N. Velychkivska, J. P. Hill, and J. Labuta, “NMR lineshape analysis using analytical solutions of multi-state chemical exchange with applications to kinetics of host–guest systems”, *Sci. Rep.*, [10.1038/s41598-022-20136-4](https://doi.org/10.1038/s41598-022-20136-4) (2022).

Other publications of the author

- [1] N. Velychkivska, L. Starovoytova, V. Březina, L. Hanyková, J. P. Hill, and J. Labuta, “Improving the Colloidal Stability of Temperature-Sensitive Poly(N-isopropylacrylamide) Solutions Using Low Molecular Weight Hydrophobic Additives”, *ACS Omega* **3**, 11865–11873 (2018).
- [2] M. K. Chahal, J. Labuta, V. Březina, P. A. Karr, Y. Matsushita, W. A. Webre, D. T. Payne, K. Ariga, F. D’Souza, and J. P. Hill, “Knock-on synthesis of tritopic calix[4]pyrrole host for enhanced anion interactions”, *Dalton Trans.*, [10.1039/C9DT02365H](https://doi.org/10.1039/C9DT02365H) (2019).
- [3] D. T. Payne, M. K. Chahal, V. Březina, W. A. Webre, K. Ariga, F. D’Souza, J. Labuta, and J. P. Hill, “Diporphyrin tweezer for multichannel spectroscopic analysis of enantiomeric excess”, *Front. Chem. Sci. Eng.*, [10.1007/s11705-019-1869-1](https://doi.org/10.1007/s11705-019-1869-1) (2020).
- [4] K. Takimoto, S. Ishihara, J. Labuta, V. Březina, D. T. Payne, J. P. Hill, K. Ariga, M. Sumita, S. Mori, and H. Sato, “Enantiomeric Excess Dependent Splitting of NMR Signal through Dynamic Chiral Inversion and Coligand Exchange in a Coordination Complex”, *J. Phys. Chem. Lett.*, 8164–8169 (2020).
- [5] D. T. Payne, J. Labuta, Z. Futera, V. Březina, L. Hanyková, M. K. Chahal, and J. P. Hill, “Molecular rotor based on an oxidized resorcinarene”, *Organic Chemistry Frontiers* **9**, 39–50 (2022).