UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 2262

Bringing Structure to Drug Discovery

HERMINA WIESKE

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2023

ISSN 1651-6214
ISBN 978-91-513-1794-6
URN urn:nbn:se:uu:diva-500237

Dissertation presented at Uppsala University to be publicly examined in Room A1:111a, BMC, Husargatan 3, Uppsala, Friday, 2 June 2023 at 09:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Craig Butts (University of Bristol - School of Chemistry).

Abstract

Wieske, H. 2023. Bringing Structure to Drug Discovery. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 2262. 69 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-1794-6.

Drug discovery is aided by structural information. The type of structural information needed depends on the question at hand. The methods that can be used to determine the absolute configuration of a newly synthesised compound are different from those needed to study ligand binding. This thesis employs a set of structural techniques to study a variety of research questions.

NMR methods were used to understand the binding of an inhibitor to an enzyme responsible for antibiotic resistance. This thesis describes the backbone resonance assignment of the enzyme and the investigation of the protein-ligand interaction. The binding-site as well as the binding-affinity were investigated.

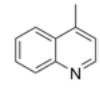
Obtaining insights into the passive membrane permeability of unconventionally large drugs was achieved by looking at their solution ensembles in polar and apolar environments. NMR experiments were used to obtain the solution ensembles of eight antimicrobial and antiviral drugs.

One of the antiviral drugs was studied by MicroED, a new methodology capable of obtaining crystal structures. MicroED requires less material, smaller crystals and the crystals can be of lower quality as compared to conventional X-ray diffraction. This thesis shows that MicroED can be used to obtain the crystal structure of a flexible small molecule that is challenging to elucidate by X-ray crystallography.

The final study of this thesis investigated the applicability of time-saving sampling schemes for the acquisition of quantitative NOESY data. It explores different variables, but none of the investigated conditions lead to the level of accuracy needed for NOE-based distance determination.

Hermina Wieske, Department of Chemistry - BMC, Organic Chemistry, Box 576, Uppsala University, SE-75123 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-513-1794-6
URN urn:nbn:se:uu:diva-500237 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-500237)


N

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I. Wieske, L.H.E., Bogaerts, J., Leding, A.A.M., Wilcox, S., Andersson Rasmussen, A., Leszczak, K., Turunen, L., Herrebout, W.E., Hubert, M., Bayer, A. and Erdélyi, M.* NMR Backbone Assignment of VIM-2 and Identification of the Active Enantiomer of a Potential Inhibitor. *ACS Med. Chem. Lett.*, **2022**, *13*, 257-261

II. Danelius, E., Poongavanam, V., Peintner, S., Wieske, L.H.E., Erdélyi, M.* and Kihlberg, J.* Solution Conformations Explain the Chameleonic Behaviour of Macrocyclic Drugs. *Chem. Eur. J.*, **2020**, *26*, 5231–5244

III. Wieske, L.H.E., Atilaw, Y., Poongavanam, V., Erdélyi, M.* and Kihlberg, J.*, Going Viral: An Investigation into the Chameleonic Behaviour of Antiviral Compounds. *Chem. Eur. J.*, **2023**, e202201280

IV. Bu, G.,† Wieske L.H.E.,† Danelius, E.,* and Gonen, T.* Simeprevir: The First Macrocyclic Drug Elucidated *Ab Initio* by MicroED. *Manuscript*

V. Wieske, L.H.E., Erdélyi, M.* NUS for NOESY? A Case Study on Spiramycin. *Magn. Reson. Chem*., **2021**, *59*. 723-737

Reprints were made with permission from the respective publishers.

† These authors contributed equally; * Corresponding author(s)

This thesis is a continuation of the authors' half-time thesis (January 2021). Sections 1.3, 5.1 and 5.3 are adaptations from the licentiate thesis.

Contributions

Contribution from the author to the listed papers.

I. NMR backbone assignment, titration experiments, and data-analysis. Wrote the first version of the manuscript and contributed to its improvements.

II. NMR assignment and structural elucidation of rifampicin and NMR signal assignment of spiramycin.

III. NMR assignment and solution ensemble determination of asunaprevir and simeprevir and computational refinement of three out of the four antivirals. Wrote the first version of the manuscript and contributed to its improvements.

IV. Simeprevir crystallization, data-collection and manuscript writing.

V. Experimental work, wrote the first version of the manuscript and contributed to its improvements.

The following papers are not included in this thesis.

Poongavanam, V., Atilaw, Y., Ye, S., Wieske, L.H.E., Erdélyi, M., Ermondi, G., Caron, G.* and Kihlberg, J.* Predicting the Permeability of Macrocycles from Conformational Sampling – Limitations of Molecular Flexibility. *J. Pharm. Sci.,* **2020**, *110*, 301-313

Kuhn, S.*, Wieske, L.H.E., Trevorrow, P., Schober, D., Schlörer, N.E., Nuzillard, J.M., Kessler, P., Junker, J., Herráez, A., Farés, C., Erdélyi, M., Jeannerat, D. NMReDATA: Tool and Applications. *Magn. Reson. Chem*, **2020**, *59*, 792-803

Papagna, R.; Engelage, E.; Wieske, L.H.E.; Erdélyi, M.; Huber, S.* Self-complementary Dimers Based on Zwitterionic Halogen Bond Donors. *ChemRxiv.* Cambridge: Cambridge Open Engage; DOI:10.26434/chemrxiv-2022-0xf4t

Wieske, L.H.E., Peintner, S., Erdélyi, M.* Ensemble Determination by NMR Data Deconvolution. *Nat. Chem. Rev.*, **2023**, DOI:10.1038/s41570-023-00494-x

Mahambo, E.T., Uwamariya, C., Miah, M., da Costa Clementino, L., Salazar Alvarez, L.C., Paula Di Santo Meztler, G., Trybala, E., Said, J., Wieske, L.H.E., Ward, J.S., Rissanen, K., Munissi, J.J.E., Costa, F.T.M., Sunnerhagen, P., Bergström, T., Nyandoro, S.S., Erdélyi, M.,* Crotonfolane Diterpinoids and Other Constituents Isolated from *Croton kilwae*. *J. Nat. Prod.*, **2023**, *86*, 380-389

Contents

Abbreviations

α-state	Low-energy spin-state, parallel to the magnetic field
β-state	High-energy spin-state, antiparallel to the magnetic field
η	Observable NOE enhancement
η_{max}	Maximum observable NOE enhancement
σ	Build-up rate
τ_c	Correlation time
Φ	Dihedral angle
ω	Spectrometer frequency
3D	Three-dimensional
$^3J_{HH}$	Three-bond proton-proton scalar coupling
AB	Apical-to-basolateral
bRo5	Beyond rule of five
cLogP	Calculated LogP
CP	Cross-peak
Cryo-EM	Cryogenic electron microscopy
CSP	Chemical shift perturbation
DISCON	Distribution of conformations
DMSO	Dimethyl sulfoxide
DP	Diagonal-peak
ECD	Electronic circular dichroism
FDA	Food and drug administration
FID	Free induction decay
FT	Fourier transform
HBA	Hydrogen-bond acceptor
HBD	Hydrogen-bond donor
HCV	Hepatitis C virus
HIV	Human Immunodeficiency virus
HPLC	High-pressure liquid chromatography
HSQC	Heteronuclear single-quantum correlation
I	Normalized intensity
I_{ref}	Normalized intensity of the reference proton-pair
IST	Iterative soft thresholding
K_d	Dissociation constant
MBL	Metallo-β-lactamase
MCMM	Monte Carlo multiple minimum

MDD	Multidimensional decomposition
ME	Maximum entropy
MicroED	Microcrystal electron diffraction
MIST	Modified iterative soft thresholding
MW	Molecular weight
NAMFIS	NMR analysis of molecular flexibility in solution
NESTA	Nesterov's algorithm
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
NS	Non-structural
NUS	Non-uniform sampling
PANIC	Peak amplitude normalization for improved cross-relaxation
PROTAC	Proteolysis targeting chimera
R^2	Coefficient of determination
r_{ab}	Interproton distance of proton-pair of interest
R_{gyr}	Radius of gyration
RMSD	Root-mean-square deviation
RNA	Ribonucleic acid
Ro5	Rule of five
ROA	Raman optical activity
ROESY	Rotating frame nuclear Overhauser effect spectroscopy
r_{ref}	Interproton distance of reference proton-pair
S/N	Signal-to-noise ratio
SA 3D PSA	Solvent accessible 3D polar surface area
SARS-Cov-2	Sever acute respiratory syndrome coronavirus 2
SCRUB	Scrupulous cleaning to remove unwanted baseline artifacts
Spin I	NOE enhanced spin
Spin S	Perturbed spin (by saturation or inversion)
t_{mix}	Mixing time
VCD	Vibrational circular dichroism
VIM	Verona-integron encoded metallo-β-lactamase
VIM-2	Verona-integron encoded metallo-β-lactamase 2
W_0	Probability for zero-quantum relaxation
W_1	Probability for single-quantum relaxation
W_2	Probability for double-quantum relaxation
XRD	X-ray diffraction

1 Introduction

Treating a disease is more than inhibition of the target receptor or enzyme. When developing a drug one has to take multiple factors into consideration.[1] The drug does not only need to bind its target, but it must be able to reach it, while disrupting the homeostasis of the patient as little as possible.[2] The most popular drugs can be administered orally. The journey the drug takes from pill to target faces many difficulties: it must survive the low pH of the stomach, be able to be absorbed by the gastrointestinal tract and to solubilize in the blood in order to reach the affected tissue through the blood-stream. Once at the tissue of interest, the drug must often overcome a final challenge: it must permeate the lipophilic cell-membrane in order to reach its molecular target, which is frequently found inside the cell.[3]

Designing a new drug needs to consider all these different aspects. Some can be predicted with reasonable accuracy, whereas for other properties medicinal chemists are still shooting in the dark.[4] For all of the different steps, knowing the structure plays an important role. For instance, one needs to know which enantiomer is active. Optimisation of a drug from a co-crystal structure where it is bound to its target requires careful evaluation of how it is bound and in which direction the binder can potentially be expanded.[5, 6] Predicting membrane permeability can be aided by understanding the solution conformations of the drug in a polar and an apolar environment. Every step in drug development is aided by structural information.

There are several methods out there to gain structural information. Each providing its own benefits and downsides, giving different insights into the compound's properties. The work presented in this thesis will address various research questions, by employing several structural techniques. The binding site and binding affinity of a set of inhibitors of a metallo-beta-lactamase (MBL) enzyme were studied by nuclear magnetic resonance (NMR) spectroscopy (Paper I). The interplay between solubility and membrane permeability was explored for a series of complex antimicrobial and antiviral drugs (Papers II-III). Paper IV describes the applicability of microcrystal electron diffraction (MicroED) for obtaining crystal structures of compounds that are hard to elucidate by conventional single crystal X-ray diffraction (XRD). In order to look into the mechanism behind membrane permeability,

nuclear Overhauser effect (NOE) experiments have been employed. These experiments are time consuming and therefore, this thesis addresses whether or not it is possible to speed up data-acquisition by employing non-uniform sampling (NUS) schemes without losing accuracy (Paper V).

1.1 Structure in drug discovery

It is rare that a drug is administered as a racemic mixture, because enantiomers tend to display different bioactivities.[7] Therefore, it is important to know the absolute stereochemistry of a lead compound. The most popular way of determining absolute configuration and three-dimensional (3D) structure (conformation) is by single crystal XRD. This method is well established and obtaining a 3D structure is relatively straightforward once crystals of sufficient size and quality have been grown.[8, 9] It is, however, not always possible to grow the high-quality crystals needed for XRD. Even if it is possible to obtain a single crystal XRD structure, the obtained 3D conformation might not be physiologically relevant. The studies discussed in this thesis employ three different techniques for structure elucidation: NMR spectroscopy (sections 2, 3 and 5), chiroptical spectroscopy (section 2) and MicroED (section 5). Each technique comes with its own set of advantages and disadvantages.

Two-types of NMR experiments are used this thesis: the first set is used to look at ligand-target interactions and the second involves NOE experiments to determine ligand conformations in solution.

The binding strength and whether the target undergoes a conformational change upon binding can be detected by solution-state NMR spectroscopy.[10, 11] $^{1}H,^{15}N$-HSQC (heteronuclear single-quantum correlation) titration is a straightforward way of mapping the binding interaction between a lead compound and its target (section 1.2).[10-13] It shows N-H correlations, allowing one to trace the backbone of a ^{15}N-labelled protein or peptide.

NMR is not only a suitable technique to study protein-ligand binding interactions, but also to identify the solution conformations of organic compounds (300-2000 Da).[14-23] Conformational ensembles can be derived from NOE-based interproton distances[24] and scalar couplings.[25, 26] Determining solution ensembles can be useful to study membrane permeability, for example, but takes several days of spectrometer time and several weeks of data-analysis. It gives the conformations the compound can adopt, and if measured under well-chosen conditions, these are physiologically relevant.

Several types of NMR experiments can capture 3D information, but these are unable to differentiate between enantiomers. In order to determine absolute configuration by NMR, the absolute stereochemistry of at least one

chiral centre needs to be known. This can either be achieved by using a chiral starting material for which the stereochemistry does not change under any of the chemical transformations, or by derivatisation of the final product with a chiral compound of known configuration.[27, 28] NMR requires large amount of sample (0.5–20 mg) and is non-destructive, allowing for the post-measurement recovery of the sample.

Electronic circular dichroism (ECD), vibrational circular dichroism (VCD) and Raman optical activity (ROA) are chiroptical methods that study the three-dimensional structure of a compound in its solution-state.[29, 30] Additionally, they are able to determine the absolute configuration without the need for crystals. ECD measurements are relatively quick (1–10 min) and don't require a large amount of sample (0.001–1 mg). However, in order to assign the absolute configuration, a reference spectrum is needed, either from the literature or from quantum chemical calculations.[31, 32] VCD and ROA measurements require large amounts (2–20 mg) of sample and high concentrations, and take several hours to days to acquire. The assignment of stereochemistry by VCD and ROA require reference spectra as well. For all chiroptical methods, the calculations are computationally demanding. The ones for ECD require more computational power than those needed for VCD and ROA, because the electronically excited states need to be calculated. Similar to NMR, chiroptical methods are non-destructive, which allows for recovery of the sample once the measurements are done.[29-32]

MicroED is a cryogenic electron microscopy (cryo-EM) method that is highly similar to XRD in data-processing and obtained solution.[33] It requires crystalline material, but the crystals only have to be a billionth in size of those needed for XRD.[33] The crystal integrity usually is compromised as it grows. Significantly smaller crystals are more likely to be of high enough quality. Another advantage of MicroED is that it uses a very small amount of material (0.5 mg is sufficient).[34-36] A MicroED data-set is recorded in a matter of minutes and data-analysis can be quick. From sample-prep to data-collection it might take at most a full day of work, collecting several data-sets for a routine sample. High-quality data can be processed in a couple of hours, leading to a structure within 2 days.

1.2 $^{1}H,^{15}N$-HSQC titration studies

NMR spectroscopy can be used to study ligand-target binging in solution,[10] allowing for elucidation of the binding site and binding dynamics at the same time.[11-13] $^{1}H,^{15}N$-HSQC spectra show at least one signal for every amino-acid residue, apart from proline, and hence visualise the entire protein while showing reduced spectral overlap when compared to $^{1}H,^{13}C$-HSQC spectra. This allows for tracking of chemical shift changes upon additions of a binder, and may even give insight into the binding affinity.[11-13] The exchange-rate between the bound and free forms of the protein affects the appearance of the spectra. For very weak binding, the spectra may only show the free form, since the binding event does not occur frequent enough and is too short-lived to be observed by NMR.[10] On the contrary, the binding event of strong binders lasts longer than the NMR time-scale.[10, 37] For a strong binding event, both, the free and the bound forms of the protein are observed for all the spectra where the ligand is present and the protein is not saturated yet. The positions of the peaks don't change upon further addition of the ligands, but the intensity of the peaks are affected by the increasing ligand concentration. Gradual shift changes are observed when the binding-event is short and the binding even occurs frequently, resulting in a fast exchange rate (10^4 s^{-1}).[11]

From $^{1}H,^{15}N$-HSQC titration experiments the binding site and the binding affinity in the form of the dissociation constant (K_d) can be obtained.[11] The binding-site can be determined by identification of the residues that undergo the largest chemical shift perturbation (CSP).[38] Upon addition of the ligand the peak can shift in both the ^{1}H and the ^{15}N dimensions, which is typically averaged to a single value describing the spectral change (Equation 1).[38]

$$\Delta\delta_{NH} = \sqrt{\Delta\delta_H{}^2 + (\frac{\Delta\delta_N}{R_{scale}})^2}\ ;\ R_{scale} = 6.5 \qquad \text{Equation 1}$$

The CSP is indicated by $\Delta\delta_{NH}$ in the above equation, $\Delta\delta_H$ and $\Delta\delta_N$ are the perturbation in the f2 and f1 dimension, respectively and R_{scale} is a scaling factor to account for the different chemical shift ranges of the two nuclei involved.[38] K_d can be estimated by plotting the $\Delta\delta_{NH}$ as a function of ligand concentration (Equation 2).[11]

$$\delta_{obs}([\mathrm{L}]) = \delta_{max} * \frac{([P]+K_d+[L])-\sqrt{([P]+K_d+[\mathrm{L}])^2-4[P][L]}}{2[P]} \qquad \text{Equation 2}$$

In the above equation, $\Delta\delta_{obs}([L])$ is the observed chemical shift change at a given ligand concentration $[L]$, $\Delta\delta_{max}$ is the maximum chemical shift change observed for the residue, K_d is the dissociation constant and $[P]$ is the protein

concentration.[11] Equation 2 only applies to systems with a single binding-site.

1.3 The nuclear Overhauser effect

The nuclear Overhauser effect was first observed in 1953 by Albert Overhauser in the form of a polarisation transfer between the electron spin magnetic moment and the nuclear spin magnetic moment of metals.[39] The NOE later found its application in NMR spectroscopy, was described to be the result of dipole-dipole relaxation and could be used to obtain spatial information of organic compounds.[40-43] Development of the application of the NOE for biological macromolecules was pioneered by Kurt Wüthrich in the late 1970s and early 1980s.[44, 45] Carefully carried out NOE experiments can give interatomic distances with high accuracy, 3.0-6.9% average error depending on the solvent viscosity.[46-48]

The NOE is a signal enhancement that arises from a through-space dipole-dipole interaction. Herein the homonuclear NOE of spin-1/2 nuclei (I=1/2), typically ^{1}H-spins, is discussed. There are two spin orientations associated with a spin-1/2 nucleus when it is exposed to a static magnetic field. It is either parallel (α-state) or antiparallel to the magnetic field (β-state).[49] The α-state has a lower energy than the β-state, and is therefore slightly more populated at equilibrium.[50] When the system is perturbed from equilibrium, for example by spin inversion, as is the case for transient NOE experiments (Figure 1), it falls back to equilibrium through various relaxation pathways. Single-quantum relaxation, with probability W_1, involves the flip of a single spin. Such single-quantum relaxation pathways do not give rise to the NOE and generally occur slower than the zero-quantum and the double-quantum relaxation pathways, which have probabilities W_0 and W_2, respectively. These, in contrast, give rise to the NOE and are known as the cross-relaxation pathways. Depending on which relaxation pathway dominates, the observed NOE is either positive or negative.[49-51] The zero-quantum relaxation pathway ($\beta\alpha \leftrightarrow \alpha\beta$) has no net energy change associated with it for a homonuclear system. The double-quantum relaxation pathway ($\beta\beta \rightarrow \alpha\alpha$) moves two spins into a lower energy state, and is therefore associated with a larger energy transition. Compounds that tumble slowly in solution don't have the energy needed for a double-quantum transition and are therefore dominated by zero-quantum transitions, displaying negative NOEs.[50, 52]

The NOE has originally been defined as a difference experiment, where a reduction in the population difference between the α- and the β-states resulted in an intensity decrease (negative NOE), whereas an increase in the population difference resulted in an increase of signal intensity (positive NOE).[49, 50, 52]

A transient NOE experiment starts with the inversion of spin S which moves the spin-state away from the equilibrium (where α>>β) to a state where α<<β (Figure 1).

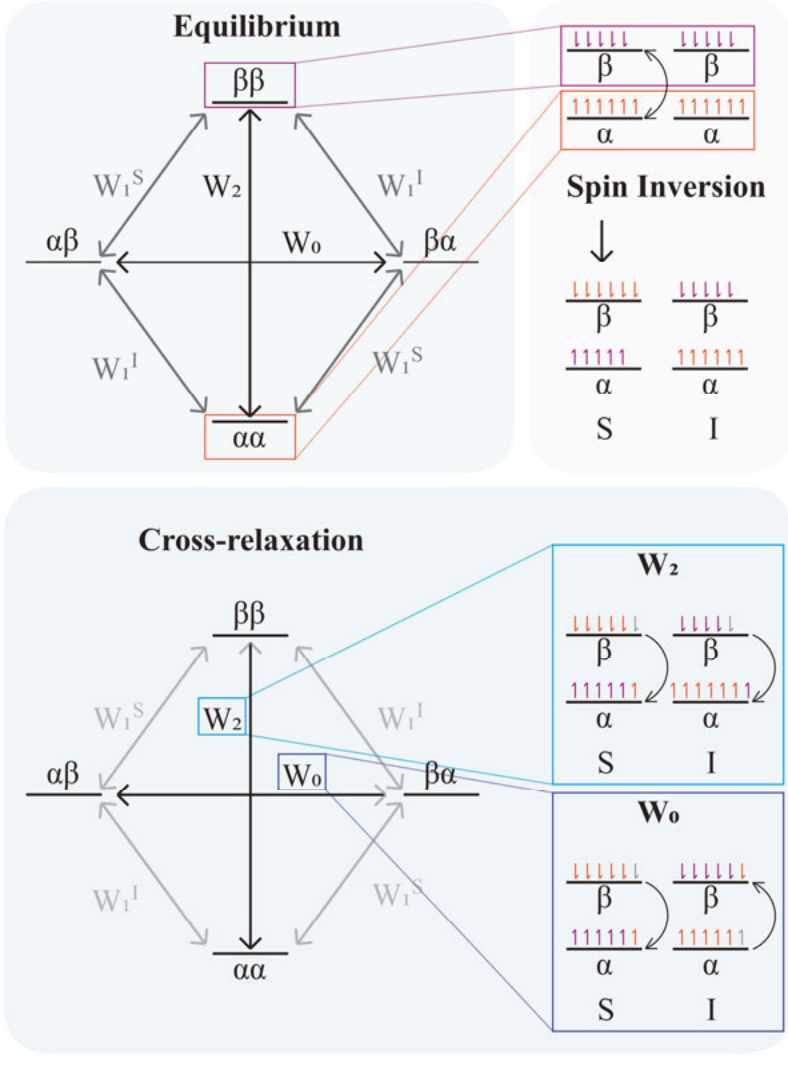


Figure 1. Energy diagram of a homonuclear two-spin system. At equilibrium, the lower energy α-state is slightly more populated than the β-state. Upon spin-inversion of spin S, the system falls back to equilibrium via multiple possible relaxation pathways. The zero-quantum and double-quantum relaxation pathways, with probabilities W_0 and W_2, give rise to the NOE.[53-56]

In order to regain equilibrium conditions by cross-relaxation pathways, the system can either undergo the double-quantum relaxation pathway, which moves spins from the β-state back into the α-state for spin S, while at the same

time also moving β→α for spin I. This causes a population change for spin I from α>>β to α>>>β resulting in an increased population difference and therefore a positive NOE.[57-59] Alternatively, spin S falls back to equilibrium through the zero-quantum relaxation pathway, where β→α transitions for spin S occur in combination with α→β transition for spin I. As a result, the population difference for spin I shifts from α>>β to α>β, resulting in a smaller population difference, and hence a negative NOE.[56, 60-62]

The probabilities for the different relaxation pathways W_1, W_2 and W_0 depend on the spectrometer frequency (ω) and the correlation time (τ_c).[50, 51, 56, 63] Larger magnetic fields result in a larger the gap between the α- and β-state. The speed at which a molecule tumbles in the magnetic field is associated with a certain energy. If this energy matches the energy gap between the ββ and the αα states, a double-quantum transition can occur. Large compounds and/or compounds in highly viscous solvents have long correlation-times and therefore lower W_2. Whereas W_2 depends on both ω and τ_c, W_0 only depends on τ_c and scales linearly with it (a change in W_1 and W_2 affects W_0).[50, 51] Figure 2 shows the dependency of the maximum observable NOE (η_{max}) as a function of τ_c for various field strengths. If W_0 and W_2 are about equal, no net NOE is observed. This happens when $\omega\tau_c$ roughly equals 1.12 and is known as the zero-crossover point. NOEs around the zero-crossover point are very weak and therefore difficult to observe.[51, 52]

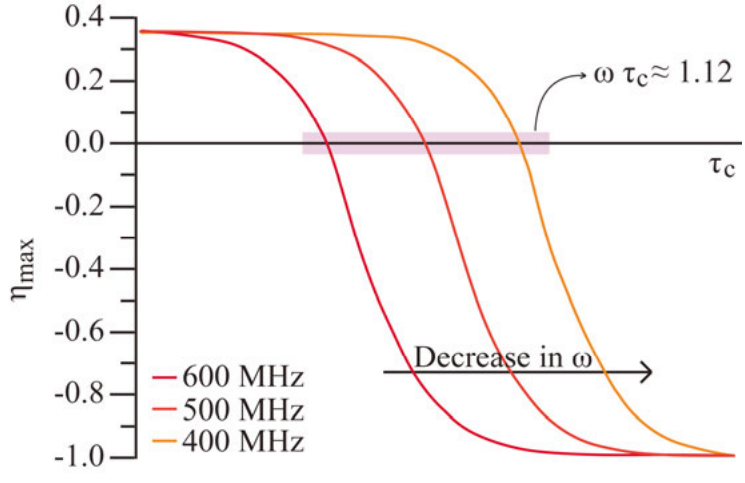


Figure 2. Maximum NOE (η_{max}) as a function of the spectrometer frequency (ω) and the correlation time (τ_c). Small compounds in low-viscosity solvents have a short τ_c resulting in positive NOEs. Larger compounds, such as biologically macromolecules, have a long τ_c and give rise to large negative NOEs. When $\omega\tau_c \approx 1.12$, the contribution of the positive NOE is about equal to the negative NOE, giving rise to no net observable NOE.[24, 56]

Transient NOE experiments rely on spin-inversion rather than spin-saturation and can range from +0.38 for positive NOEs to -1 for negative NOEs. NOESY is a 2D transient NOE experiment. Dipolar couplings do not give rise to splitting of the signals in a sample, where the molecules tumble isotropically, since the free movement of the molecules average the dipolar coupling to zero.[52]

1.4 The initial-rate approximation

When recorded appropriately, transient NOE spectra can be used to calculate accurate internuclear distances.[24] The NOE intensities can be converted into distances using the initial rate approximation, the assumption that the interproton distance is proportional to the inverse sixth power of the NOE in the initial, linear part of an NOE build up curve. Here, the magnitude of the NOE depends on the mixing time (t_{mix}), used at data acquisition. The longer the mixing time, the stronger the NOE (η) until the mixing time becomes so long that other relaxation effects start to significantly influence the signal intensity (Figure 3). [24, 64]

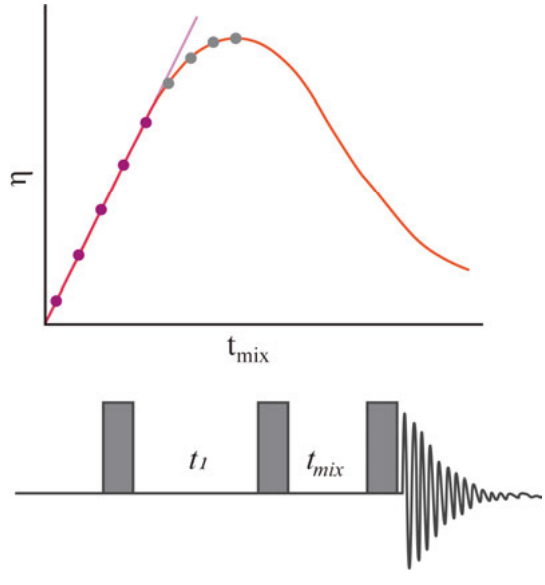


Figure 3. The NOE build-up curve (top), with the observable NOE enhancement (η) plotted as a function of the mixing time (t_{mix}). Intensities from NOESY spectra recorded at different t_{mix} that fall within the linear-regime are indicated as purple dots. Grey dots represent intensities from NOESY spectra that were recorded beyond the linear regime and therefore the initial-rate approximation is no longer valid. The pulse-sequence for a transient 2D NOESY sequence showing the placement of the mixing time, t_{mix}, in the pulse-sequence is shown in the bottom.

The interproton distance can either be calculated as a function of the normalised intensities (I) or as a function of the build-up rate (σ). The build-up rate can be established by determination of the slope of the build-up curve (linear fit to purple data-points of Figure 3). In order to be able to plot the build-up curve for a given proton-pair, NOESY spectra are acquired at different mixing times. The observed intensities are normalised against the diagonal peaks (a.k.a. the PANIC [peak amplitude normalisation for improved cross-relaxation] method)[24, 47] and the normalised intensities (Equation 3) are plotted as a function of the mixing time.

$$I_{ab} = \sqrt{\frac{CP_{ab} * CP_{ba}}{DP_a * DP_b}} \qquad \text{Equation 3}$$

In the above equation, I_{ab} is the normalised intensity between spin H_a and H_b, CP_{ab} and CP_{ba} are the intensities of the two cross-peaks and the two diagonal peak intensities of H_a and H_b are given by DP_a and DP_b, respectively.

At shorter mixing times the NOE intensity is proportional to the interproton distance ($r_{ab} \propto I^{1/6}$). This is where the build-up of the NOE intensity is linearly dependent on the mixing time, and dictates the mixing-time interval where the initial-rate approximation is valid (Equation 4).[48, 65] With the help of the normalised intensity measured for a proton-pair (I_{ref}), referenced to a known distance (such as geminal[46] or ortho aromatic protons[17], r_{ref}), the interproton distance (r_{ab}) can be determined. Beyond the linear regime of the build-up curve, the observed intensity is no longer solely the result of the NOE and Equation 4 can no longer be applied for distance quantification.

$$r_{ab} = r_{ref} * \left(\frac{I_{ref}}{I_{ab}}\right)^{\frac{1}{6}} = r_{ref} * \left(\frac{\sigma_{ref}}{\sigma_{ab}}\right)^{\frac{1}{6}} \qquad \text{Equation 4}$$

1.5 Scalar couplings

$^3J_{HH}$ scalar couplings give spatial information, since their magnitude is correlated to the dihedral angle by a trigonometric function, as described by Martin Karplus in 1959.[26] This relationship, known as the Karplus equation, was improved by Haasnoot and co-workers (Equation 5), and is in wide use today.[25]

$$J(\Phi) = A\cos^2(\Phi) + B\cos(\Phi) + C \qquad \text{Equation 5}$$

In this equation the constants A, B and C are dependent on the atom-type the proton is attached to as well as on the nearby substituents, whereas Φ is the

dihedral angle. Every vicinal proton-pair is best described by a unique Karplus equation. However, there are some general constants that work well for certain vicinal relationships.[25, 26, 66-68] Examples of more elaborately studied vicinal relationships are HC-CH,[25] and HC-NH,[69, 70] dihedrals.

1.6 Aim of the thesis

The aim of this thesis is to give structural understanding for a set of medicinal chemistry related challenges.

Section 2 aims to assign the NMR backbone resonances of VIM-2 in order to allow solution studies, such as performing binding studies with its inhibitors (Paper I). It aims to identify the active enantiomer of a known inhibitor that was previously only studied as a racemic mixture.

Section 3 aims to use NMR ensemble determination to shed light on the origin of the passive membrane permeability of a set of beyond rule of 5 drugs (Papers II and III). It explores the conformational ensembles of eight antiviral and antimicrobial agents, and aims to identify the extent of chameleonic behaviour they display.

Section 4 aims to elucidate the crystal structure of simeprevir by MicroED (Paper IV). This study showcases the applicability of MicroED for flexible mid-sized molecules, and addresses the sample preparation needed to obtain data.

Section 5 aims to assess the applicability of non-uniforms sampling schemes for quantitative NOE studies (Paper V). The findings of this study are directly applicable to projects looking at solution structures by NMR, such as the ones employed in section 3.

2 Inhibiting VIM-2: a fight against resistance (Paper I)

Micro-organisms are in constant competition for resources with one and another. The production of chemical agents specifically targeting bacteria is a strategy that several fungi and even some bacteria employ to win the battle for the limited amount of resources. Alexander Fleming discovered the effect of such an agent for the first time in 1929.[71] The agent he isolated from the fungus *P. rubrum* is now known as penicillin, the first known antibiotic. After the discovery of penicillin, its medical application was quickly found, and more antibiotics were actively searched for.[72] However, their success could only be celebrated for a limited period of time, because development of resistance soon followed after their introduction.[73, 74] The discovery and development of new antibiotic agents became less frequent as time moved on, until the point that antibiotic resistance occurs more rapidly than the development of a new antibiotic agent.[72-74] More multi-drug resistant strains appear in clinical settings, posing a real threat to our global health.[75, 76]

2.1 Antibiotic resistance mechanisms

Antibiotic agents affect components of cell processes that are unique to prokaryotic cells and leave eukaryotic cells unaffected. Antibiotics commonly disrupt the peptidoglycan cell-wall, interfere with the prokaryotic ribosome, or with the replication mechanisms.[72-74] In order to reach the ribosome and replication machinery, the drug has to pass at least one membrane. In case of antibiotics targeting the cell-wall, the drug must pass one membrane for gram-negative bacteria, but none for gram-positive bacteria.[54, 77, 78]

In response to the exposure to antibiotics, bacteria can develop resistance in multiple ways. Bacteria can obtain resistance by i) alteration of the target,[79-83] ii) de-activation of the antibiotic either by degradation or by derivatisation,[79, 84, 85] iii) over-expression of efflux pumps, making it near impossible for the drug to reach its target,[74, 83, 86] or iv) limiting uptake by alteration of the cell-wall or biofilm formation, resulting in reduced permeability of the antibiotic.[87-89] It is not uncommon for multi-drug resistant

strains to employ a combination of antibiotic resistance mechanisms, introducing yet another complication to the treatment of infections.[72-74]

2.2 Verona-integron encoded metallo-β-lactamase 2 (VIM-2)

Most of the clinically relevant strains express beta-lactamases as one of their main resistance mechanisms.[75] Beta-lactamases are a class of enzymes that hydrolyse the beta-lactam ring of beta-lactam antibiotics. Penicillins, cephalosporins and the last-resort antibiotic carbapenems are some of the beta-lactam antibiotics in use today (Figure 4).[75] The active-site of a beta-lactamase can either hydrolyse the beta-lactam ring with the help of a serine residue, or by having one or two Zn^{2+} ions coordinated by a catalytic water molecule which performs the hydrolysis.[75, 90, 91] The first group of enzymes are known as serine-beta-lactamases, whereas the latter are known as metallo-beta-lactamases (MBLs).[75]

Figure 4. Core structures of penicillins, cephalosporins, and carbapenems with the beta-lactam moiety highlighted in orange.

Beta-lactamases can be divided into four Ambler classes:[92] class A are the extended spectrum beta-lactamases, B are the metallo-beta-lactamases, C are the AmpC beta-lactamases, and D are the oxocillinases.[61, 75, 91] This classification is based on the structure of the beta-lactamases, and therefore, all beta-lactamases within a certain class have similar active-sites. Of the four different classes, class B (the MBLs) are the clinically most relevant beta-lactamases and they can be further divided into three sub-classes based on their catalytic site.[75, 90] This division relies on the number of zinc-ions, and on the residues that coordinate the Zn^{2+}. The B1 MBLs are the most prevalent MBLs found in multi-drug resistant infections, and include New-Delhi metallo-beta-lactamases, imipenemases, German imipenemases and Verona-integron encoded metallo-beta-lactamases (VIMs).[93, 94] This work focusses on VIM-2, since no NMR data was available prior to the study published in 2022 (Paper I). Solution-state NMR studies can provide structural and kinetic information at the same time. Making the backbone resonances (^{1}H, ^{15}N, $^{13}C_{\alpha}$,

$^{13}C_\beta$) available is the first step needed to obtain relevant solution-state NMR data.

2.3 NMR backbone assignment

In order to be able to map the binding-site by ^{1}H,^{15}N-HSQC titration experiments, the resonances of the individual amides need to be known. NMR backbone assignment is done by sequential assignment where a series of 3D spectra are used to map out the backbone resonances (Figure 5).[95] The spectra are named after the correlations they show: HNCO,[96-98] HNcaCO,[99] HNCA,[96-98] HNcoCA,[96, 97] HNCACB,[100, 101] and HNcoCACB.[102]. By combining information on common chemical shifts from these spectra, it is possible to identify neighbouring residues. The HNCO, HNcoCA and HNcoCACB spectra show correlations between the ^{1}H and ^{15}N resonances of one residue and the ^{13}C resonance of residue (n-1). The HNCO spectrum of residue i and the HNcaCO spectrum of residue (i+1) share the ^{1}H and ^{15}N resonances. Similarly, the HNCA spectrum of residue j shares the $^{13}C_\alpha$ resonance of the HNcoCA spectrum of residue (j+1).

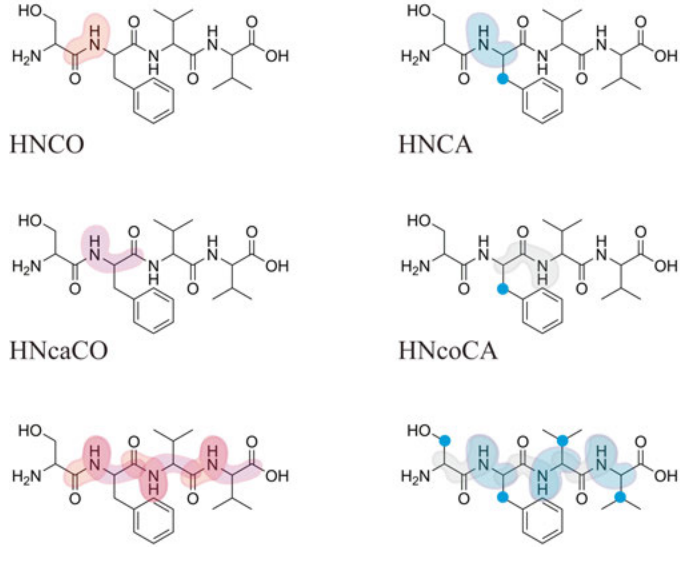
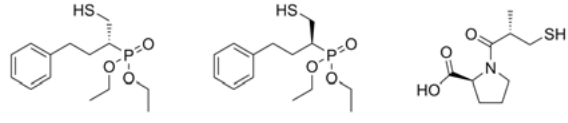


Figure 5. Schematic representation of sequential backbone assignment on the tetrapeptide Ser-Phe-Val-Val. A combination of an HNCO and HNcaCO or HNCA and HNcoCA spectra can be used to link the amino-acid residues. A combination of HNCACB and HNcoCACB gives rise to the same correlations as the HNCA/HNcoCA combination, but with an additional resonance for the $^{13}C_\beta$ resonances (blue dots). The bottom sequence shows the overlapping resonance that these two sets of spectra have in common.

Since no gradual shift change was observed for VIM-2 (it showed slow exchange) a total of three set of backbone assignments were performed. The first one was recorded under conditions set to maximise the number of visible resonances: at 37 ˚C in absence of any ligand. The other two sets were recorded under the same conditions as the titration experiments were performed: at 25 ˚C with an additional 4.8% ethanol and in absence and presence of 1.5 equivalents of L-captopril.

2.4 Inhibitor scope

Once the NMR backbone assignment of VIM-2 was known, $^{1}H,^{15}N$-HSQC titration studies could be performed (section 1.2).[11, 103] A set of known binders (Figure 6) was used for the first titration studies, which revealed some interesting behaviour of VIM-2.



1-*R* **1**-*S* L-captopril

Figure 6. The three binders used for $^{1}H,^{15}N$-HSQC titration studies. The two enantiomers of compound **1**, which have been previously been studied as a racemic mixture,[104] and L-captopril.[105]

L-Captopril has been reported to bind to VIM-2 in the low μM range (IC_{50} 4.4 μM),[105] whereas compound **1** had only been studied as a racemic mixture previously (IC_{50} 0.89 μM).[104] The two enantiomers were separated using chiral chromatography, and studied by NMR to allow for identification of the active enantiomer.

2.5 Results and discussion

Under optimal conditions a total of 193 backbone amides were assigned (84%) for VIM-2 (Figure 7). Of these, 178 could be traced back for the sample without L-captopril at 25 ˚C and 190 residues were confirmed for VIM-2 in presence of L-captopril at 25 ˚C. These backbone resonances were used for the subsequent $^{1}H,^{15}N$-HSQC titration studies.

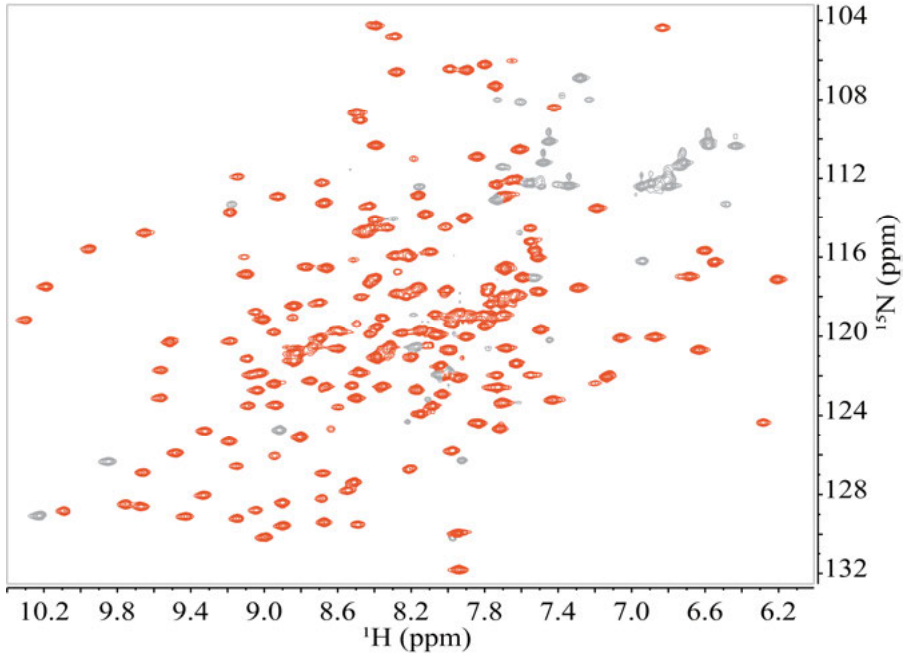


Figure 7. The ^{1}H,^{15}N-HSQC spectrum for VIM-2 at 37 ˚C, with assigned correlations highlighted in orange.

Compound **1** was re-synthesised and successfully separated by chiral HPLC (high-pressure liquid chromatography) after which the absolute stereochemistry was determined by chiroptical methods. To increase solubility, the two enantiomers were first dissolved in small amounts of ethanol and subsequently further diluted to the desired concentration in buffer, giving rise to a total ethanol concentration of 4.8%. All three compounds displayed slow exchange behaviour and generally affected the same amino acid residues, indicating that they occupy the same binding-site (Figure 8). Titration experiments on VIM-2 with L-captopril revealed dissociation constants similar to those reported in literature (0.63–6.8 μM).[105, 106]

Furthermore, the titration experiments for the two enantiomers of compound **1** revealed that the *S*-enantiomer is the active enantiomer, with K_d values between 0.23 and 3.9 μM, whereas the *R*-enantiomer showed affinities in the mM range.

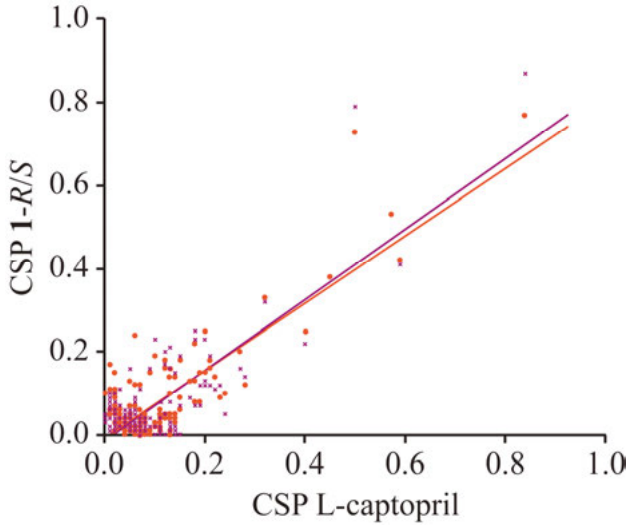


Figure 8. The chemical shift perturbations (CSPs) of VIM-2 upon titration of the two enantiomers of compound **1** (**1**-*S* in orange and **1**-*R* in purple) plotted against those observed upon titration of L-captopril. The linear fit for the correlations are given as solid lines **1**-*S* (R^2=0.82), **1**-*R* (R^2=0.80).

2.6 Summary

The emergence of multi-drug resistant microbial infections is an increasing problem worldwide, and requires new treatment strategies.[76] One solution can be combination therapy where antibiotics and inhibitors, blocking enzymes that give rise to resistance, are combined. This is considered to be a realistic short-term solution to the problem.[90] Such a combination treatment should ideally include a broad-spectrum MBL inhibitor. One of the MBLs that should be targeted is VIM-2.

Paper I reports the first NMR backbone assignment for VIM-2, which will open up the possibility to study VIM-2 by NMR to a wider audience. The resonance assignment can aid in the discovery of VIM-2 inhibitors and subsequent broad-spectrum MBL inhibitors. Furthermore, the identification of the active enantiomer of a known inhibitor will help in the further design and development of similar binders.

3 Molecular chameleons: the discovery of a new species (Papers II-III)

Section 2 has looked at the interaction between a set of small organic compounds and an enzyme giving rise to antibiotic resistance (Paper I). The ligands that were investigated all follow the Lipinski's rule of five (Ro5).[107] However, not all diseases can be treated with drugs that conform to the Ro5,[60, 108-111] some require inhibitors that show a larger degree of flexibility. These beyond rule of five (bRo5) drugs cannot be represented accurately by a single 3D structure in solution, but are more likely to exist as a set of conformations that are in rapid exchange with one and another.[17, 112] The conformations they adopt is dependent on both the degree of flexibility and the environment the drug is in. Looking at these compounds with solution-based methods can shed light on the dynamic processes they are involved in. Solution-state NMR contains information about all conformations in solution, even though it most frequently gives rise to only one set of signals.[112] This one set of signals contains the time-averaged information of the drug in solution. Deconvolution algorithms such as DISCON (distribution of conformations),[14] NAMFIS (NMR analysis of molecular flexibility in solution)[17] and StereoFitter[20] are able to select a set of conformations from time-averaged NMR data such as NOE-derived interproton distances and coupling constants. Papers II and III looked at the chameleonic properties of eight antimicrobial and antiviral drugs.

3.1 The space in between

Traditional drugs are small organic compounds that mostly fall within a chemical space described by a set of descriptors, which was first published by Lipinski in 1997.[107] These are the so-called Lipinski's rule of five (Ro5) and consist of a set of four guidelines involving multiples of the number five: the compound has a mass of 500 Da or smaller, a maximum of 5 hydrogen-bond donors and 10 hydrogen-bond acceptors and a calculated LogP value of 5 or lower.[107] Targets that can be drugged with Ro5 complying drugs most often have a well-defined binding pocket. On the far other end of the spectrum are biological therapeutics.[113] Biological therapeutics have been around for a long time, one of the most well-known examples being insulin, whose effects

were reported as early as 1922.[114, 115] These types of drugs are generally highly specific to their target, are larger and contain many hydrogen bond acceptors and donorf, therefore requiring administration through injection and don't display passive membrane permeability, preventing them from reaching intracellular targets.[113] However, the space in between the small organic compound and biological therapeutics seemed to have been overshadowed by the findings of Lipisnki in 1997, preventing unhindered exploration of larger and more complex drug-like compounds for a while.[116] The set of molecular properties Lipinski[107, 117] and (subsequently) others[118-121] came up with were never intended to be strict rules for drug design. Instead they were intended as guidelines that increases the chance of designing a drug that is simultaneously water soluble and membrane permeable, and shows metabolic stability while acting as a potent binder.[55] Thankfully, in more recent years, medicinal chemists ventured more frequently into the space beyond the rule of five, resulting in an increase in drugs non-compliant with the Ro5 approved by the FDA (food and drug administration).[59, 108-111] These non-compliant molecules are also known as "beyond rule of five" (bRo5) drugs. As people started to venture into the space in between, (peptidic) macrocycles were discovered amongst the first as drugs that show oral availability and membrane permeability while at the same time inhibiting difficult-to-drug targets.[122, 123] After the discovery of more and more potent orally available bRo5 drugs, their membrane permeability became a topic of interest.[59, 109, 116, 122] One of the main hypotheses today is that most orally bioavailable bRo5 drugs benefit from displaying chameleonic behaviour, where they are able to adjust their solvent accessible 3D polar surface area (SA 3D PSA) and radius of gyration (R_{gyr}) to their environment.[60, 124] This adjustment is achieved by forming intramolecular interactions (such as hydrogen bonds) that will shield polar groups from an apolar environment.[108, 109, 123-129] Additional descriptors such as the number of rotatable bonds[118] and the Kier's flexibility index[130] can be used to assess whether a potential inhibitor may display chameleonic behaviour. However, chameleonic behaviour, and thus oral availability and membrane permeability, is notoriously hard to predict, especially from 2D descriptors.[131] Obtaining 3D structures of bRo5 drugs helps unravelling the mechanisms by which chameleonic behaviour is achieved.

3.2 Antimicrobial and antiviral agents

In nature, several antimicrobial agents are produced that don't comply with the Ro5, yet have proven to be powerful orally available antibiotics.[109, 125, 127, 128] Medicinal chemists have taken advantage of this gift and modified some of the naturally occurring antimicrobials to increase their potency. A subset of these bRo5 antimicrobials contain a macrocyclic core (12 or more non-hydrogen atoms in the ring).[132, 133] Four orally available membrane permeable

macrocyclic antimicrobials were selected to be further studied for their chameleonic properties (Paper II). These antimicrobials have a biosynthetic origin, however, a set of synthetic antiviral compounds, that don't comply with the Ro5 either while displaying favourable pharmacokinetic properties, were studied in addition to the antimicrobials (Paper III, Table 1). In total eight bRo5 antimicrobial and antiviral agents (Figure 9), both cyclic and linear, were investigated in polar and apolar environments to assess their chameleonic behaviour.

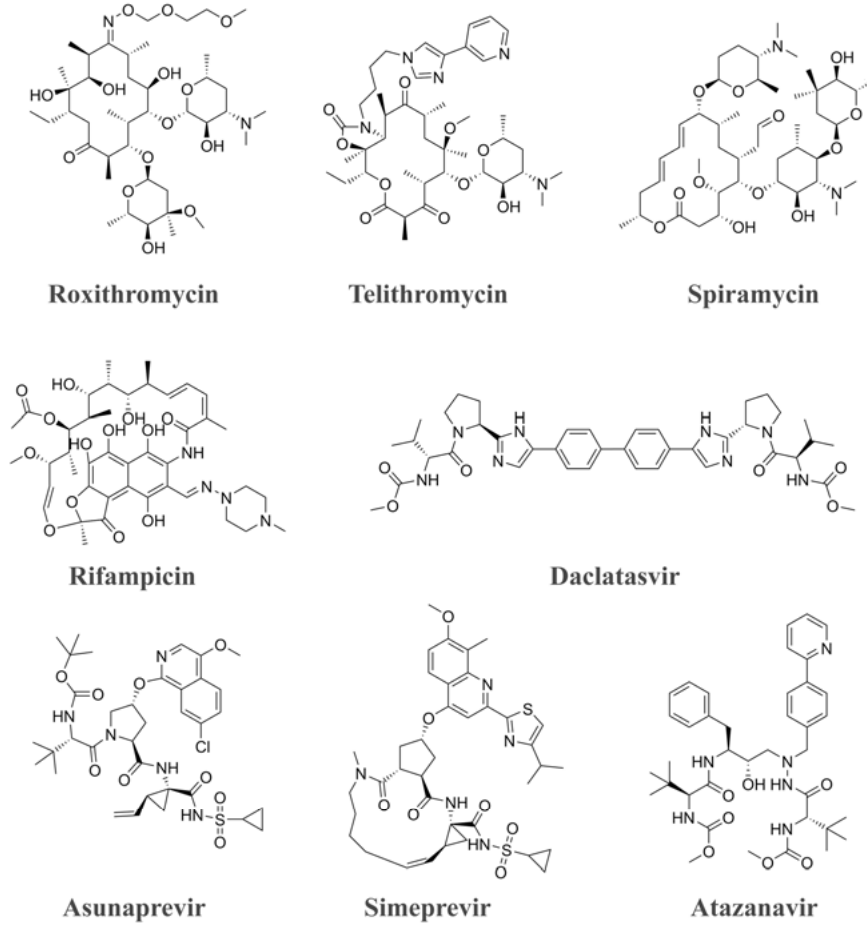


Figure 9. The eight antimicrobial and antiviral compounds studied for their chameleonic properties. The antimicrobials consist of the macrocycles roxithromycin, telithromycin, spiramycin and rifampicin. Daclatasvir, asunaprevir, simeprevir and atazanavir are all antivirals and, apart from simeprevir, are all linear.

Roxithromycin, telithromycin and spiramycin are known to interfere with the protein synthesis of prokaryotic cells by binding to the 50S ribosomal

subunit.[134-137] Rifampicin binds to the β subunit of the RNA polymerase thereby interfering with the RNA synthesis of prokaryotic cells.[138] Of the antiviral compounds, daclatasvir, asunaprevir and simeprevir all target the hepatitis C virus (HCV),[139-143] whereas atazanavir inhibits the aspartyl protease of human immunodeficiency virus (HIV).[60, 144] Asunaprevir and simeprevir both bind to the HCV NS3/NS4a serine protease, thereby inhibiting synthesis of the polyprotein chain.[143] Daclatasvir inhibits the NS5a protein, which is involved in RNA replication.[140]

Table 1. The molecular weight (MW), number of hydrogen-bond donors (HBD), hydrogen-bond acceptors (HBA) and a calculated lipophilicity (cLogP) value for the Ro5 upper limits the eight investigated compounds.

	MW (Da)	HBD	HBA	cLogP
Ro5	500	5	10	5.0
Roxithromycin	837	5	16	3.0
Telithromycin	812	1	11	5.4
Spiramycin	843	4	15	2.5
Rifampicin	823	6	14	2.8
Daclatasvir	739	4	8	5.5
Asunaprevir	748	3	8	4.4
Simeprevir	750	2	9	4.0
Atazanavir	705	5	7	4.7

Despite their difference in origin (natural product derivative *versus* synthetic) and their variation in functional groups, these eight drugs have three properties in common: they all violate the Ro5, show some degree of passive membrane permeability (Table 2) and are administered orally. All eight compounds are significantly larger than 500 Da, but asunaprevir, simeprevir and atazanavir stay within the limits of the Ro5 when it comes to number of hydrogen-bond donors, acceptors and the cLogP value. Telithromycin exceeds all limits except for the number of hydrogen-bond donors, making it on paper the most non-compliant drug of the set.

The solubility of the drugs (Table 2) varies greatly from having good aqueous solubility, as is the case for roxithromycin and telithromycin, to having poor solubility, in particular daclatasvir and atazanavir. Despite the poor solubility of these drugs, they are all part of existing treatment against bacterial,[145-148] HCV[149, 150] or HIV[151] infections. Besides a spread in solubility, a spread in membrane permeability was observed. Even though the passive membrane permeability might be low for some, the fact that these drugs still display some passive membrane permeability makes them interesting to investigate.

Table 2. Solubility data and Caco-2 permeability data for the investigated drug set.

	Solubility (μM)[a]	P_{app} AB + inhibitors ($x10^{-6}$ cm/s)[b]
Roxithromycin	1510 (24)	11.9 (1.6)
Telithromycin	1960 (141)	4.3 (0.1)
Spiramycin	327 (46)	0.19 (0.0029)
Rifampicin	183 (8)	1.0 (0.1)
Daclatasvir	0.50 (0.1)	15 (1.6)
Asunaprevir	160 (49)	48 (9.9)
Simeprevir	19 (3)	0.50 (0.1)
Atazanavir	2.3 (1.1)	72 (3.6)

[a] Determined in aqueous potassium phosphate buffer at pH 7.4. Values are means ± std from three repeats (Papers II and III).

[b] P_{app} AB + inhibitors: permeability in the apical-to-basolateral (AB) direction across Caco-2 cell monolayers, determined in the presence of a cocktail of three inhibitors that target the three major efflux transporters. Values are means ± std from three repeats.

3.3 Obtaining solution conformations

To obtain the solution ensembles of the eight bRo5 drugs, NAMFIS (NMR analysis of molecular flexibility in solution)[17] was used to deconvolute the experimental data. NAMFIS fits population-weighted conformations to the experimental data. It shows the best-fit solution together with the back-calculated interproton distances and scalar coupling of the fit for the given experimental input data.[17] In order to run NAMFIS, two text-files are needed: one summarising the experimental data, consisting of NOE-derived interproton distances and scalar coupling constants, and a second with the corresponding interproton distances and scalar coupling for every conformation present in the theoretical input ensemble (Figure 10). There are several methods that can be used to generate the theoretical input ensemble. For NMR data deconvolution, it is most important to sample the entire feasible conformational space. The energies calculated for the corresponding conformations are neglected for the fitting procedure, which is experimental data driven. Monte Carlo multiple minimum (MCMM) methods are excellent for this task because they are computationally inexpensive and not energy driven, making it likely that they sample the entire feasible conformational space.[152-154] To maximise sampling of the conformational space, several MCMM calculations, using different force-fields and solvation models, can be combined.

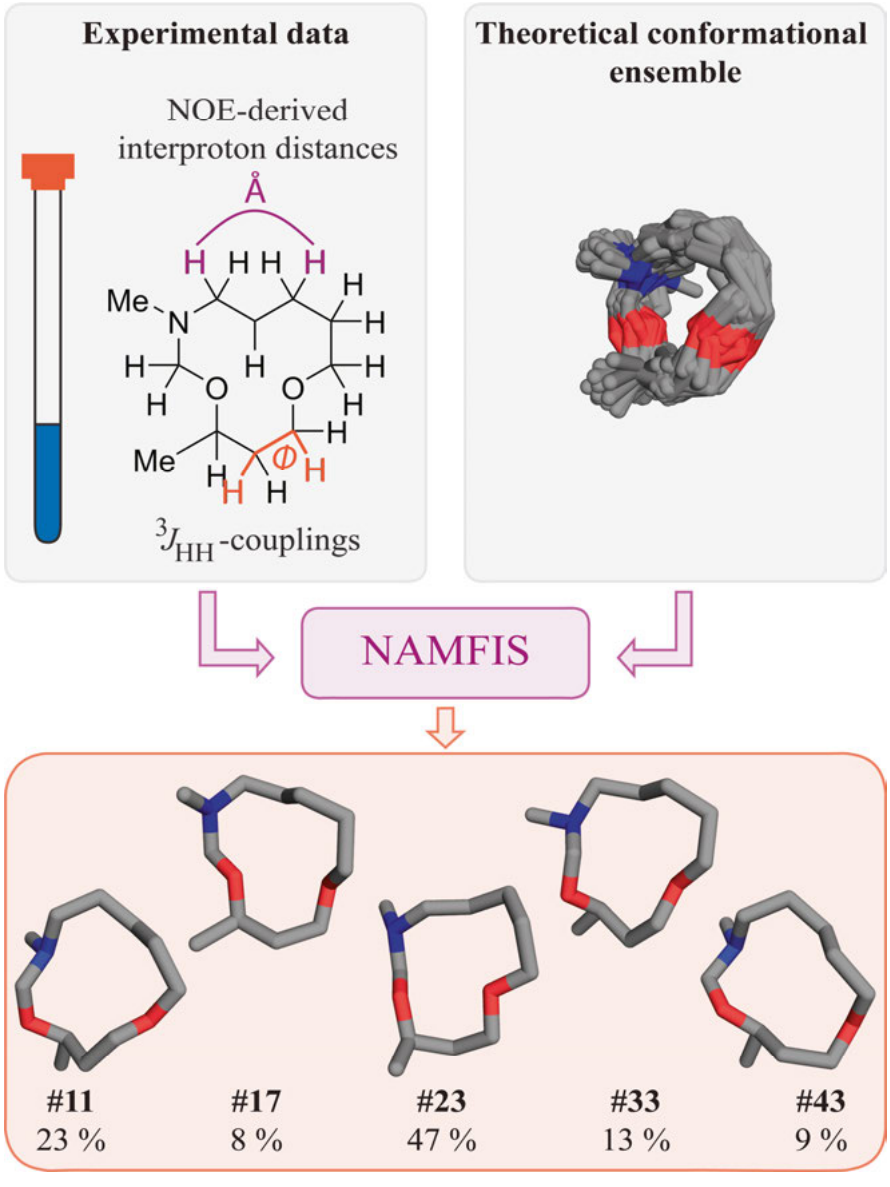


Figure 10. The work-flow for ensemble determination by NAMFIS.[17] Experimentally determined coupling constants and NOE-derived interproton distances are used for determination of the solution ensemble. Besides the experimental data, a theoretical ensemble is generated covering the entire feasible conformational space. The experimental and theoretical data are presented to the NAMFIS algorithm, which will show the best-fit solution in the form of a combination of population weighted conformations.

3.4 Results and discussion

Different chameleonic species. The solution ensembles of roxithromycin, telithromycin, spiramycin, rifampicin, daclatasvir, asunaprevir, simeprevir and atazanavir all varied between polar and apolar environments. For each conformation the SA 3D PSA and R_{gyr} was determined, in order to observe any trends between polar and apolar solvents. These two properties represent the size and polarity of the conformations.[129, 155] More closed conformations (smaller R_{gyr}) might have more intramolecular hydrogen bonds and therefore a lower SA 3D PSA. By looking at these two properties the compounds could be clustered into three groups: full molecular chameleons, partial molecular chameleons and non-chameleonic compounds.

Full molecular chameleons. Roxithromycin, telithromycin and spiramycin all show a significant difference in their size and polarity between solvents (Figure 11). Their conformational ensembles show a spread in R_{gyr} and SA 3D PSA and generally show a larger R_{gyr} and SA 3D PSA in a polar than in an apolar environment. Characteristic for these compounds is that their conformations show a variation in these two properties and the most populated conformation in a polar environment is distinctly separated from the most populated conformation in an apolar environment. The most populated conformation in a polar environment consistently has a larger R_{gyr} and SA 3D PSA than the most populated conformation in an apolar environment.

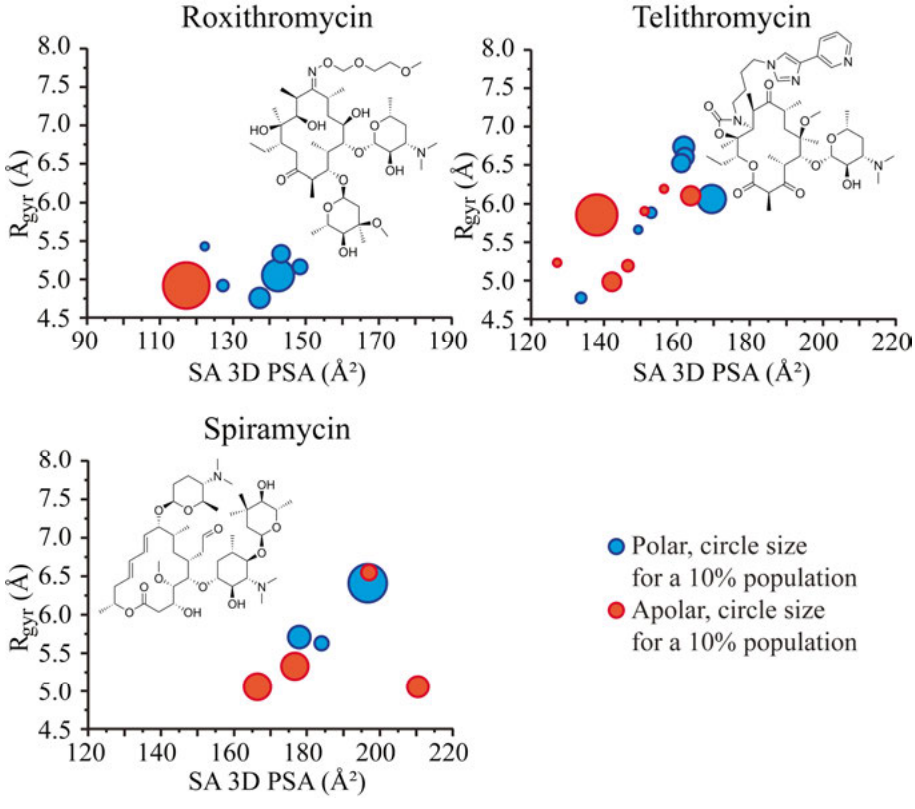


Figure 11. The R_{gyr} and SA 3D PSA for the full molecular chameleons roxithromycin, telithromycin and spiramycin. Full molecular chameleons display a spread in R_{gyr} and SA 3D PSA. Their conformational ensembles found in a polar environment are shifted towards larger R_{gyr} and SA 3D PSA values. The sizes of the circles are scaled according their populations. The size of a 10% population is indicated on the right.

Partial molecular chameleons. Despite the fact that simeprevir, asunaprevir and atazanavir show a large spread in R_{gyr} and SA 3D PSA in both polar and apolar environments (Figure 12), they cannot be classified as full molecular chameleons like roxithromycin, telithromycin and spiramycin. For simeprevir asunaprevir and atazanavir there is no distinction between the space the conformational ensembles in polar environment occupy as compared to those in an apolar environment. The most populated conformations from the two different solvents are often found to have a similar R_{gyr} and SA 3D PSA. However, the fact that these compounds show a large variation in conformations within a given solvent reveals adaptivity, which could contribute to improving solubility and cell permeability as compared to rigid, non-chameleonic compounds.

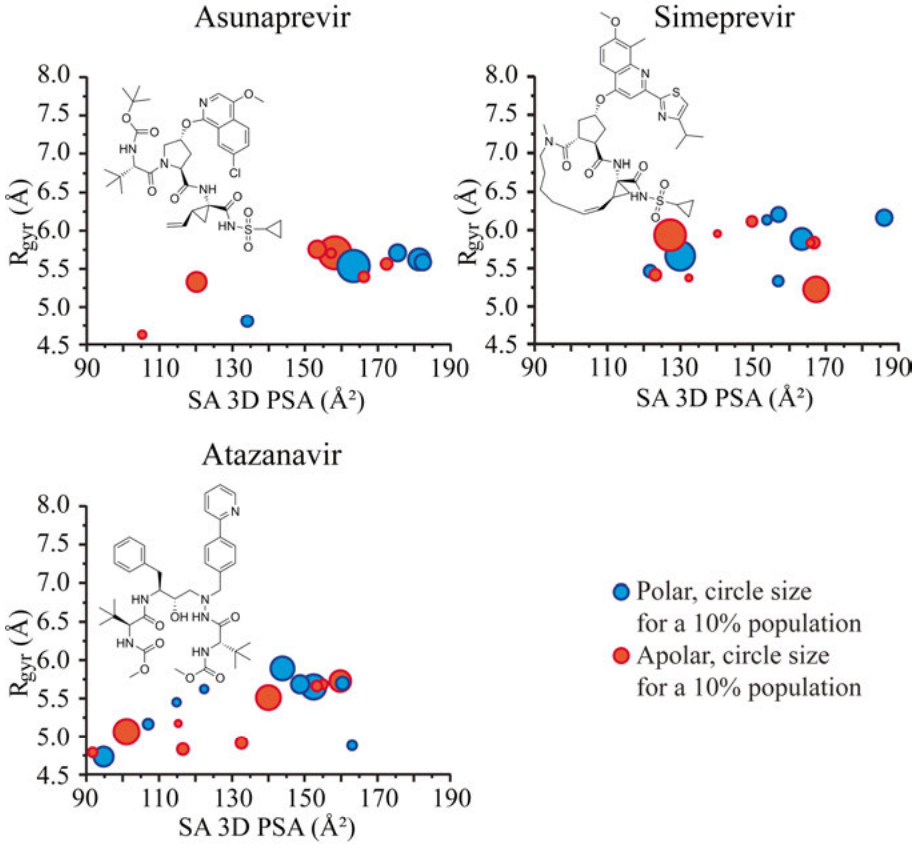


Figure 12. The R_{gyr} and SA 3D PSA for the partial molecular chameleons asunaprevir, simeprevir and atazanavir. Partial molecular chameleons display a spread in R_{gyr} and SA 3D PSA. However, the conformational ensemble found in a polar environment has a similar distribution in R_{gyr} and SA 3D PSA as the conformational ensemble found in an apolar environment. The sizes of the circles are scaled according their populations. The size of a 10% population is indicated on the right.

Non-chameleonic compounds. The last two compounds studied in Papers II and III are rifampicin and daclatasvir (Figure 13). These two compounds show a decreased degree of flexibility as displayed by the lack of spread in R_{gyr} and SA 3D PSA for their different conformations. In both solvents the R_{gyr} does not vary more than 1 Å and the SA 3D PSA generally varies within a limited range of about 10-20 Å^2. This is significantly smaller than the spread observed for the full and partial chameleons that varies roughly between 1.5 and 2.5 Å in R_{gyr} and between 40 and 90 Å^2 in SA 3D PSA. Furthermore, is there no clear separation between the space that conformations from a polar environment occupy as compared to the space occupied by those found in an apolar environment. Therefore, rifampicin and daclatasvir are considered to

not act as molecular chameleons, and are expected to rely on a different mechanism to permeate the cell-membrane.

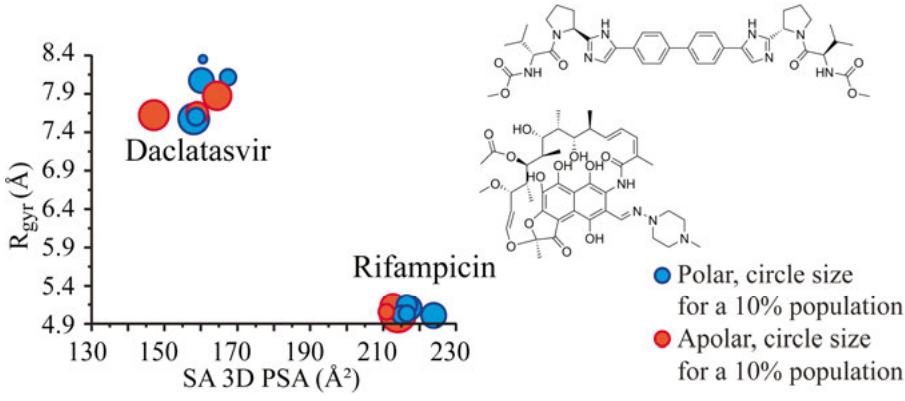


Figure 13. The R_{gyr} and SA 3D PSA for the non-chameleonic rifampicin and daclatasvir. Non-chameleonic compounds are characterised by solution ensembles that do not display a spread in R_{gyr} and SA 3D PSA, and by the fact that conformations found in a polar environment have similar R_{gyr} and SA 3D PSA values as those found in an apolar environment. The sizes of the circles are scaled according their populations. The size of a 10% population is indicated on the right.

3.5 Summary

Drugs that are able to bind difficult-to-drug targets are often non-compliant with the Ro5. Despite their non-compliance, a large portion of these drugs displays passive membrane permeability, which opens up the possibility to have bRo5 orally available drugs for difficult-to-drug targets. However, it is far from trivial to predict which bRo5 molecule is orally available. Since synthesising bRo5 compounds involves long and complicated synthetic routes, it is desirable to have a good understanding of what facilitates their passive membrane permeability.

Papers II and III investigate a series of orally available bRo5 drugs for their chameleonic properties. The ability to adjust their size (R_{gyr}) and polarity (SA 3D PSA) to the environment by adopting different conformations can explain the passive membrane permeability of six out of the eight investigated drugs. Three types of chameleonic behaviour were observed:

i. Full molecular chameleons
ii. Partial molecular chameleons
iii. Non-chameleonic compounds

i) Full molecular chameleons
Roxithromycin, telithromycin and spiramycin all behave like full molecular chameleons. This is characterised by a spread in R_{gyr} and SA 3D PSA for the different conformations and generally lower R_{gyr} and SA 3D PSA values for conformations found in an apolar environment than those found in a polar environment.

ii) Partial molecular chameleons
Much like the full molecular chameleons, partial molecular chameleons display a spread in R_{gyr} and SA 3D PSA for their various conformations. However, unlike full molecular chameleons, there is no distinct difference between the R_{gyr} and SA 3D PSA for the conformations found in different polarities. Asunaprevir, simeprevir and atazanavir were found to behave like partial molecular chameleons.

iii) Non-chameleonic compounds
Rifampicin and daclatasvir did not behave like molecular chameleons and were therefore classified as non-chameleonic. The conformations found for these two compounds did not display a large spread in R_{gyr} and SA 3D PSA and did not show a difference in R_{gyr} and SA 3D PSA between the two different solvents.

Being able to accurately predict membrane permeability of bRo5 compounds is expected to save time and money in the discovery of new (bRo5) drugs. Understanding the behaviour of this type of compounds in polar and apolar environments is the first step towards formulating predictive models. Obtaining insights into the different types of chameleonic behaviour is a key step in this process.

4 Structure elucidation of simeprevir by MicroED (Paper IV)

The previous section (section 3) explored the behaviour of bRo5 drugs in solution. These drugs are not only more challenging to synthesise than the conventional small molecule drugs, but also face difficulties for structure elucidation. Their flexibility complicates crystallisation, whereas conformational elucidation by NMR is a long and elaborate process. Paper IV described the elucidation by MicroED of simeprevir (Figure 9), a macrocyclic drug introduced in section 3. Furthermore, the implications it has for bRo5 drugs were discussed.

4.1 Macrocyclic drugs

A macrocycle is a compound with a cyclic core of 12 non-hydrogen atoms or more.[132, 133] Many macrocyclic drugs fall in the bRo5 chemical space, giving rise to increased flexibility in comparison to small molecule drugs. This degree of flexibility is, however, limited as compared to a linear equivalent as a result of cyclisation. The benefit of this is that it causes a reduced entropic penalty upon binding.[110, 132, 133] These drugs are often good at targeting poorly-defined binding-sites, such as protein-protein interaction, which seem to be inaccessible with conventional small molecules drugs.[110] Macrocycles have also shown promising results against rapidly mutating targets, as they retain their activity despite the mutations.[156] Cyclisation does not only lead to a reduced entropic penalty upon binding, but has also shown to increase the stability and binding specificity,[157-159] and can in some cases result in improved membrane permeability.[160, 161] Out of the roughly 80 approved macrocyclic drugs, 30 are orally available.[110]

Most of the orally available macrocyclic drugs are natural product derivatives. Simeprevir is a macrocyclic drug that is not based on a natural product. No small molecule XRD structure is available of simeprevir, but a co-crystal with its target, the HCV NS3/4a protease, is available.[157] Besides the target-bound structure, the NMR solution ensembles of simeprevir are available (Paper III). There are many (macrocyclic) bRo5 drugs with a similar scaffold to simeprevir that show promising activities against the HCV NS3/4a serine

protease and against the RNA dependent RNA polymerase of the SARS-CoV-2.[157] Few of these compounds have a small molecule XRD structure.

4.2 Structure elucidation by MicroED

MicroED is a cryo-EM method, where electron diffraction data are collected from small crystals, typically with a thickness in the nanometer range (Figure 14). MicroED has the potential to enable studies of new and important chemical and biological structures, which were previously beyond reach due to difficulties in growing the appropriate large and well-ordered crystals needed for X-ray diffraction. The biggest obstacle in streamlining the MicroED work-flow was not to account for the difference in wavelength and energy that comes with a different energy source (X-rays *versus* electrons) during data-analysis, but instead the way in which the data was recorded.[58] X-ray data is obtained by collecting high-speed diffraction patterns while continuously rotating the crystal in the beam that is mounted on the goniometer. This results in a wedge of data being recorded for every frame, something that is accounted for in the X-ray analysis software. The stage of the electron microscope can continuously rotate from -70˚ to +70˚, generating a wedge of up to 140˚ that can be sampled, which is enough to reach high-completeness data-sets.[162] The first MicroED protein structure that was solved, was recorded with still diffraction images by tilting the sample stage.[163] This means that the crystal is not rotated during data-recording, which complicated data-indexing with conventional X-ray software.[163] Improved cameras with high sensitivity and shutterless data-recording allowed for recording of low-dose (less than 0.1 $e^{-1}Å^{-1}s^{-1}$) diffraction movies,[164] enabling multiple exposures of a single crystal.[58] The speed of data read-out was another obstacle that was overcome with improved cameras. Newer cameras are developed with faster read-outs and higher sensitivity was achieved with the development of direct electron detectors.[164] Modern instruments are capable of fast tilt rotations of the stage and wedges of 0.5-1 ˚/frame are typically recorded.[164]

MicroED has been used to elucidate a wide variety of compounds. It was initially developed from 2D electron crystallography which was mostly used to study membrane proteins.[33, 57, 163, 165] Obtaining a membrane protein crystal structure is not at all trivial, but recent method developments within MicroED have smoothened the process of obtaining one.[166-168] Despite the focus of MicroED being on protein structures during its development, it is becoming increasingly popular for small molecules[35, 169-178] and non-biological (macromolecular) complexes.[179-188] Nearly all small-molecules and complexes elucidated by MicroED so far are on rather rigid systems. In many cases the powders obtained after synthesis (or obtained from a vendor) have a microcrystalline state that can be applied directly to the grid, meaning no

further crystallisation of the compound is required.[35, 169, 172, 189, 190] Elucidating traditional small molecules or rigid systems can be trivially easy, but the applicability of MicroED for larger small molecules, with increased flexibility, such as bRo5 compounds, is relatively unexplored. Especially for compounds that are hard to crystallise, MicroED can be advantageous, since it requires smaller and lower-quality crystals.

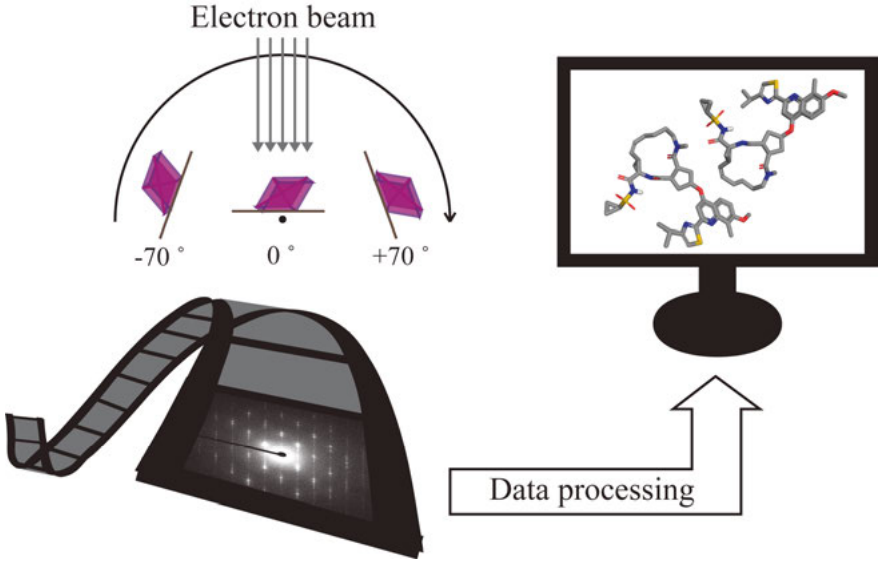


Figure 14. Tiny crystals of the object of interest are exposed to an electron beam while the object is continuously rotated inside the beam. This gives rise to a movie of changing diffraction patterns, due to the rotation, which can be analysed in a similar way to conventional X-ray diffraction data, resulting in a 3D structure of the studied object.[164]

4.3 Results and discussion

The structure of simeprevir was determined by MicroED. Microcrystals were obtained by dissolving simeprevir in minimal amounts of methanol and evaporating the solvent over 20 h at room temperature. The microcrystals appeared as needles with sub-micrometre thickness, but could reach lengths of several µm (Figure 15).

Data of simeprevir were collected on a 200 keV FEI Talos Arctica microscope as 60˚ wedges on a Falcon III camera. The diffraction patterns were recorded under continuous rotation at a tilt-speed of ~0.6˚/s and integration times of 2 s per frame. In total 8 data-sets were processed in XDS, merged in XSCALE and the final structure was solved with SHELX.[191]

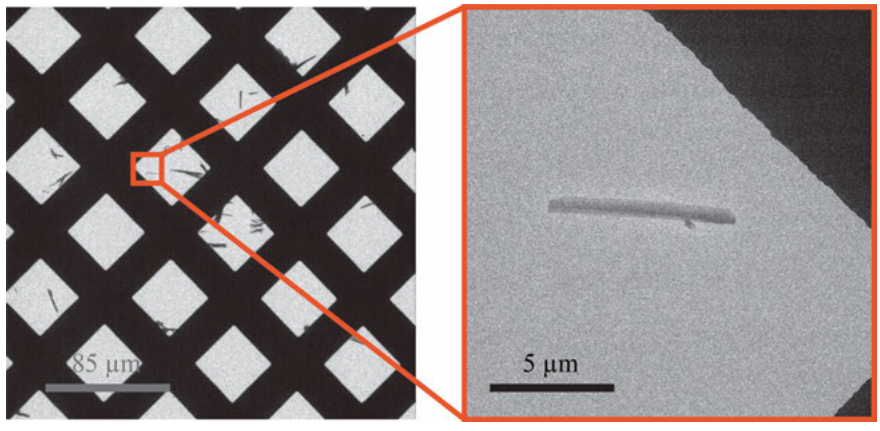


Figure 15. Grid of simeprevir at 200x (left) and 3400x (right), showing small needles dispersed over the carbon film of the copper grid.

The structure of simeprevir showed two near-identical conformations in the asymmetric unit (Figure 16). The RMSD (root-mean-square deviation) value between the heavy atoms of the macrocyclic core of the two MicroED conformations was 0.196 Å, with the side-chains pointing in the same directions with a similar conformation.

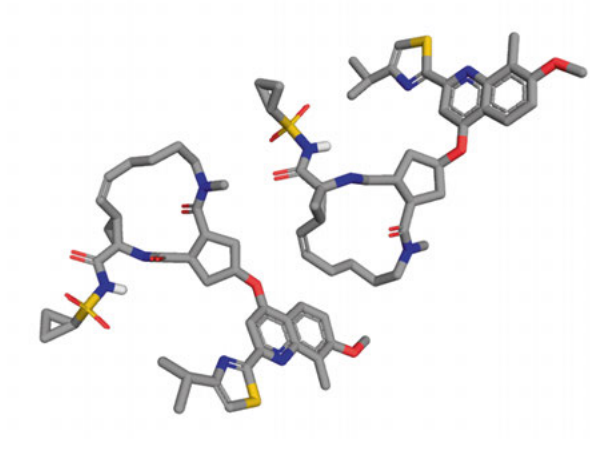


Figure 16. The MicroED structure of the asymmetric unit of simeprevir. Non-polar hydrogens are omitted for clarity.

The two MicroED conformations (MicroED_01 and MicroED_02) of simeprevir are not only similar to one and another, but the macrocyclic cores are similar to the cores of the target-bound conformations that were obtained by X-ray crystallography[157] (Table 3, Figure 17). The open and rather flat conformations observed for the MicroED and target-bound structures are

likely a result of crystal packing and the fact that the HCV serine protease has a shallow and rather exposed binding-site for simeprevir, respectively. In contrast, the solution ensembles (Paper III), display a varied set of conformations, with no clear difference in overall size and solvent exposed polar moieties between the two investigated solvents. The solution conformations are more folded than the solid-state conformations and adopt rather dissimilar core conformations (Paper III).

Table 3. RMSD values (Å) between the macrocyclic core of the two MicroED conformations and the target-bound (PDB code 3KEE)[157] and solution ensemble conformations (Paper III).

	MicroED_01	**MicroED_02**
MicroED_02	0.196	
PDB: 3KEE chain I	0.217	0.349
PDB: 3KEE chain K	0.250	0.383
PDB: 3KEE chain M	0.239	0.381
PDB: 3KEE chain P	0.264	0.406
DMSO_01	0.613	0.666
DMSO_02	0.636	0.683
DMSO_03	0.672	0.697
DMSO_04	0.639	0.641
DMSO_05	0.634	0.679
DMSO_06	0.524	0.542
DMSO_07	0.703	0.663
$CDCl_3$_08	0.760	0.731
$CDCl_3$_09	0.585	0.582
$CDCl_3$_10	0.666	0.677
$CDCl_3$_11	0.494	0.489
$CDCl_3$_12	0.447	0.446
$CDCl_3$_13	0.757	0.795
$CDCl_3$_14	0.761	0.800
$CDCl_3$_15	0.822	0.881

Superposition of the macrocyclic cores of the MicroED conformations to the most similar solution structures 11 and 12 (Figure 17) reveals a similar overall conformation with a local dissimilarity in the core conformation. The need for an open and flat conformation in the solid-state structures seems to force the macrocyclic core in this specific conformation.

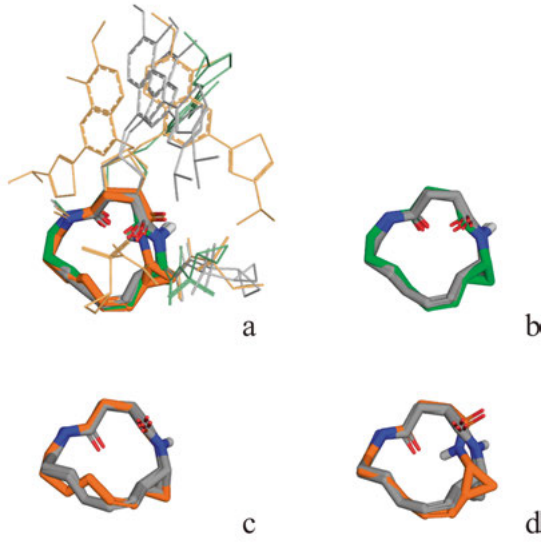


Figure 17. The MicroED conformations (grey structures), superimposed on the macrocyclic core (atoms highlighted in bold) with the target-bound structure (PDB: 3KEE chain I, in green)[157] and the two most similar solution ensemble structures (in orange); a) overlay of all structures, with the side-chains displayed as sticks, b) overlay of the macrocyclic cores of the MicroED conformations and the target-bound structure, c) superposition of the macrocyclic core of the MicroED structures and solution conformer 11 and d) solution conformer 12 (Paper III).

4.4 Summary

The MicroED structure of simeprevir has been solved with relative ease, whereas this compound has not been solved by XRD. Obtaining the structure of simeprevir was, however, significantly less straightforward than for smaller compounds, such as ibuprofen and brucine.[35] The main challenge was faced with the generation of the microcrystalline structure, which required slow evaporation from methanol.

The conformations obtained by MicroED were similar to the target-bound structure,[157] showing near-identical core conformations. All MicroED and target-bound conformations adopted an open and rather flat structure, either as a result of crystal packing or of binding to the shallow binding pocket of the serine protease. The solution structures obtained by NMR (Paper III), adopted more folded and varied conformations. Two solution structures displayed similar core conformation where only two heavy atoms oriented differently.

The high accessibility of electron microscopes and the relative ease of sample preparation make MicroED suitable to elucidate the structures of bRo5 drugs.

5 No NUS for NOESY (Paper V)

Section 3 dealt with ensemble determination of drugs in the bRo5 space (Papers II and III). These analyses are elaborate and rely on accurate NOE data. To ensure that high-quality data is obtained, a series of time-consuming NOE experiments are typically acquired, which take days of spectrometer time. This section explored the possibility to record quantitative NOESY spectra using non-uniform sampling (NUS) to shorten data-acquisition.

5.1 Non-uniform sampling

Multi-dimensional NMR experiments, such as NOESY, often require a lot of spectrometer time to yield a spectrum of decent resolution with a proper signal-to-noise ratio (S/N). The more dimensions that are being sampled, the longer the experiment takes. In order to obtain an NMR signal, the recorded free induction decay (FID) has to fulfil the Nyquist theorem for proper sampling of the wave function.[192] Without fulfilment of the Nyquist theorem, the signals recorded on a time domain cannot be converted into their corresponding frequency domain by Fourier transformation (FT). Conventional sampling schemes for multi-dimensional experiments consist of incremental stepwise collection of the data points.[193] Data collection of the direct dimension is a quick process in comparison to the other, indirect dimensions. One method to speed up data-collection is by making use of NUS for the indirect dimension(s) of the experiment.[194] NUS collected data cannot be directly subjected to FT to give rise to the signals in the frequency domain. Instead, the wave-function has to be reconstructed prior to FT. Collecting data with varying increment steps has been suggested in the late 1980s,[194-196] but has only become routine in the past ten years.[197-199] One of the first reported ways of implementing NUS was to sample the indirect dimension exponentially, where more data-points are collected in the high S/N regime of the FID and fewer at the lower S/N part of the FID.[194] The maximum entropy (ME) method was used for reconstruction of the indirect dimension.[194]. Several programs have been developed for the reconstruction of the data, MDD (Multidimensional Decomposition)[200] and ME[194] were amongst the firsts. Later IST (Iterative Soft Thresholding),[62, 201-203] SCRUB (Scrupulous Cleaning to Remove Unwanted Baseline Artifact)[53] and NESTA-NMR

(NESTerov's Algorithm)[204] have been developed and all five are used for the reconstruction of FIDs. Spectra recorded with 50% or 25% NUS use half or a quarter of the time a uniformly sampled experiment would need. Implementation of NUS schemes have become a standard way of recording spectra for macromolecular NMR studies and has contributed to the applicability of NMR experiments on biological samples.[48, 67, 95] The wide usage of NUS schemes in biomolecular NMR stands in shrill contrast to NMR experiments used in organic chemistry, where application of NUS is an outlier rather than routine.[205-207] NUS has been shown to be a powerful time-saving tool with little-to-no downsides on some occasions. However, implementation of NUS gives rise to observable artifacts for experiments that deal with low-intensity signals.[208] This is typically the case for NOESY experiments, where the most informative long-range NOEs display the lowest intensity signals. These long-range interactions are often key in interpreting 3D information of the studied object.

5.2 Accuracy of the NOE-derived distances

The accuracy of an NOE-derived interproton distance, derived from a uniformly sampled spectrum, is dependent on multiple factors. Larger interproton distances give rise to weaker signals, resulting in a lower signal-to-noise ratio.[46, 48, 65, 209] The accuracy does not only depend on the interproton distance but also on the solvent viscosity,[46] the mixing time and the type of experiment that has been selected.[48] Transient NOE experiments are most reliably used for quantitative purposes. Steady-state experiments cannot be used for distance determination, because three-spin effects (spin diffusion) are unavoidable.[64] Signal overlap in the 1D trace can, but does not necessarily have to, lead to overlap of the NOE cross-peaks. Since every NOE cross-peak has a unique build-up rate, troublesome overlap can be identified by looking at the linearity of the initial part of the NOE build-up curve (Figure 18).[24] Cross-peaks suffering from signal overlap are generally less linear and should be excluded from further analyses.

In order to assess the linearity of an NOE build-up curve, multiple NOESY spectra with varying mixing times are required.[143] Plotting of the NOE build-up curve allows for several quality checks. One can ensure that only data-points that were acquired in the linear regime of the NOE build-up curve are included in the analysis. By looking at the coefficient of determination (R^2),[210, 211] the amount of noise to the data-points can be quantified. The NOE build-up curves occasionally show an offset from the origin of the graph. As a result, any single data-point on the curve will have an over- or under-estimated peak-intensity. However, the slopes of the curves with and without offset are identical (Figure 18).

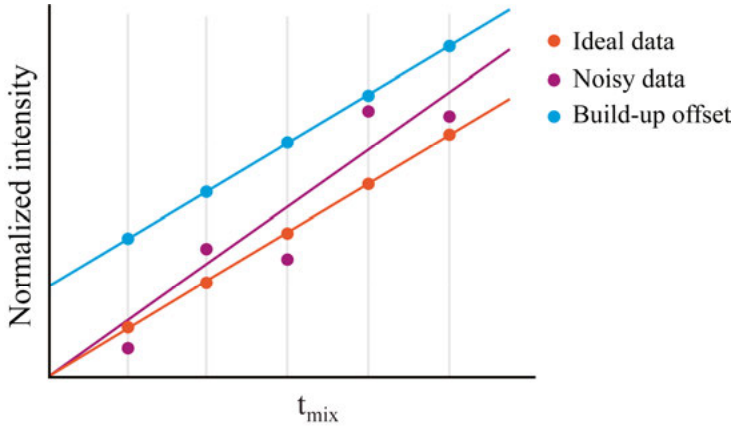


Figure 18. Representative build-up curves with their linear regressions, highlighting common errors. The ideal build-up is plotted in orange, whereas noisy data (in purple) gives rise to less reliable data and can lead to slope errors. An offset from the origin (in blue) can be observed when there is background intensity that is independent of the mixing time. Build-up curves with an offset from the origin can still be used for quantification since the slope is not affected by the offset.

One final point to consider is the pulse homogeneity along the spectral window of the acquired data. ROESY (rotating frame nuclear Overhauser effect spectroscopy) experiments have the advantage that they lack a zero-cross over point, and will always show a correlation if there is a ROE.[63] However, ROESY spectra make use of a spin-lock pulse, introducing pulse inhomogeneity along the spectral window.[212] This inhomogeneity can be accounted for, but does introduce an additional source of error.[213] Another seemingly advantageous development resulting in pulse inhomogeneity affecting the spectrum in a frequency-dependent way is solvent suppression.[214-218] Solvent suppression and ROESY sequences can be used in a quantitative way, but should be used with care, keeping potential artifacts in mind.

5.3 Accuracy of NUS recorded NOE-data

In a recent study, the effect of NUS on 2D-NOESY spectra for the model compound strychnine has been studied for five proton-proton interactions.[219] The interproton distances of these five correlations corresponding to distances up to 2.60 Å from spectra sampled with as low as 17% NUS were investigated.[219] Despite this being a good starting point for the evaluation of quantitativity of NUS recorded NOESY spectra, strychnine may not be a representative model compound for most compounds where solution structures are of interest. Strychnine is a rigid molecule, whereas most drugs

and biomolecules show some degree of flexibility. Besides the lack of flexibility of strychnine, this pioneering study has only looked at short distances, which falls within the distance range where one expects an NOE signal and therefore it does not give new information about the 3D structure of a studied compound. A major complication the authors reported with the implementation of NUS was the increased t1-noise.[219]

Many of the quantitative NOE experiments are performed on biological macromolecules.[209] These macromolecules show a higher degree of flexibility than strychnine even in their most ordered form.[48] On top of the larger degree of flexibility of macromolecules, these large structures can show valuable long-distance correlations, which give insights into inter-strand relationships, protein-ligand interactions or protein-protein interactions.

Quantitative NOE experiments are also often used for medium-sized organic compounds such as peptides,[220-222] macrocycles[131, 223] and PROTACs[224, 225] with a mass of a couple of hundred Da. These compounds are known to display a certain degree of flexibility, and their 3D conformations play a key role in understanding their bioactivities. Their conformations can be elucidated with the help of quantitative NOE experiments (Section 3). Figure 19 shows the structures and solution ensembles of strychnine and spiramycin (Paper III).

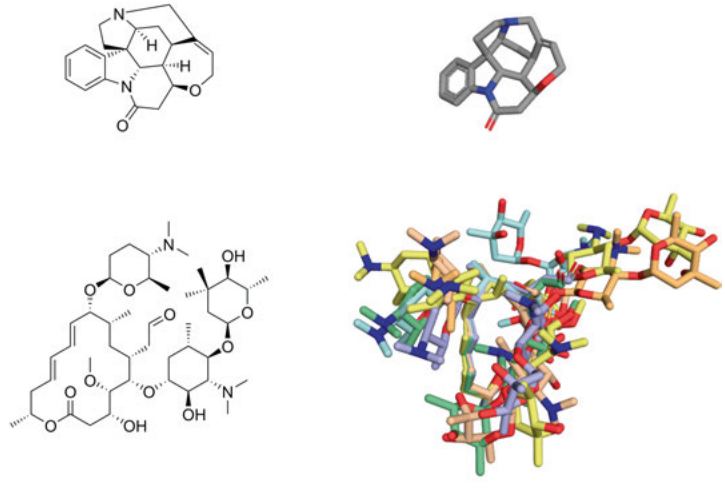


Figure 19. Structures of strychnine (on top) and spiramycin (below) and their corresponding solution conformations to the right. The major conformer (97.5%) of strychnine and the conformers of spiramycin in solution (Paper II) are displayed, hydrogens are omitted for clarity.[226, 227]

5.4 Results and discussion

In this study, the quantitativity of NOESY spectra recorded on spiramycin (Section 3, Figure 19) in DMSO-d_6 were determined under different conditions. The most accurate data were derived from the uniformly sampled build-up series recorded with mixing times of 100-700 ms (100 ms increments). The normalised intensities and interproton distances from this series were used as a reference for the evaluation of the quality of the data acquired with NUS. Several aspects were investigated: sampling density, reconstruction algorithm, relaxation delay and number of sampled points. The multi-dimensional decomposition (MDD)[200] and modified iterative soft thresholding (MIST),[197] as implemented in MestReNova, were investigated. Increasing the relaxation delay (allowing for full recovery of the magnetisation vector) and the number of sampled points (by recording the same number of NUS increments as for the uniformly sampled data, aiming to gain resolution instead of saving time) did not significantly improve NUS recorded data. A final analysis was to evaluate the number of interproton distances and their errors, as compared to the reference data-set, from a single NOESY spectrum recorded with a mixing time of 500 ms. This last analysis was done to determine the importance of recording an entire build-up series. Recording just one data-point resulted in larger errors (because the effect of outliers will not be reduced by inclusion of more data, Figure 18) and inclusion of artifacts that were mistakenly treated as true NOE correlations.

The two most significant variables investigated were the sampling density and the reconstruction algorithm used. The uniformly sampled build-up series was analysed first, resulting in 62 correlations that showed linear build-up curves (with $R^2 \geq 0.95$). Build-up series were obtained for data recorded with 75, 50 and 25 % NUS and each series was reconstructed twice: once with MIST and once with MDD in parallel. For the same 62 correlations, build-up curves were generated and a first comparison was made. For all six NUS reconstructed data-sets, fewer correlations showed a linear build-up curve as compared to the uniformly sampled data (Table 4).

The number of correlations with a linear build-up decreased as the sampling density decreased. This was observed for interproton correlations of all distances, but was more pronounced for long-range than short-range interactions. The implications of this discrepancy are that long-range interactions are affected by implementation of NUS more severely than short-range interactions. Furthermore, Table 4 shows that fewer correlations have a linear build-up with MDD reconstructed data than with MIST reconstructed data.

Table 4. Number of correlation with a linear build-up ($R^2 \geq 0.95$) for the uniformly sampled data and the NUS reconstructed data-sets. In total 62 correlations were analysed which corresponded to interproton distances of various lengths. The number of correlations belonging to each distances group are given for the different data-sets.

Distance range (Å)	Uniformly sampled	75 % NUS		50% NUS		25% NUS	
		MIST	MDD	MIST	MDD	MIST	MDD
All distances	62	37	21	28	9	28	0
0.00-2.49	12	11	6	8	4	8	0
2.50-2.99	27	19	10	11	4	16	0
3.00-3.49	10	4	2	6	1	2	0
3.50-3.99	8	2	3	3	0	2	0
4.00-5.00	5	1	0	0	0	0	0

Assessing the linearity of the build-up curves only gave insight into how many correlations could be used for interproton distance determination based on build-up rates. For structural studies, non-linear build-up curves would be excluded from further analysis. Therefore, in order to assess the accuracy of the data, interproton distances were only determined for correlations with a linear build-up curve. The obtained distances were compared to the reference data-set and their errors were determined (Figure 20).

Despite the build-up curves showing linear correlations, the average absolute errors increased as the sampling density decreased. This observation was consistent for all distance ranges. The average absolute error was larger for long-range distances than for short-range distances. This meant that there were not only fewer long-range distances observed with NUS recorded data, but the long-range distances were also less accurate. The MDD-reconstructed data showed fewer linear build-ups and larger errors on the corresponding interproton distances than the MIST reconstructed data.

Analysis of the spectra recorded with a mixing time of 500 ms for each data-set revealed the same trends as were visible for the build-up series: the average absolute error increased as the sampling density decreased and the effect was more pronounced for long-range than for short-range correlations. Since the linearity of the build-up curve could not serve as a quality check with data coming from a single NOESY spectrum, more correlations were evaluated. Only correlations showing the opposite phase were excluded from the analysis. The error calculated for the interproton distances were consistently larger for the distances coming from a single NOESY spectrum than for those calculated from the corresponding build-up series. Interproton distances calculated from the uniformly sampled NOESY spectrum with a mixing time

of 500 ms deviated less than 6.9% for all distance ranges, but was larger for long-range correlations than for short-range correlations.

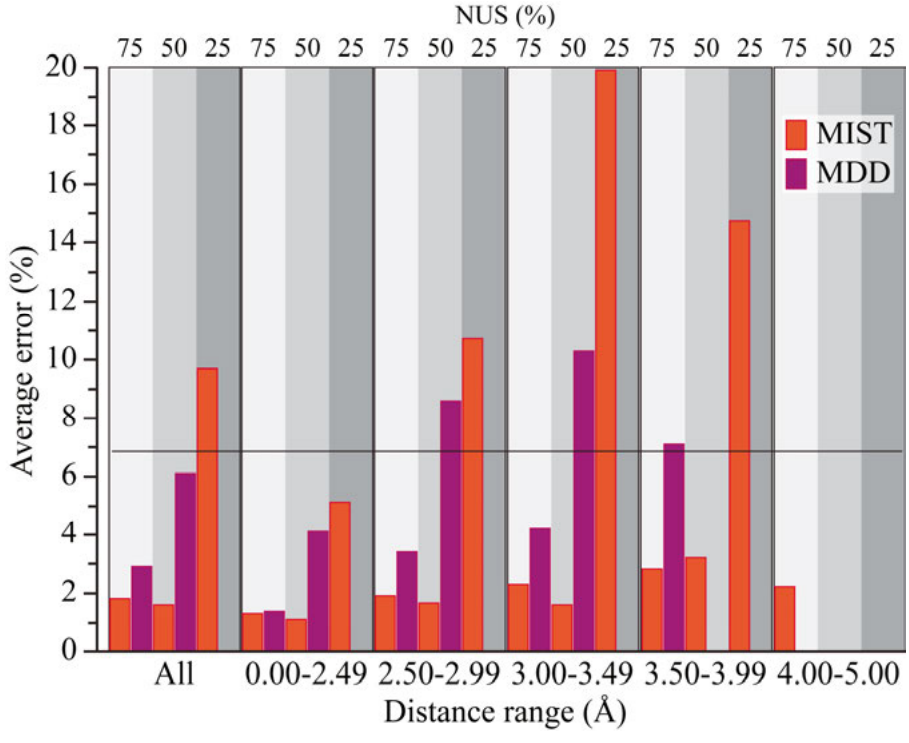


Figure 20. The average absolute error on the distances for each data-set. The average error for all correlations within a data-set are given as well as the correlations grouped on distance range (bottom axis) and sampling density (top axis). The line at 6.9 % corresponds to the average error observed for interproton distance determination of strychnine from 1D NOE-data in DMSO-d_6.[46] Only distances with a linear build-up curve (Table 4) are included in this plot, no linear correlations were found for the 25% NUS MDD reconstructed data-set.

5.5 Summary

Paper V investigated the accuracy of NUS recorded NOE-data. Four different sampling densities (i), 25% NUS, 50% NUS, 75% NUS and uniformly sampled (100% NUS) data, were recorded. For all data-sets build-up curves with mixing times from 100 ms to 700 ms, with 100 ms steps in between, were recorded. Distances were determined based on the build-up rate and the normalised intensity obtained from the spectrum recorded at 500 ms (ii). Two different reconstruction algorithms (iii), the modified iterative soft thresholding (MIST) and the multi-dimensional decomposition (MDD), were investigated in order to assess whether the reconstruction algorithm plays an

important role for the accuracy. Different relaxation delays (iv), 1.2×T1 and 5×T1 were employed to see whether the relaxation delay has an influence on NUS recorded NOE-data. Lastly, the number of sampled points was investigated (v). Spectra recorded with the same number of sampled points in the indirect dimension, but recorded with NUS affected the data-quality. A uniformly sampled spectrum with 512 sampled points was compared to a 50% NUS spectrum recorded with 1024 points and a 25% NUS spectrum recorded with 2048 points (both NUS spectra had 512 sampled points). The uniformly sampled spectrum gave the best results.

i) Sampling density
The sparser the sampling, the fewer number of correlations were observed and the less accurate the interproton distances were. This result was consistent even if other parameters were varied.

ii) Build-up rate *versus* single NOESY spectrum
Under none of the investigated conditions did data coming from a single NOESY spectrum give equally good or better quality results than those obtained from uniformly sampled build-up curves. However, a uniformly sampled single NOESY spectrum gives more reliable results than data derived from build-up rates but recorded with any degree of NUS.

iii) Reconstruction algorithm
Much better results were obtained with the MIST reconstruction algorithm than with MDD. However, neither of the algorithms result in data-quality that can reliably be used for quantitative assessment of NOE-data.

iv) Relaxation delay
No significant difference in accuracy was observed for data recorded with 5×T1 and with 1.2×T1.

v) Number of sampled points
An increased number of sampled points did not lead to more accurate NUS reconstructed NOE-data. Despite having a larger number of points in the indirect dimension after NUS reconstruction, the data-accuracy was significantly lower than the uniformly sample data.

NOE data with the highest accuracy comes from a series of uniformly sampled NOE experiments. Under none of the investigated conditions did NUS recorded data match or exceed the accuracy of uniformly sampled data. Out of all the correlations with a linear build-up curve for the uniformly sampled data, only 60% or less of the correlations still showed a linear build-up for the NUS data-set. The MIST algorithm was capable of reconstructing the remaining correlation recorded with 75% NUS to an acceptable quality, but

only for distances up to 3.00 Å. For the 50% NUS reconstructed MIST data, the quality of the correlations with a linear build-up was only sufficient for distances up to 2.5 Å. Any longer-range correlations could no longer reliably be established. Given that the long-range interaction are the most informative and the time-gain with 50% or 75% NUS is limited, implementation of NUS for quantitative NOE studies is not recommended. One way of saving time is by reducing the relaxation delay to 1.2×T1. If time is truly the limiting factor, recording fewer NOESY spectra is recommended over the implementation of NUS.

6 Concluding remarks and future perspectives

This thesis describes the use of various structure elucidation strategies to answer a scope of research questions.

The findings of section 2 (Paper I) open up for the study of new inhibitor candidates for VIM-2 using solution-state NMR, by reporting the backbone resonance assignments of this bacterial enzyme. Given the clinical relevance of VIM-2 in antibiotic resistance, the development of clinically applicable inhibitors is much-desired.

Difficult-to-drug targets are often best modulated using inhibitors that occupy the beyond rule of 5 chemical space. It is challenging to design these kinds of drugs, so that they are simultaneously membrane permeable and water soluble. Section 3 (Papers II and III) determines the solution ensembles of eight beyond rule of 5 drugs, laying an experimental ground for understanding the mechanisms behind their passive membrane permeability. Understanding the behaviour of these compounds will aid future drug design by expanding the chemical space that medicinal chemists have at their disposal.

Section 4 (Paper IV) described the structure elucidation of simeprevir, a large and flexible compound, which occupies the beyond rule of 5 chemical space, by MicroED. It showcases the ability of MicroED in obtaining solid-state structures of flexible organic compounds. Being able to solve the structures of molecules that don't grow crystals large enough for traditional X-ray crystallography will enable medicinal chemists to explore a larger scope of compounds more easily. Ultimately, this will lead to more diverse drug design.

Section 5 (Paper V) emphasises the importance of careful parameter selection for data-collection of quantitative NOESY. It shows that non-uniform sampling schemes, as they exist today, are incompatible with NOE-based interproton distance determination. This serves as a cautionary tale for those who aim to perform quantitative NOE experiments.

Obtaining structural information of organic compounds and biological macromolecules at atomic resolution has for a long time been limited to solid-state experiments. These experiments require large amounts of sample and

lack information about the dynamics of the system. The obtained conformation(s) might not be physiologically relevant, because crystallisation can distort the conformation to facilitate packing. Despite the fact that crystal structures lack dynamic information, they are useful for research questions where physiological relevance is not the goal. Until there is a method that can quickly and reliably solve the 3D structure of compounds in their non-crystalline form, there will be a need for methods like X-ray diffraction.

MicroED is a solid-state method that offers some advantages over X-ray diffraction. For a start, the equipment needed for structure elucidation by MicroED is much more accessible than the few synchrotrons available for X-ray diffraction of biological macromolecules. Furthermore, an electron microscope that can record MicroED data is capable of recording data for any type of compound. No distinction is made between biological samples and small organic compounds. The fact that MicroED makes use of X-ray diffraction software for data processing, allows for analysis of the data coming from either energy source with already established software. Despite major developments still taking place, its advantages are already clearly visible. MicroED has been shown to be useful in all fields of chemistry.

Being able to obtain solution-state structures allows for studying physiologically relevant conformations as well as looking at the dynamics of a given system. Since the 1990s, NMR spectroscopists have been aiming to develop software that can elucidate conformations of (medically relevant) compounds in solution. Unfortunately, until the current day, the types of NMR experiments and corresponding analyses giving conformational information are only accessible to experienced users. The main obstacle for an average chemist to perform these types of analyses is the lack in understanding of the parameters needed to obtain quantitative data, and the user unfriendliness of the deconvolution software. Even with NMR data deconvolution software becoming more accessible, and user-friendly, obtaining sensible results relies on the recording of accurate data.

X-ray crystallography is the most routinely used method to obtain structural information. X-ray facilities are available in most parts of the developed world. These platforms help with data-collection and processing, facilitating easy elucidation of a crystal structure. X-ray crystallography has been around for many decades, making it a well-established method and its existence is known to every chemist. Other methods, such as NMR and cryo-EM are still largely performed by specialists and lack big user platforms. It is often unclear what these methods can offer, since both of them provide more than a single type of experiment.

All structural methods are complementary to each other, rather than in competition with one and another. This thesis uses methods other than X-ray diffraction and shows the large variety of problems that can be addressed by alternative techniques to obtain structural information.

7 Sammanfattning på svenska

Struktur i läkemedelsutveckling

Läkemedelsutveckling är en lång och komplicerad process där hänsyn behöver tas till en rad olika faktorer. Ett läkemedel måste ha tillräcklig effekt samtidigt som det bör uppvisa minimala biverkningar. Läkemedel fungerar genom att binda till specifika molekyler i kroppen. De flesta läkemedel binder till proteiner, men det finns även läkemedel som binder till DNA- eller RNA-molekyler. Dessa kallas också biologiska makromolekyler. Varje biologisk makromolekyl har en specifik funktion, men funktionen kan påverkas av interaktioner med andra ämnen såsom läkemedel.

För att förstå hur ett läkemedel binder till en biologisk makromolekyl krävs en noggrann analys av interaktionerna mellan dem. Att få en tredimensionell bild av läkemedlet när det är bundet till den biologiska makromolekylen är oerhört användbart för att förstå vilka faktorer som gör att läkemedlet binder till makromolekylen och hur läkemedlet kan förändras för att bindningen ska blir starkare och mer specifik. Starkare bindning betyder starkare effekt, och en ökad specificitet i bindningen kan minska biverkningarna. Det finns flera olika metoder för att studera bindning. Denna avhandling fokuserar på ett läkemedel som binder till ett protein som ger upphov till antibiotikaresistens. Metoden som har använts för att bestämma bindningsstyrkan och bindningsplats är känd som kärnmagnetisk resonansspektroskopi.

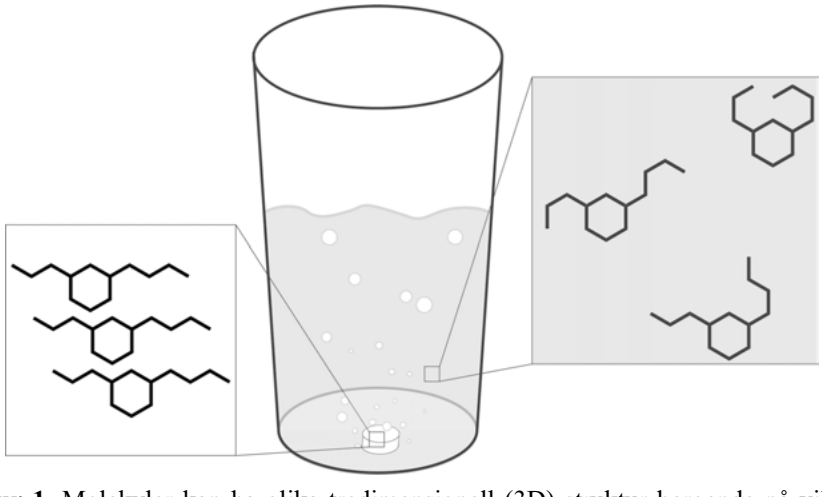
Innan bindning mellan läkemedel och makromolekyl kan studeras måste läkemedlet först tillverkas och dess struktur bestämmas. Läkemedelsmolekyler är för små för att kunna observeras med ett vanligt mikroskop, så det behövs alternativa analysmetoder där strukturen kan bestämmas på molekylär nivå. I denna avhandling beskrivs en metod som kan användas för att detektera 3D-strukturer av nya molekyler. Metoden är relativt ny och är värdefull för framtida läkemedelsutveckling.

Utöver att binda till specifika makromolekyler behöver läkemedelsubstanser också nå fram till dem för att ha effekt. Biologiska makromolekyler finns oftast inuti kroppens celler, vilket medför att läkemedelsmolekylerna måste passera cellmembranet för att kunna binda till dem. Små molekyler kan ofta

passera cellmembranet relativt enkelt, men för större molekyler kan denna passage vara en utmaning. Det finns dock stora läkemedelsmolekyler som kan passera cellmembran tvärtemot alla förväntningar, och det är av stort intresse att förstå hur det går till eftersom det kan ge information om hur nya och bättre läkemedel kan designas i framtiden. I denna avhandling har 3D-strukturer för ett antal större läkemedelsmolekyler i miljöer som efterliknar cellmembran jämförts med strukturer som de har i miljöer som liknar miljön inne i cellerna. Läkemedelsmolekyler som har en mer flexibel struktur verkar vara bättre lämpade för att passera cellmembran än molekyler som är mindre flexibla, vilket är en viktig insikt för framtida läkemedelsutveckling.

Den typ av experiment som användes i studien av hur läkemedel passerar genom cellmembran var mycket tidskrävande och det krävdes flera dagar för att samla in data. Sista delen av denna avhandling utreder om datainsamlingen kan påskyndas. När datainsamlingen utfördes snabbare blev resultatet att 3D-strukturerna inte kunde fastställas med tillräcklig noggrannhet.

Arbetet som presenteras i denna avhandling utforskar olika metoder för att undersöka 3D-strukturer i samband med läkemedelsdesign. Avhandlingen visar på hur viktigt det är att kunna undersöka molekyler med olika metoder och även vikten av strukturbestämning vid läkemedelsutveckling.



Figur 1. Molekyler kan ha olika tredimensionell (3D) struktur beroende på vilken miljö de befinner sig i. I figuren visas ett exempel: Miljön i en tablett är helt olik den som molekylen möter när den löses upp i vatten. I tabletten (vänster ruta) är flexibiliteten lägre, i lösning (höger ruta) är den högre och molekylerna kan förekomma i fler konformationer.

8 Acknowledgements

Máté, I have thoroughly enjoyed working with you (and still do). We seem to communicate well and I have always felt that no question was too stupid to ask (and I know that I have asked many stupid questions). There are plenty of things I can write here, explaining why working with you has been such a positive experience, but that will make this last bit long and tedious to read, and I feel like you know it either way, so there is no need for it. I am still trying to figure how many apple pies I need to bake to express my gratitude.

Adolf, you still are and will remain the NMR master to me. Your two index fingers always managed to tame the spectrometers when they misbehaved. You gave me the unique opportunity to learn some basic tips and tricks on how to fix malfunctioning NMRs. Thank you for your support over the past couple of years.

Hanna, you have been around mostly during the beginning of my PhD. You were someone I could talk to about science, but also about personal issues and were my go-to person for practicing my Swedish early on. Your kindness and openness are something I will remember.

Jan, we started working together even before I started my PhD and have been keeping this up essentially until the current day. Thank you for teaching me everything I know about macrocycles and most, if not everything, I know about medicinal chemistry. The biggest section (section 3) of this thesis is dedicated to our work together, which is telling for the contribution you had to my PhD education. Thank you for your guidance, patience and kindness.

Emma, thank you for training me during my master's and again at the end of my PhD. You have not only taught me NAMFIS and MicroED, but have also given me ample of advice about the "next step". Especially during my visit to UCLA our chats about postdoc grant applications have been very insightful. Thank you for everything.

For the VIM-2 project, I would like to thank **Jonathan**, **Albin**, **Scott**, **Anna**, **Lotta**, **Bono** and **Wolfgang**. The molecular chameleon project was a large collaborative project, thank you **Stefan**, **Yoseph**, **Emma**, **Nathan**, and **Sofie**. The MicroED part of my thesis would have not been there without **Emma**,

Guanhong and **Tamir**, thank you for welcoming me into your group and teaching me a new method. For the NUS project I would like to thank **Adolf** and **Maxim** for their help along the way.

Throughout my PhD I have worked on a varied set of projects and have received help from many people. For this I would like to give a special thanks to **Mauricio**, **Zoltan**, **Ulrika**, **Cecilia**, **Maxim**, **Filip**, **Rui**, **Scott**, **Kinga**, **Manuel**, **Fabio**, **Jonathan** and **Andrey**, you have always been willing to help me even if there was nothing in it for you.

I would also like to thank all the people who involved me into their projects through a collaboration. I enjoyed every single project and met some great people because of it. Thank you for the lovely discussions and trusting me with your projects **Raffaella**, **Emanuel**, **Ivan**, **Billy**, **Eric**, **Stefan** (Huber), **Stephen**, **Abiy**, **Alethea**, **Orion** and **Andreas**.

During my undergraduate studies I met many people who taught me a thing or two: **Laura**, **Marc**, **Gert**, **Lisa**, **Jonas**, **Martin**, **Dirk** and **Doreen**. Thank you **Sandra**, **Valérie**, **Pieter** and **Romeo**, we had fun studying together and I appreciate that we are still in touch.

The past couple of years involved more than just research and required that a lot of thing ran smoothly. Thank you **Helena** (Grennberg), **Francoise**, **Johanna** (Andersson), **Helena** (Danielson), **Rikard**, **Jan** (Streuff), **Luke**, **Roland**, **Christine**, **Farshid**, **Gunnar**, **Emil**, **Johanna** (Johansson), **Hanna** (Oskarsson), **Linda**, **Eva**, **Gabriella**, **Krisitina**, **Anela** and **Kevok**.

During my PhD I have had several student helping me with my projects: **Sofie**, **Albin**, **Bono**, **Erik**, **Vasily**, **Manuel**, **Yordanos** and **Gintare**, thank you for the hard work.

Every other week we met up in smaller groups to discuss our projects. The meetings with "Team NAMFIS" have always been great, with **Stefan** and **Yoseph** as stable participants. Thank you **Yoseph** for your kindness and limitless willingness to help, it has been a pleasure working with you. **Stefan**, I have regularly described you as my best college. You were always the one I could count on for nerdy NMR discussion. With you I could discuss things no one else at our department seemed interested in. You are not only a great colleague, but also a dear friend. I am sure we will stay in touch.

Pap, **mam**, **Abel** en **Margreet**, hoewel jullie geen directe wetenschappelijk bijgedragen hebben geleverd, zijn jullie op de achtergrond altijd aanwezig geweest. Het gegeven dat ik altijd op jullie kan rekenen heeft er toe geleid dat ik mij zonder zorgen volledig op het werk kon en kan storten. Bedankt voor jullie steun.

9 References

[1] Y. R. Lai, X. Y. Chu, L. Di, W. Gao, Y. Y. Guo, X. R. Liu, C. Lu, J. L. Mao, H. Shen, H. P. Tang, C. Q. Xia, L. Zhang, X. X. Ding, *Acta Pharm. Sin. B* **2022**, *12*, 2751-2777.

[2] I. Doytchinova, *Molecules* **2022**, *27*, 1496.

[3] L. Di, P. Artursson, A. Avdeef, L. Z. Benet, J. B. Houston, M. Kansy, E. H. Kerns, H. Lennernäs, D. A. Smith, K. Sugano, *Chem. Med. Chem.* **2020**, *15*, 1862-1874.

[4] Y. Chang, B. A. Hawkins, J. J. Du, P. W. Groundwater, D. E. Hibbs, F. Lai, *Pharmaceutics* **2023**, *15*, 49.

[5] D. Narayanan, K. T. Tran, J. S. Pallesen, S. M. Ø. Solbak, Y. Qin, E. Mukminova, M. Luchini, K. O. Vasilyeva, D. González Chichón, G. Goutsiou, C. Poulsen, N. Haapanen, G. M. Popowicz, M. Sattler, D. Olagnier, M. Gajhede, A. Bach, *J. Med. Chem.* **2022**, *65*, 14481-14526.

[6] L. Maveyraud, L. Mourey, *Molecules* **2020**, *25*, 1030.

[7] A. R. Fassihi, *Int. J. Pharm.* **1993**, *92*, 1-14.

[8] S. Parsons, *Tetrahedron Asymmet.* **2017**, *28*, 1304-1313.

[9] H. D. Flack, G. Bernardinelli, *J. Appl. Cryst.* **2000**, *33*, 1143-1148.

[10] M. J. Harner, L. Mueller, K. J. Robbins, M. D. Reily, *Arch. Biochem. Biophys.* **2017**, *628*, 132-147.

[11] K. Teilum, M. B. Kunze, S. Erlendsson, B. B. Kragelund, *Protein Sci.* **2017**, *26*, 436-451.

[12] A. Furukawa, T. Konuma, S. Yanaka, K. Sugase, *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *96*, 47-57.

[13] M. P. Latham, G. R. Zimmermann, A. Pardi, *J. Am. Chem. Soc.* **2009**, *131*, 5052-5053.

[14] N. H. Andersen, C. P. Chen, T. M. Marschner, S. R. Krystek, D. A. Bassolino, *Biochemistry* **1992**, *31*, 1280-1295.

[15] M. L. M. Beckers, L. M. C. Buydens, J. A. Pikkemaat, C. Altona, *J. Biomol. NMR* **1997**, *9*, 25-34.

[16] M. J. Blackledge, R. Bruschweiler, C. Griesinger, J. M. Schmidt, P. Xu, R. R. Ernst, *Biochemistry* **1993**, *32*, 10960-10974.

[17] D. O. Cicero, G. Barbato, R. Bazzo, *J. Am. Chem. Soc.* **1995**, *117*, 1027-1033.

[18] A. E. Cleves, A. N. Jain, *J. Comput. Aided Mol. Des.* **2017**, *31*, 419-439.

[19] C. Landis, V. S. Allured, *J. Am. Chem. Soc.* **1991**, *113*, 9493-9499.

[20] Y. Z. Liu, A. Navarro-Vazquez, R. R. Gil, C. Griesinger, G. E. Martin, R. T. Williamson, *Nat. Protoc.* **2019**, *14*, 217-247.

[21] D. A. Pearlman, *J. Biomol. NMR* **1996**, *8*, 49-66.

[22] J. J. Wang, R. S. Hodges, B. D. Sykes, *J. Am. Chem. Soc.* **1995**, *117*, 8627-8634.

[23] S. Wang, K. Krummenacher, G. A. Landrum, B. D. Sellers, P. Di Lello, S. J. Robinson, B. Martin, J. K. Holden, J. Y. K. Tom, A. C. Murthy, N. Popovych, S. Riniker, *J. Chem. Inf. Model.* **2022**, *62*, 472-485.

[24] S. Macura, B. T. Farmer, L. R. Brown, *J. Magn. Reson.* **1986**, *70*, 493-499.
[25] C. A. G. Haasnoot, F. A. A. M. De Leeuw, C. Altona, *Tetrahedron* **1980**, *36*, 2783-2792.
[26] M. Karplus, *J. Chem. Phys.* **1959**, *30*, 11-15.
[27] M. Raban, K. Mislow, *Tetrahedron Lett.* **1965**, *6*, 4249-4253.
[28] T. J. Wenzel, in *Differentiation of Enantiomers II* (Ed.: V. Schurig), Springer International Publishing, Cham, **2013**, pp. 1-68.
[29] J. Bogaerts, R. Aerts, T. Vermeyen, C. Johannessen, W. Herrebout, J. M. Batista, *Pharmaceuticals* **2021**, *14*.
[30] M. Górecki, J. Frelek, *TrAC* **2021**, *144*, 116428.
[31] P. Changenet, F. Hache, *Eur. Phys. J. Spec. Top.* **2022**.
[32] S. Zhu, M. Sun, *Appl. Spectrosc. Rev.* **2021**, *56*, 553-587.
[33] M. W. Martynowycz, T. Gonen, *Curr. Opin. Colloid Interface Sci.* **2018**, *34*, 9-16.
[34] E. Danelius, S. Halaby, W. A. van der Donk, T. Gonen, *Nat. Prod. Rep.* **2021**, *38*, 423-431.
[35] C. G. Jones, M. W. Martynowycz, J. Hattne, T. J. Fulton, B. M. Stoltz, J. A. Rodriguez, H. M. Nelson, T. Gonen, *ACS Cent. Sci.* **2018**, *4*, 1587-1592.
[36] C. P. Ting, M. A. Funk, S. L. Halaby, Z. Zhang, T. Gonen, W. A. van der Donk, *Science* **2019**, *365*, 280-284.
[37] G. Rivière, S. Oueslati, M. Gayral, J.-B. Créchet, N. Nhiri, E. Jacquet, J.-C. Cintrat, F. Giraud, C. van Heijenoort, E. Lescop, S. Pethe, B. I. Iorga, T. Naas, E. Guittet, N. Morellet, *ACS Omega* **2020**, *5*, 10466-10480.
[38] F. A. A. Mulder, D. Schipper, R. Bott, R. Boelens, *J. Mol. Biol.* **1999**, *292*, 111-123.
[39] A. W. Overhauser, *Phys. Rev.* **1953**, *92*, 411-415.
[40] F. A. L. Anet, A. J. R. Bourn, *J. Am. Chem. Soc.* **1965**, *87*, 5250-5251.
[41] R. A. Bell, J. K. Saunders, *Can. J. Chem.* **1970**, *48*, 1114-1122.
[42] R. E. Schirmer, J. H. Noggle, J. P. Davis, P. A. Hart, *J. Am. Chem. Soc.* **1970**, *92*, 3266-3273.
[43] I. Solomon, *Phys. Rev.* **1955**, *99*, 559-565.
[44] A. Kumar, R. R. Ernst, K. Wuthrich, *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1-6.
[45] R. Richarz, K. Wuthrich, *J. Magn. Reson.* **1978**, *30*, 147-150.
[46] C. P. Butts, C. R. Jones, E. C. Towers, J. L. Flynn, L. Appleby, N. J. Barron, *Org. Biomol. Chem.* **2011**, *9*, 177-184.
[47] H. T. Hu, K. Krishnamurthy, *J. Magn. Reson.* **2006**, *182*, 173-177.
[48] B. Vögeli, *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *78*, 1-46.
[49] J. Keeler, *Understanding NMR Spectroscopy*, Wiley, **2010**.
[50] D. Neuhaus, M. P. Williamson, *The Nucleas Overhauser Effect in structural and conformational analysis*, VCH Publishers, **1989**.
[51] N. R. Krishna, D. G. Agresti, J. D. Glickson, R. Walter, *Biophys. J.* **1978**, *24*, 791-814.
[52] J. H. Noggle, R. E. Schirmer, *The Nuclear Overhauser Effect*, Academic Press, **1971**.
[53] B. E. Coggins, J. W. Werner-Allen, A. Yan, P. Zhou, *J. Am. Chem. Soc.* **2012**, *134*, 18619-18630.
[54] S. Combo, S. Mendes, K. M. Nielsen, G. J. da Silva, S. Domingues, *Biomedicines* **2022**, *10*, 2399.
[55] C. A. Lipinski, *Adv. Drug. Deliv. Rev.* **2016**, *101*, 34-41.
[56] T. D. W. Claridge, *High-resolution NMR techniques in organic chemistry*, 2nd ed., Elsevier, Amsterdam ; Boston, **2009**.

[57] T. Gonen, Y. Cheng, P. Sliz, Y. Hiroaki, Y. Fujiyoshi, S. C. Harrison, T. Walz, *Nature* **2005**, *438*, 633-638.
[58] L. Jiang, D. Georgieva, I. Nederlof, Z. Liu, J. P. Abrahams, *Microsc. Microanal.* **2011**, *17*, 879-885.
[59] S. Stegemann, C. Moreton, S. Svanbäck, K. Box, G. Motte, A. Paudel, *Drug Discov. Today* **2023**, *28*, 103344.
[60] M. Egbert, A. Whitty, G. M. Keserű, S. Vajda, *J. Med. Chem.* **2019**, *62*, 10005-10025.
[61] S. D. Kotsakis, C.-F. Flach, M. Razavi, D. G. J. Larsson, *Antimicrob. Agents Chemother.* **2019**, *63*, e01817-01818.
[62] A. S. Stern, D. L. Donoho, J. C. Hoch, *J. Magn. Reson.* **2007**, *188*, 295-300.
[63] B. T. Farmer, S. Macura, L. R. Brown, *J. Magn. Reson.* **1988**, *80*, 1-22.
[64] S. Boros, Z. Gáspári, G. Batta, in *Annu. Reports NMR Spectrosc., Vol. 94* (Ed.: G. A. Webb), Academic Press, **2018**, pp. 1-39.
[65] C. R. Jones, C. P. Butts, J. N. Harvey, *Beilstein J. Org. Chem.* **2011**, *7*, 145-150.
[66] A. E. Aliev, A. A. Sinitsyna, *Bull. Russ. Acad. Sci. Chem. Sci.* **1992**, *41*, 1143-1160.
[67] D. Strotz, J. Orts, M. Minges, B. Vogeli, *J. Magn. Reson.* **2015**, *259*, 32-46.
[68] A. A. Wu, D. Crerner, *Int. J. Mol. Sci.* **2003**, *4*, 158-192.
[69] S. Ludvigsen, K. V. Andersen, F. M. Poulsen, *J. Mol. Biol.* **1991**, *217*, 731-736.
[70] P. Schmieder, H. Kessler, *Biopolymers* **1992**, *32*, 435-440.
[71] A. Fleming, *Br. J. Exp. Pathol.* **1929**, *10*, 226-236.
[72] G. A. Durand, D. Raoult, G. Dubourg, *Int. J. Antimicrob. Agents* **2019**, *53*, 371-382.
[73] D. R. Dodds, *Biochem. Pharmacol.* **2017**, *134*, 139-146.
[74] T. M. Uddin, A. J. Chakraborty, A. Khusro, B. M. R. M. Zidan, S. Mitra, T. B. Emran, K. Dhama, M. K. H. Ripon, M. Gajdács, M. U. K. Sahibzada, M. J. Hossain, N. Koirala, *J. Infect. Public. Heal.* **2021**, *14*, 1750-1766.
[75] B. Suay-Garcia, M. T. Perez-Gracia, *Antibiotics* **2019**, *8*, 16.
[76] R. A. Bonomo, E. M. Burd, J. Conly, B. M. Limbago, L. Poirel, J. A. Segre, L. F. Westblade, *Clin. Infect. Dis.* **2017**, *66*, 1290-1297.
[77] H. C. J. Gram, C. Friedlaender, *Ueber die isolirte Färbung der Schizomyceten: in Schnitt-und Trockenpräparaten. Fortschritte der Medizin, Vol. 2*, Theodor Fischer's medicinischer Buchhandlung, **1884**.
[78] K. M. Lehman, M. Grabowicz, *Antibiotics* **2019**, *8*, 18.
[79] K. Bush, P. A. Bradford, *Cold Spring Harb. Perspect. Med.* **2016**, *6*.
[80] T. J. Foster, *FEMS Microbiol. Rev.* **2017**, *41*, 430-449.
[81] F. Nainu, A. Masyita, M. A. Bahar, M. Raihan, S. R. Prova, S. Mitra, T. B. Emran, J. Simal-Gandara, *Antibiotics* **2021**, *10*, 822.
[82] E. Peterson, P. Kaur, *Front. Microbiol.* **2018**, *9*.
[83] W. C. Reygaert, *AIMS Microbiol.* **2018**, *4*, 482-501.
[84] J. M. A. Blair, M. A. Webber, A. J. Baylay, D. O. Ogbolu, L. J. V. Piddock, *Nat. Rev. Microbiol.* **2015**, *13*, 42-51.
[85] J. Lin, K. Nishino, M. C. Roberts, M. Tolmasky, R. I. Aminov, L. Zhang, *Front. Microbiol.* **2015**, *6*.
[86] F. Nguyen, A. L. Starosta, S. Arenz, D. Sohmen, A. Donhofer, D. N. Wilson, *Biol. Chem.* **2014**, *395*, 559-575.
[87] C. W. Hall, T.-F. Mah, *FEMS Microbiol. Rev.* **2017**, *41*, 276-301.
[88] J. Iredell, J. Brown, K. Tagg, *BMJ* **2016**, *352*, h6420.

[89] Z. Pang, R. Raudonis, B. R. Glick, T.-J. Lin, Z. Cheng, *Biotechnol. Adv.* **2019**, *37*, 177-192.
[90] N. Reddy, M. Shungube, P. I. Arvidsson, S. Baijnath, H. G. Kruger, T. Govender, T. Naicker, *Expert. Opin. Ther. Pat.* **2020**, *30*, 541-555.
[91] A. M. Somboro, J. O. Sekyere, D. G. Amoako, S. Y. Essack, L. A. Bester, *Appl. Environ. Microbiol.* **2018**, *84*, e00698-00618.
[92] R. P. Ambler, *Philos. Trans. R. Soc. B-Biol. Sci.* **1980**, *289*, 321-331.
[93] J. S. Kang, A. L. Zhang, M. Faheem, C. J. Zhang, N. Ai, J. D. Buynak, W. J. Welsh, P. Oelschlaeger, *J. Chem. Inf. Model.* **2018**, *58*, 1902-1914.
[94] G. Cornaglia, H. Giamarellou, G. M. Rossolini, *The Lancet Infect. Dis.* **2011**, *11*, 381-393.
[95] D. P. Frueh, *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *78*, 47-75.
[96] S. Grzesiek, A. Bax, *J. Magn. Reson.* **1992**, *96*, 432-440.
[97] L. E. Kay, G. Y. Xu, T. Yamazaki, *J. Magn. Reson. Ser. A* **1994**, *109*, 129-133.
[98] J. Schleucher, M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky, S. J. Glaser, O. W. Sorensen, C. Griesinger, *J. Biomol. NMR* **1994**, *4*, 301-306.
[99] R. T. Clubb, V. Thanabal, G. Wagner, *J. Magn. Reson.* **1992**, *97*, 213-217.
[100] D. R. Muhandiram, L. E. Kay, *J. Magn. Reson. Ser. B* **1994**, *103*, 203-216.
[101] M. Wittekind, L. Mueller, *J. Magn. Reson. Ser. B* **1993**, *101*, 201-205.
[102] T. Yamazaki, W. Lee, C. H. Arrowsmith, D. R. Muhandiram, L. E. Kay, *J. Am. Chem. Soc.* **1994**, *116*, 11655-11666.
[103] J. H. Lee, *J. Kor. Magn. Reson. Soc.* **2018**, *22*, 125-131.
[104] S. Skagseth, S. Akhter, M. H. Paulsen, Z. Muhammad, S. Lauksund, O. Samuelsen, H. K. S. Leiros, A. Bayer, *Eur. J. Med. Chem.* **2017**, *135*, 159-173.
[105] J. Brem, S. van Berkel Sander, D. Zollman, Y. Lee Sook, O. Gileadi, J. McHugh Peter, R. Walsh Timothy, A. McDonough Michael, J. Schofield Christopher, *Antimicrob. Agents Chemother.* **2015**, *60*, 142-150.
[106] T. Christopeit, T. J. O. Carlsen, R. Helland, H.-K. S. Leiros, *J. Med. Chem.* **2015**, *58*, 8671-8682.
[107] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv. Drug. Deliv. Rev.* **1997**, *23*, 3-25.
[108] D. A. DeGoey, H.-J. Chen, P. B. Cox, M. D. Wendt, *J. Med. Chem.* **2018**, *61*, 2636-2651.
[109] B. C. Doak, B. Over, F. Giordanetto, J. Kihlberg, *Chem. Biol.* **2014**, *21*, 1115-1142.
[110] B. C. Doak, J. Zheng, D. Dobritzsch, J. Kihlberg, *J. Med. Chem.* **2016**, *59*, 2312-2327.
[111] M. D. Shultz, *J. Med. Chem.* **2019**, *62*, 1701-1714.
[112] R. G. Bryant, *J. Chem. Educ.* **1983**, *60*, 933-935.
[113] A. Henninot, J. C. Collins, J. M. Nuss, *J. Am. Chem. Soc.* **2018**, *61*, 1382-1414.
[114] F. G. Banting, C. H. Best, J. B. Collip, W. R. Campbell, A. A. Fletcher, *Canad. Med. Ass. J.* **1922**, *7*, 141-146.
[115] A. A. Fletcher, *Canad. Med. Ass. J.* **1962**, *87*, 1052-&.
[116] M.-Q. Zhang, B. Wilkinson, *Curr. Opin. Biotechnol.* **2007**, *18*, 478-488.
[117] C. A. Lipinski, *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235-249.
[118] D. F. Veber, S. R. Johnson, H.-Y. Cheng, B. R. Smith, K. W. Ward, K. D. Kopple, *J. Med. Chem.* **2002**, *45*, 2615-2623.
[119] T. W. Johnson, K. R. Dress, M. Edwards, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 5560-5564.
[120] S. Schneckener, S. Grimbs, J. Hey, S. Menz, M. Osmers, S. Schaper, A. Hillisch, A. H. Göller, *J. Chem. Inf. Model.* **2019**, *59*, 4893-4905.
[121] N. A. Meanwell, *Chem. Res. Toxicol.* **2016**, *29*, 564-616.

[122] F. Giordanetto, J. Kihlberg, *J. Med. Chem.* **2014**, *57*, 278-295.
[123] A. Whitty, M. Q. Zhong, L. Viarengo, D. Beglov, D. R. Hall, S. Vajda, *Drug Discov. Today* **2016**, *21*, 712-717.
[124] M. Tyagi, V. Poongavanam, M. Lindhagen, A. Pettersen, P. Sjö, S. Schiesser, J. Kihlberg, *Org. Lett.* **2018**, *20*, 5737-5742.
[125] A. Alex, D. S. Millan, M. Perez, F. Wakenhut, G. A. Whitlock, *Med. Chem. Comm.* **2011**, *2*, 669-674.
[126] G. Caron, G. Ermondi, *Drug Discov. Today* **2017**, *22*, 835-840.
[127] P. Matsson, B. C. Doak, B. Over, J. Kihlberg, *Adv. Drug Deliv. Rev.* **2016**, *101*, 42-61.
[128] P. Matsson, J. Kihlberg, *J. Med. Chem.* **2017**, *60*, 1662-1664.
[129] M. Rossi Sebastiano, B. C. Doak, M. Backlund, V. Poongavanam, B. Over, G. Ermondi, G. Caron, P. Matsson, J. Kihlberg, *J. Med. Chem.* **2018**, *61*, 4189-4202.
[130] L. B. Kier, *Quant. Struct. Act. Relat.* **1989**, *8*, 221-224.
[131] V. Poongavanam, Y. Atilaw, S. Ye, L. H. E. Wieske, M. Erdelyi, G. Ermondi, G. Caron, J. Kihlberg, *J. Pharm. Sci.* **2021**, *110*, 301-313.
[132] E. M. Driggers, S. P. Hale, J. Lee, N. K. Terrett, *Nat. Rev. Drug Discov.* **2008**, *7*, 608-624.
[133] A. K. Yudin, *Chem. Sci.* **2015**, *6*, 30-49.
[134] S. Mabe, J. Eller, W. S. Champney, *Curr. Microbiol.* **2004**, *49*, 248-254.
[135] W. K. Yam, H. A. Wahab, *J. Chem. Inf. Model.* **2009**, *49*, 1558-1567.
[136] W. S. Champney, J. Pelt, *Curr. Microbiol.* **2002**, *45*, 328-333.
[137] W. S. Champney, C. L. Tober, *Curr. Microbiol.* **2000**, *41*, 126-135.
[138] E. A. Campbell, N. Korzheva, A. Mustaev, K. Murakami, S. Nair, A. Goldfarb, S. A. Darst, *Cell* **2001**, *104*, 901-912.
[139] R. Bartenschlager, *Antivir. Chem. Chemother.* **1997**, *8*, 281-301.
[140] S. Beldar, M. S. S. Manimekalai, N. J. Cho, K. Baek, G. Gruber, H. S. Yoon, *J. Gen. Virol.* **2018**, *99*, 194-208.
[141] S. M. Lambert, D. R. Langley, J. A. Garnett, R. Angell, K. Hedgethorne, N. A. Meanwell, S. J. Matthews, *Protein Sci.* **2014**, *23*, 723-734.
[142] S. Venkatraman, F. G. Njoroge, W. Wu, V. Girijavallabhan, A. J. Prongay, N. Butkiewicz, J. Pichardo, *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1628-1632.
[143] D. I. Soumana, N. Kurt Yilmaz, A. Ali, K. L. Prachanronarong, C. A. Schiffer, *J. Am. Chem. Soc.* **2016**, *138*, 11850-11859.
[144] A. Farajallah, R. T. Bunch, N. A. Meanwell, in *Antiviral Drugs*, **2011**, pp. 1-17.
[145] M. Nguyen, E. P. Chung, *Clin. Thera.* **2005**, *27*, 1144-1163.
[146] A. Cervin, B. Wallwork, *Curr. Allergy. Asthma Rep.* **2014**, *14*, 416.
[147] B. Bortone, C. Jackson, Y. Hsia, J. Bielicki, N. Magrini, M. Sharland, *PLoS One* **2021**, *16*, e0241899.
[148] M. Grobbelaar, G. E. Louw, S. L. Sampson, P. D. van Helden, P. R. Donald, R. M. Warren, *Infect. Genet. Evol.* **2019**, *74*, 103937.
[149] Y. Teraoka, T. Uchida, M. Imamura, N. Hiraga, M. Osawa, H. Kan, Y. Saito, M. Tsuge, H. Abe-Chayama, C. N. Hayes, G. N. Makokha, H. Aikata, D. Miki, H. Ochi, Y. Ishida, C. Tateno, K. Chayama, *J. Gen. Virol.* **2018**, *99*, 1058-1065.
[150] S. M. McConachie, S. M. Wilhelm, P. B. Kale-Pradhan, *Expert Rev. Clin. Pharmacol.* **2016**, *9*, 287-302.
[151] E. DeJesus, J. K. Rockstroh, K. Henry, J.-M. Molina, J. Gathe, S. Ramanathan, X. Wei, K. Yale, J. Szwarcberg, K. White, A. K. Cheng, B. P. Kearney, *The Lancet* **2012**, *379*, 2429-2438.

[152] G. Chang, W. C. Guida, W. C. Still, *J. Am. Chem. Soc.* **1989**, *111*, 4379-4386.
[153] P. C. D. Hawkins, *J. Chem. Inf. Model.* **2017**, *57*, 1747-1756.
[154] G. Olanders, H. Alogheli, P. Brandt, A. Karlén, *J. Comput. Aided Mol. Des.* **2020**, *34*, 231-252.
[155] C. R. W. Guimarães, A. M. Mathiowetz, M. Shalaeva, G. Goetz, S. Liras, *J. Chem. Inf. Model.* **2012**, *52*, 882-890.
[156] D. I. Soumana, N. Kurt Yilmaz, K. L. Prachanronarong, C. Aydin, A. Ali, C. A. Schiffer, *ACS Chem. Biol.* **2016**, *11*, 900-909.
[157] M. D. Cummings, J. Lindberg, T.-I. Lin, H. de Kock, O. Lenz, E. Lilja, S. Felländer, V. Baraznenok, S. Nyström, M. Nilsson, L. Vrang, M. Edlund, Å. Rosenquist, B. Samuelsson, P. Raboisson, K. Simmen, *Angew. Chem. Int. Ed.* **2010**, *49*, 1652-1655.
[158] J. E. DeLorbe, J. H. Clements, M. G. Teresk, A. P. Benfield, H. R. Plake, L. E. Millspaugh, S. F. Martin, *J. Am. Chem. Soc.* **2009**, *131*, 16758-16770.
[159] A. Mann, in *The Practice of Medicinal Chemistry (Third Edition)* (Ed.: C. G. Wermuth), Academic Press, New York, **2008**, pp. 363-379.
[160] M. D. Cummings, T.-I. Lin, L. Hu, A. Tahri, D. McGowan, K. Amssoms, S. Last, B. Devogelaere, M.-C. Rouan, L. Vijgen, J. M. Berke, P. Dehertogh, E. Fransen, E. Cleiren, L. van der Helm, G. Fanning, O. Nyanguile, K. Simmen, P. Van Remoortere, P. Raboisson, S. Vendeville, *J. Med. Chem.* **2014**, *57*, 1880-1892.
[161] T. Rezai, J. E. Bock, M. V. Zhou, C. Kalyanaraman, R. S. Lokey, M. P. Jacobson, *J. Am. Chem. Soc.* **2006**, *128*, 14073-14080.
[162] D. Shi, B. L. Nannenga, M. J. de la Cruz, J. Liu, S. Sawtelle, G. Calero, F. E. Reyes, J. Hattne, T. Gonen, *Nat. Protoc.* **2016**, *11*, 895-904.
[163] D. Shi, B. L. Nannenga, M. G. Iadanza, T. Gonen, *eLife* **2013**, *2*, e01345.
[164] J. Hattne, M. W. Martynowycz, P. A. Penczek, T. Gonen, *IUCrJ* **2019**, *6*, 921-926.
[165] D. Shi, M. R. Lewis, H. S. Young, D. L. Stokes, *J. Mol. Biol.* **1998**, *284*, 1547-1564.
[166] M. J. Gallenito, T. Gonen, *Biochem. Soc. Trans.* **2022**, *50*, 231-239.
[167] M. W. Martynowycz, A. Shiriaeva, M. T. B. Clabbers, W. J. Nicolas, S. J. Weaver, J. Hattne, T. Gonen, *Nat. Commun* **2023**, *14*, 1086.
[168] C. Nguyen, H.-T. Lei, L. T. F. Lai, M. J. Gallenito, D. Matthies, T. Gonen, *bioRxiv* **2022**, 2022.2005.2031.494244.
[169] J. F. Bruhn, G. Scapin, A. C. Cheng, B. Q. Mercado, D. G. Waterman, T. Ganesh, S. Dallakyan, B. N. Read, T. Nieusma, K. W. Lucier, M. L. Mayer, N. J. Chiang, N. Poweleit, P. T. McGilvray, T. S. Wilson, M. Mashore, C. Hennessy, S. Thomson, B. Wang, C. S. Potter, B. Carragher, *Front. Mol. Biosci.* **2021**, *8*.
[170] B. J. Curtis, L. J. Kim, C. J. J. Wrobel, J. M. Eagan, R. A. Smith, J. E. Burch, H. H. Le, A. B. Artyukhin, H. M. Nelson, F. C. Schroeder, *Org. Lett.* **2020**, *22*, 6724-6728.
[171] N. T. Duong, Y. Aoyama, K. Kawamoto, T. Yamazaki, Y. Nishiyama, *Molecules* **2021**, *26*.
[172] P. R. Gleason, B. L. Nannenga, J. H. Mills, *Front. Mol. Biosci.* **2021**, *7*.
[173] C. G. Jones, M. Asay, L. J. Kim, J. F. Kleinsasser, A. Saha, T. J. Fulton, K. R. Berkley, D. Cascio, A. G. Malyutin, M. P. Conley, B. M. Stoltz, V. Lavallo, J. A. Rodriguez, H. M. Nelson, *ACS Cent. Sci.* **2019**, *5*, 1507-1513.
[174] L. J. Kim, M. Z. Xue, X. Li, Z. Xu, E. Paulson, B. Mercado, H. M. Nelson, S. B. Herzon, *J. Am. Chem. Soc.* **2021**, *143*, 6578-6585.
[175] T. Kunde, B. M. Schmidt, *Angew. Chem. Int. Ed.* **2019**, *58*, 666-668.

[176] S. Sekharan, X. T. Liu, Z. C. Yang, X. Liu, L. Deng, S. G. Ruan, Y. Abramov, G. X. Sun, S. Z. Li, T. Zhou, B. M. Shi, Q. Zeng, Q. Zeng, C. Chang, Y. D. Jin, X. K. Shi, *RSC Adv.* **2021**, *11*, 17408-17412.
[177] B. Wang, Y. Q. Lin, *Chem. Comm.* **2022**, *58*, 13071-13074.
[178] Y. Watanabe, S. Takahashi, S. Ito, T. Tokiwa, Y. Noguchi, H. Azami, H. Kojima, M. Higo, S. Y. K. Ban, K. Nagai, T. Hirose, T. Sunazuka, T. Yaguchi, K. Nonaka, M. Iwatsuki, *Org. Biomol. Chem.*
[179] F. Banihashemi, G. H. Bu, A. Thaker, D. Williams, J. Y. S. Lin, B. L. Nannenga, *Ultramicroscopy* **2020**, *216*.
[180] H. J. Cho, K. S. Kim, H. Kim, T. Kim, A. G. Malyutin, D. C. Rees, B. K. Yoo, C. Song, *ACS Appl. Mater. Interfaces* **2021**, *13*, 7546-7555.
[181] E. Kapaca, J. X. Jiang, J. Cho, J. L. Jorda, M. J. Diaz-Cabanas, X. D. Zou, A. Corma, T. Willhammar, *J. Am. Chem. Soc.* **2021**, *143*, 8713-8719.
[182] T. M. Yang, T. Willhammar, H. Y. Xu, X. D. Zou, Z. H. Huang, *Nat. Protoc.* **2022**, *17*, 2389-2413.
[183] S. Gim, G. Fittolani, Y. Nishiyama, P. H. Seeberger, Y. Ogawa, M. Delbianco, *Angew. Chem. Int. Ed.* **2020**, *59*, 22577-22583.
[184] S. Halaby, M. W. Martynowycz, Z. Zhu, S. Tretiak, A. Zhugayevych, T. Gonen, M. Seifrid, *Chem. Mater.* **2021**, *33*, 966-977.
[185] M. Lightowler, S. Li, X. Ou, X. Zou, M. Lu, H. Xu, *Angew. Chem. Int. Ed.* **2022**, *61*, e202114985.
[186] Q. Liu, N. Hoefer, G. Berkbigler, Z. Cui, T. Liu, A. C. Co, D. W. McComb, C. R. Wade, *Inorg. Chem.* **2022**, *61*, 18710-18718.
[187] H. A. Mills, C. G. Jones, K. P. Anderson, A. D. Ready, P. I. Djurovich, S. I. Khan, J. N. Hohman, H. M. Nelson, A. M. Spokoyny, *Chem. Mater.* **2022**, *34*, 6933-6943.
[188] S. Vergara, D. A. Lukes, M. W. Martynowycz, U. Santiago, G. Plascencia-Villa, S. C. Weiss, M. J. de la Cruz, D. M. Black, M. M. Alvarez, X. López-Lozano, C. O. Barnes, G. Lin, H.-C. Weissker, R. L. Whetten, T. Gonen, M. J. Yacaman, G. Calero, *J. Phys. Chem. Lett.* **2017**, *8*, 5523-5530.
[189] S. Ito, F. J. White, E. Okunishi, Y. Aoyama, A. Yamano, H. Sato, J. D. Ferrara, M. Jasnowski, M. Meyer, *Cryst. Eng. Comm.* **2021**, *23*, 8622-8630.
[190] W. Zhou, B. Bammes, P. G. Mitchell, K. Betz, W. Chiu, *Ultramicroscopy* **2022**, *232*, 113417.
[191] W. Kabsch, *Acta Cryst. Sec. D* **2010**, *66*, 125-132.
[192] C. E. Shannon, *Proc. IRE* **1949**, *37*, 10-21.
[193] W. P. Aue, E. Bartholdi, R. R. Ernst, *J. Chem. Phys.* **1976**, *64*, 2229-2246.
[194] J. C. J. Barna, E. D. Laue, M. R. Mayger, J. Skilling, S. J. P. Worrall, *J. Magn. Reson.* **1987**, *73*, 69-77.
[195] B. E. Coggins, R. A. Venters, P. Zhou, *J. Am. Chem. Soc.* **2005**, *127*, 11562-11563.
[196] Ē. Kupče, R. Freeman, *J. Am. Chem. Soc.* **2003**, *125*, 13958-13959.
[197] S. G. Hyberts, H. Arthanari, G. Wagner, in *Novel Sampling Approaches in Higher Dimensional Nmr, Vol. 316* (Eds.: M. Billeter, V. Orekhov), **2012**, pp. 125-148.
[198] S. Robson, H. Arthanari, S. G. Hyberts, G. Wagner, in *Methods Enzymol., Vol. 614* (Ed.: A. J. Wand), Academic Press, **2019**, pp. 263-291.
[199] C. P. Pedersen, A. Prestel, K. Teilum, *Magn. Reson. Chem.* **2021**, *59*, 315-323.
[200] V. Y. Orekhov, V. A. Jaravine, *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59*, 271-292.
[201] D. L. Donoho, *IEEE Trans. Inf. Theory.* **2006**, *52*, 1289-1306.

[202] D. J. Holland, M. J. Bostock, L. F. Gladden, D. Nietlispach, *Angew. Chem. Int. Ed.* **2011**, *50*, 6548-6551.
[203] S. G. Hyberts, G. J. Heffron, N. G. Tarragona, K. Solanky, K. A. Edmonds, H. Luithardt, J. Fejzo, M. Chorev, H. Aktas, K. Colson, K. H. Falchuk, J. A. Halperin, G. Wagner, *J. Am. Chem. Soc.* **2007**, *129*, 5108-5116.
[204] S. J. Sun, M. Gill, Y. F. Li, M. Huang, R. A. Byrd, *J. Biomol. NMR* **2015**, *62*, 105-117.
[205] J. A. Aguilar, A. M. Kenwright, *Magn. Reson. Chem.* **2018**, *56*, 983-992.
[206] M. S. Roginkin, I. E. Ndukwe, D. L. Craft, R. T. Williamson, M. Reibarkh, G. E. Martin, D. Rovnyak, *Magn. Reson. Chem.* **2020**, *58*, 625-640.
[207] C. Zhang, Y. Idelbayev, N. Roberts, Y. W. Tao, Y. Nannapaneni, B. M. Duggan, J. Min, E. C. Lin, E. C. Gerwick, G. W. Cottrell, W. H. Gerwick, *Sci. Rep.* **2017**, *7*, 17.
[208] P. J. Nichols, A. Born, M. A. Henen, D. Strotz, D. N. Jones, F. Delaglio, B. Vögeli, *J. Biomol. NMR* **2020**, *74*, 717-739.
[209] A. Hussain, N. Paukovich, M. A. Henen, B. Vögeli, *Methods* **2022**, *206*, 87-98.
[210] I. K. L. Lawrence, *Biometrics* **1989**, *45*, 255-268.
[211] R. Anderson-Sprecher, *The American Statistician* **1994**, *48*, 113-117.
[212] A. Bax, D. G. Davis, *J. Magn. Reson.* **1985**, *63*, 207-213.
[213] E. Ämmälahti, M. Bardet, D. Molko, J. Cadet, *J. Magn. Reson.* **1996**, *122*, 230-232.
[214] R. W. Adams, C. M. Holroyd, J. A. Aguilar, M. Nilsson, G. A. Morris, *Chem. Comm.* **2013**, *49*, 358-360.
[215] P. W. A. Howe, *Anal. Chem* **2018**, *90*, 4316-4319.
[216] R. T. McKay, in *Annu. Reports NMR Spectrosc., Vol. 66*, Academic Press, **2009**, pp. 33-76.
[217] M. Piotto, V. Saudek, V. Sklenář, *J. Biomol. NMR* **1992**, *2*, 661-665.
[218] A. J. Simpson, S. A. Brown, *J. Magn. Reson.* **2005**, *175*, 340-346.
[219] R. Dass, P. Kasprzak, W. Kozminski, K. Kazimierczuk, *J. Magn. Reson.* **2016**, *265*, 108-116.
[220] E. Danelius, H. Andersson, P. Jarvoll, U. Brath, M. Erdelyi, *Abstr. Pap. Am. Chem. Soc.* **2017**, *253*, 2.
[221] E. Danelius, U. Brath, M. Erdelyi, *Synlett* **2013**, *24*, 2407-2410.
[222] S. Peintner, M. Erdélyi, *Chem. Eur. J.* **2022**, *28*, e202103559.
[223] R. Dickman, E. Danelius, S. A. Mitchell, D. F. Hansen, M. Erdelyi, A. B. Tabor, *Chem. Eur. J.* **2019**, *25*, 14572-14582.
[224] Y. Atilaw, V. Poongavanam, C. Svensson Nilsson, D. Nguyen, A. Giese, D. Meibom, M. Erdelyi, J. Kihlberg, *ACS Med. Chem. Lett.* **2021**, *12*, 107-114.
[225] V. Poongavanam, Y. Atilaw, S. Siegel, A. Giese, L. Lehmann, D. Meibom, M. Erdelyi, J. Kihlberg, *J. Med. Chem.* **2022**, *65*, 13029-13040.
[226] C. P. Butts, C. R. Jones, J. N. Harvey, *Chem. Comm.* **2011**, *47*, 1193-1195.
[227] J. Du, W. Lü, S. Wu, Y. Cheng, E. Gouaux, *Nature* **2015**, *526*, 224-229.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 2262

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-500237

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2023