universität wien

DISSERTATION | DOCTORAL THESIS

Titel | Title

Protein Epitope Analysis by NMR Spectroscopy

verfasst von | submitted by

Andreas Beier BSc MSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 796 605 419
Dissertationsgebiet lt. Studienblatt | Field of study as it appears on the student record sheet:	Chemie
Betreut von | Supervisor:	Univ.-Prof. Dr. Robert Konrat

SUPERVISOR:
Univ.-Prof. Dr. Robert Konrat

LOCATION:
Vienna, Austria

TIME FRAME:
March 2018 - September 2024

ACKNOWLEDGMENTS

Many people have accompanied and supported me along the way. I wish to express my gratitude to:

- my supervisor Prof. Robert Konrat - for his trust and confidence in me over all these years. Thank you for your personal and scientific support and your magnetic optimism. I feel grateful to have a boss who I can call a friend.
- my friend and colleague Gerald Platzer - for his guidance and support. Thank you for always being someone I can count on, both at work and beyond.
- my friend and colleague Thomas Schwarz - for teaching me what I know about labwork and his invaluable expertise.
- my friends and colleagues at the department - Daniel Braun, Clemens Kauffmann, Theresa Höfurthner, Irene Ceccolini, Julian Holzinger, Sven Brüschweiler, Sonja Knödelstorfer, Karin Ledolter, Mario Migotti, Aleksandra Ptaszek and many more - for making the office a second home to me. May we have many more coffee breaks and discussions together.
- my partner Maria - for sharing the best times and enduring the worst.
- my parents Marlene and Heinrich - for supporting me every step of the way.

PUBLICATIONS

Parts of this thesis have previously been published as scientific articles in peer-reviewed academic journals. These publications are featured as separate chapters in order of appearance:

[1] T. C. Schwarz, A. Beier, K. Ledolter, *et al.*, "High-resolution structural information of membrane-bound α-synuclein provides insight into the moa of the anti-parkinson drug ucb0599," *Proceedings of the National Academy of Sciences*, vol. 120, no. 15, e2201910120, 2023.

[2] A. Beier, G. Platzer, T. Ho¨furthner, *et al.*, "Probing protein–ligand methyl- π interaction geometries through chemical shift measurements of selectively labeled methyl groups," *Journal of Medicinal Chemistry*, vol. 67, no. 15, pp. 13 187–13 196, 2024.

ABSTRACT

Nuclear magnetic resonance (NMR) has established its role as a pivotal technique that can provide information both on protein-protein interfaces as well as on protein-ligand systems and binding epitopes. It is especially suitable for the important group of intrinsically disordered proteins (IDPs) which lack a stable tertiary structure in solution and are not accessible to many other standard structure determination techniques.

We investigate the IDP α-Synuclein which upon binding to membrane surfaces forms predominately α-helical multimeric complexes. For this purpose, we measured paramagnetic observables in the context of SDS micelles as membrane mimics. In addition, we carried out cross-linking mass-spectrometry (MS) experiments on ^{15}N labeled protein mixtures in order to determine crosslinks at the multimer interface. By combining both NMR and MS observables, we can determine monomer structures and further refine them into the most probable dimeric form.

In a second project, we investigated the chemical shift change in proteins induced by aromatic ligands. For the protein BRD4-BD1 and a set of known ligands, using a simple model of the induced shift applied to co-crystal structures of the protein-ligand complexes, we can get a good correlation between experimental and calculated chemical shift changes. In addition, we find good agreement between the interaction geometries between aromatic systems and protein methyl groups found in the protein data bank (PDB) and our model system. We combine both interaction geometry propensities and chemical shift models to fit the most probable position and orientation of aromatic centers in the bound state. As this approach is available even in the absence of an experimentally determined structure of the protein, it can provide invaluable information during structure guided drug development and optimization.

ZUSAMMENFASSUNG

Die Kernspinresonanzspektroskopie (NMR) hat sich als eine entscheidende Technik etabliert, die sowohl Informationen über Protein-Protein Interaktionen als auch über Protein-Ligand Systeme und Bindungsepitope liefern kann. Sie eignet sich besonders für die wichtige Gruppe der intrinsisch ungeordneten Proteine (IDPs), die keine stabile Tertiärstruktur aufweisen und für viele andere Standardmethoden zur Strukturbestimmung nicht zugänglich sind.

Wir untersuchen das IDP α-Synuclein, das bei der Bindung an Membranoberflächen vorwiegend α-helikale multimere Komplexe bildet. Zu diesem Zweck haben wir paramagnetische Observablen im Kontext von SDS-Mizellen als Membranmimetika gemessen. Darüber hinaus führten wir Cross-Linking Massenspektrometrie (MS) Experimente an ^{15}N-markierten Proteingemischen durch, um Crosslinks an der Multimer-Schnittstelle zu bestimmen. Durch die Kombination von NMR- und MS-Observablen können wir Monomerstrukturen bestimmen und weiter zu der wahrscheinlichsten dimeren Form verfeinern.

In einem zweiten Projekt untersuchten wir die chemische Verschiebung in Proteinen, die durch aromatische Liganden induziert wird. Für das Protein BRD4-BD1 und eine Reihe bekannter Liganden können wir unter Verwendung eines einfachen Modells der induzierten Verschiebung, das auf Ko-Kristallstrukturen der Protein-Ligand-Komplexe angewendet wird, eine gute Korrelation zwischen experimentellen und berechneten chemischen Verschiebungsänderungen erzielen. Darüber hinaus finden wir eine gute Übereinstimmung zwischen den Interaktionsgeometrien zwischen aromatischen Systemen und Protein-Methylgruppen, die in der Protein Data Bank (PDB) gefunden wurden, und unserem Modellsystem. Wir kombinieren sowohl Interaktionsgeometrie-Wahrscheinlichkeiten als auch chemische Verschiebungsmodelle, um die wahrscheinlichste Position und Orientierung aromatischer Zentren im gebundenen Zustand zu bestimmen. Da dieser Ansatz auch ohne experimentell bestimmte Struktur des Proteins verfügbar ist, kann er während der strukturgesteuerten Arzneimittelentwicklung und -optimierung unschätzbare Informationen liefern.

CONTENTS

Part I

INTRODUCTION

1

NMR: THE SWISS ARMY KNIFE OF STRUCTURAL BIOLOGY

In 1894, Emil Fischer described the relationship between enzymes and their substrates as "having to fit each other like a lock and a key" [1], [2]. Since then, the structure function paradigm has dominated structural biology, well before Francis Crick defined his central dogma of biology [3] and even before amino acids were recognized as the building blocks of proteins, again by Hofmeister and Fischer [4, p. 163-165]. Therefore, it is not surprising that structural biology research has commonly been treated as a means to an end in the past, the end being the discovery of a 3-dimensional structure of a protein of interest. However, the means have traditionally been X-ray crystallography and, after major technical advances, cryoelectron microscopy [5] (Fig. 1). Nuclear magnetic resonance (NMR), on the other hand, at least compared to the aforementioned techniques, has gained a reputation as an overcomplicated and expensive method, providing ambiguous results that are hard to interpret. This misconception comes - to adapt a common proverb - from trying to oversimplify most questions in structural biology into a nail represented by a high-resolution 3-dimensional structure, for which NMR is not the ideal hammer [6].

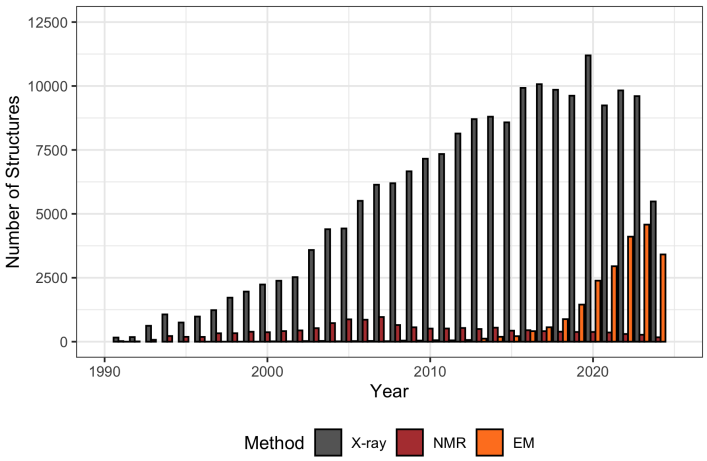


Figure 1: Number of Deposition to the PDB by year and method. Data available at the PDB [7].

Instead, NMR should be seen as a full toolbox of methods to investigate proteins not only as a single representative structure but in the full context of their inherently dynamic behavior [8], [9]. While X-ray crystallography and cryo-EM rely on a stable 3-dimensional

structure of a macromolecule, NMR reports on a per atom basis, in a population-weighted manner on all available substates of the conformational ensemble of the system and on multiple different timescales from picoseconds to seconds. While these timescales sound ambiguous at first, what this means is, that NMR is able to characterize the motions of a macromolecule both on the level of single atoms or side chains as well as on the level of segmental rearrangements and common protein-protein interactions [10], [11].

The relevance of protein dynamics on multiple levels is in turn highlighted by the recognition of the important role of intrinsically disordered proteins in recent decades and the continuous interest in their research [12]. For these highly dynamic systems lacking a stable structure, NMR remains one of the only methods to characterize them properly [13], [14]. But even on the side of classical folded proteins, the emergence of recent deep learning based methods for in-silico structure determination from primary sequences has made the availability of accurate structures a non-issue, while their dynamics still require significant experimental effort [15].

The highlighted publications of this thesis cover a broad spectrum of NMR applications from 3D structure determination to the elucidation of protein-ligand interaction geometries. Therefore, the remainder of this introduction will try to introduce the concepts necessary for the understanding of the highlighted topics.

2

PRE AND PRI

2.1 PARAMAGNETIC RELAXATION ENHANCEMENT

Paramagnetic relaxation enhancement (PRE) has proven to be a valuable technique in NMR structure elucidation [16], [17]. Especially in the field of disordered proteins, it serves as a common way to determine the presence of long range interactions and compact conformations in their structural ensemble [18]–[21]. The PRE effect arises from the presence of a paramagnetic species in the sample. These paramagnetic species posses unpaired electrons that exhibit a large magnetic moment. As the gyromagnetic ratio of a particle is the ratio between its magnetic moment and its angular momentum, the gyromagnetic ratio of an unpaired electron is expected to be in the range of $2.8 \times 10^4 MHz \cdot T^{-1}$ or approimately 650 times larger than of a proton [22]. The magnetic dipolar interaction between a nucleus of interest and the magnetic moment of the unpaired electron leads to an increase in nuclear relaxation. The increase in longitudinal and transversal relaxation rates can be described by the Solomon-Bloembergen equations (Equations 1-2) [23], [24].

$$\Gamma_1 = \frac{2}{5}(\frac{\mu_0}{4\pi})^2 \gamma_i^2 g^2 \mu_b^2 S(S+1) J_{SB}(\omega_I) \tag{1}$$

$$\Gamma_2 = \frac{1}{15}(\frac{\mu_0}{4\pi})^2 \gamma_I^2 g^2 \mu_B^2 S(S+1)\{4J_{SB}(0) + 3J_{SB}(\omega_I)\} \tag{2}$$

Here, g is the electron g-factor, μ_B is the Bohr magneton, γ_I is the gyromagnetic ratio of the nucleus under investigation, S is the electron spin quantum number, and J_{SB} is the generalized spectral density function given by equation 3. The effective correlation time τ_c is defined as $(\tau_r^{-1} + \tau_s^{-1})$, τ_r being the rotational correlation time of the protein or macromolecule and τ_s the electron spin relaxation time. At the same time, the spectral density function relates the effect to the distance (r^{-6}) between the nucleus and the paramagnetic probe.

$$J_{SB}(\omega) = r^{-6} \frac{\tau_c}{1 + (\omega\tau_c)^2} \tag{3}$$

In order to measure a PRE effect, a paramagnetic species has to be introduced into the protein of interest and measured both in its paramagnetic and diamagnetic state. Multiple different paramagnetic species have been mentioned in the literature and are commercially available [25]. By far the most popular class of spin labels are nitroxide radicals which generally take the form of 5- or 6-membered

heterocycles in which an unpaired electron is positioned at an $N-O\bullet$ group of the ring [26]. An additional thiosulfate ester group allows the formation of disulfide bonds to thiol groups in cysteine residues in the protein of interest, that are either naturally occurring or have been introduced by mutagenesis [27]. Conversion of the paramagnetic probe into its diamagnetic state is conveniently achieved by selective reduction of the free radical by the addition of ascorbic acid or other mild reduction agents. Therefore, both paramagnetic as well as diamagnetic rates can be measured on the same sample, thus removing a potential source of systematic error stemming from variations in separate sample preparation. Alternatively, various metal ions like Mn^{2+}, Fe^{3+} and a large part of the lanthanoid series are available [28], [29].

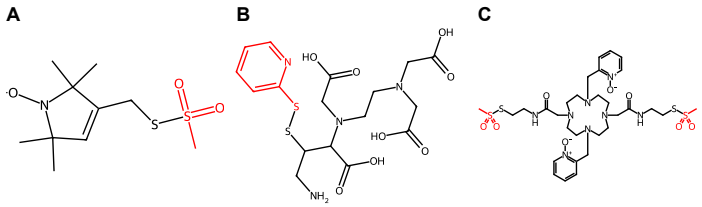


Figure 2: Examples of common protein spin probes [25]. Leaving groups that are substituted during disulfide bond formation are marked in red. A: MTSL as the most common Nitroxide spin label [26]. B: SPy-EDTA as the first EDTA based spin label [30]. C: CLaNP-5 as an example of double anchored spin probes [31].

Apart from binding to intrinsic metal binding sites of the proteins, different metal chelating tags similar to EDTA are available that can be attached to cysteine residues in the protein in a similar fashion as nitroxide spin labels [30], [32]. These tags can then be loaded, either with the chosen paramagnetic species, or with a diamagnetic reference ion, commonly Ca^{2+}. Another important factor to keep in mind when choosing a suitable paramagnetic species is the separation into probes with an isotropic or anisotropic g-factor. While nitroxide spin labels, Mn^{2+} and Gd^{3+} posses an isotropic g-factor, most other metal ions like Fe3+ and most lanthanoids posses an anisotropic g-factor [33]. This distinction is important, as in addition to the PRE effect described in this work, anisotropic spin tags lead to additional observables, namely pseudocontact shifts and residual dipolar couplings [34], [35]. While these observables can yield valuable information about the protein system, they also complicate the analysis of the data, as separate signal assignments for paramagnetic and diamagnetic samples are necessary and additional relaxation pathways via curie-spin relaxation and its cross-correlations can have a significant impact [36].

For isotropic spin probes on the other hand, analysis is relatively straightforward. The desired relaxation rate is measured for the para-

magnetic and diamagnetic state of the tagged protein system and the PRE rate is subsequently determined by calculating the difference between both rates given a specific labeling site x, $\Gamma_{x,PRE} = \Gamma_{x,para} - \Gamma_{x,dia}$. Determination of the respective rates can be achieved by measuring pseudo 3D experiments. In the case of backbone amide protons, the ^1H relaxation delay is often incorporated into the first INEPT transfer and visualized as a pseudo dimension during spectral processing [37]. The relaxation rates can then be determined via a monoexponential fit of the signal intensities in the 2D $^{15}N-^1H$ HSQC slices and their respective relaxation delays.

From Equation 1-3, it follows, that the PRE effect is dependent on the length of the nucleus-electron vector r as well as on its reorientation time. Due to the strong distance dependence of r^{-6}, the observable effect is dominated by protein conformations that exhibit short distances between the respective nucleus and the label position. This make it possible to determine the existence of encounter complexes between protein-protein and protein-ligand systems as well as the presence of minor states in exchanging systems [38]. This has proven especially useful in IDP systems to show the existence of long range interactions that are overall lowly populated in the broad conformational ensemble but dominate the PRE due to the strong distance dependence [39].

Due to the form of the reduced correlation function in Equation 3, the effective correlation time τ_c is dependent on both the nucleus-electron vector correlation time τ_r and the electron spin relaxation time τ_s. In the case of nitroxide spin probes and most isotropic probes in general, the electron spin relaxation time lies in the range of $10^{-7} - 10^{-8}$ and, for small to medium protein sizes, the effective correlation time is therefore dominated by τ_r [40], [41]. An important assumption in this case is, that the r-vectors are rigid in the molecular frame. While there are methods available that incorporate fast internal motions in a model-free-like approach, $^1H-\Gamma_2$ rates are generally rather insensitive to these internal motions. This is especially important, as most spin tags contain multiple rotatable bonds that allows them to sample a large number of conformations in solution. Both the Solomon-Bloembergen as well as the Model-Free Solomon-Bloembergen models for the PRE have been incorporated into the popular simulated annealing package Xplor-NIH as energy functions in order to use experimental PRE rates in structure calculation [42]–[44].

2.2 PARAMAGNETIC RELAXATION INTERFERENCE

While the PRE allows for the observation of residual structures in highly dynamic systems like IDPs, all thermally accessible substates contribute equally and therefore, the detection of concerted or corre-

lated compaction is not possible [45], [46]. Paramagnetic Relaxation Interference (PRI) aims to close this gap. In contrast to the PRE, for the PRI, two spin labels at different positions are incorporated into the protein of interest simultaneously. Apart from the already described PRE effect, this leads to an additional dipole-dipole cross-correlation effect between both spin-label electrons and the nucleus under investigation. To quantify this effect, the single spin label PREs for both spin label positions are measured separately ($\Gamma_{2,X1}$ and $\Gamma_{2,X2}$) as well as the increase in transverse relaxation rate when both spin labels are present $\Gamma_{2,X1,X2}$. The PRI effect is then calculated as the difference between double and single spin labelled relaxation enhancement [45].

$$\Delta\Gamma_{2,X1,X2} = (\Gamma_{2,X1} + \Gamma_{2,X2}) - \Gamma_{2,X1,X2} \tag{4}$$

In general, cross-correlated relaxation effects can be observed if 2 or more relaxation effects of the same rank are active at the same time and in a correlated manner. For dipole-dipole cross-correlation effects in our case, the rate is dependent on the distance between the nucleus and each of the spin labels by the factor r^{-3} as well as on the angle θ formed by the vectors connecting the nucleus to each of the spin labels. This dependence takes the form of the second order Legendre polynomial.

$$r_{X1,i}^{-3} r_{X2,i}^{-3} < 3cos(\theta)^2 - 1 > \tag{5}$$

In the past we have measured PRI rates in several different IDP systems like quail Osteontin (qOPN) and human Brain-Acid-Soluable-Protein (BASP). For different label positions we determined significantly different PRI rates that in each case prove the simultaneous spatial proximity of the labeling sites and the nucleus under investigation. It follows from equation 5, that we can expect a different sign for the measured rate depending on the spatial geometry of the 3-spin system. Indeed for qOPN, depending on the label sites, we find either predominantly positive or negative PRI rates and this was in good agreement with earlier research hinting at the existence of a globally compacted state in the conformational ensemble of qOPN [46]–[48].

While we used Flory's polymer theory to explain the PRI rates in qOPN and prototypical IDPs in general, for folded proteins or disordererd proteins that undergo folding upon binding, it should be possible to use the PRI rates for the calculation of 3-dimensional structures, similar to PREs or NOEs. In order to simplify the problem, we ignore the distance dependence between the labels and the nucleus. We feel that this is warranted, as this distance dependence is already partly encoded in the PRE rates that are necessary for the PRI rate calculations and can be included in any structure determination as separate PRE potentials. Instead, we focus on the angular dependence and convert it into a step-like potential [49].

$$E_{PRI} = S \times \begin{cases} \frac{3}{4}|3cos^2\theta - 1| - \frac{3cos^2\theta - 1}{4}, & \text{if } cPRI \times ePRI < 0 \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

For this we separate the PRI rates into 2 broad angle segments that according to Eq. 5 lead to either positive or negative PRI rates. Accordingly we convert the experimental PRI into an expected PRI (ePRI) that is either -1 in the case of negative rates or +1 in the case of positive rates. At each timestep of the calculation, we calculate the result of the term $3cos^2(\theta) - 1$ for each rate as the calculated PRI (cPRI) and depending on the sign of the product between ePRI and cPRI we apply an energy penalty according to Eq. 6 and calculate the respective gradient to guide the system during simulated annealing.

In one of the articles featured in this thesis, we incorporated this novel energy potential into Xplor-NIH and used it to calculate the monomer and dimer structures of α-synuclein, which folds into a compact α-helical structure upon interaction with biological membranes as well as common membrane mimics.

3

CH-π INTERACTIONS AND THE CHEMICAL SHIFT

3.1 THE CHEMICAL SHIFT

The chemical shift is probably the first observable that one comes into contact with when being introduced to NMR. As such, it contains a tremendous amount of information about the system under investigation that is often overlooked. Indeed, the chemical shift encodes information on a per atom level about the molecules covalent and non-covalent structure, its local conformation and the presence of functional groups. Therefore, it is not surprising that there are multiple approaches available that predict backbone dihedral angles, secondary structures, oxidation and ionization states and others purely from chemical shift information [50]–[54].

The isotropic chemical shift δ in ppm, that is most commonly meant when talking about the chemical shift in general, is defined as follows [55].

$$\delta_{iso} = 10^6(\nu - \nu_{ref})/\nu_{ref} \tag{7}$$

Here, ν is the resonance frequency of a nucleus and ν_{ref} is its reference resonance frequency. The resonance frequency in turn is determined by B_{eff}, the effective magnetic field experienced by the nucleus and its gyromagnetic ratio γ.

$$\nu = \frac{\gamma}{2\pi} B_{eff} \tag{8}$$

The effective magnetic field around a nucleus results from the applied magnetic field B_0 being attenuated by induced magnetic fields created by circulation of surrounding electrons. These circulations are by themselves an effect of the applied magnetic field and therefore the induced fields are proportional to the applied field. This shielding or deshielding of the applied field is summarized in the nuclear shielding tensor σ [56].

$$B_{eff} = B_0(1 - \sigma) \tag{9}$$

As nuclear shielding is a tensor, the chemical shift is itself a tensor quantity. The tensor nature of the chemical shift takes into account, that the induced field may point in a different direction than the applied field. The subelement δ_{xz} therefore represents the part of the induced field pointing in the x-direction if the applied field is pointing in the z-direction. The tensor is commonly diagonalized which provides the 3 principal components of the chemical shift tensor δ_{xx}, δ_{yy}

and δ_{zz}. In liquid NMR, the fast molecular tumbling of the molecules leads to an effective averaging of the tensor components. Therefore, the measured isotropic chemical shift is the trace of the 3 principal components [57], [58].

$$\delta_{iso} = \frac{1}{3}(\delta_{xx} + \delta_{yy} + \delta_{zz}) \tag{10}$$

In the remainder of this thesis we will only consider the isotropic chemical shift as measured in liquid state NMR. As stated previously, the chemical shielding of a nucleus depends on its environment and can in turn be split into various contributions.

$$\sigma = \sigma_d + \sigma_p + \sigma_{rc} + \sigma_m + \sigma_{el} + \sigma_s \tag{11}$$

These contributions are, in order, diamangetic, paramagnetic, ring current, magnetic anisotropy, electric fields and the solvent effect. While it is not trivial to determine these components separately or calculate them from structural data, simplified models are often available.

The ring current effect is of special interest in this case. It leads to large chemical shift changes in the vicinity of aromatic systems due to additional strong induced magnetic fields. These shift changes are anisotropic and large, often in the order of 0.3-1.2 ppm and therefore ideally suited to identify aromatic interactions and gain information about their geometry.

3.2 RING CURRENT MODELS

Linus Pauling was the first one to assume that the chemical shielding of protons in aromatic hydrocarbons was due to a Larmor precession of the delocalized π-electrons [59]. While he did not coin the term ring current himself, to the knowledge of the author, that was John Pople, he nonetheless laid the groundwork for multiple models emerging in the 1960s and 1970s that tried to model the shielding and deshielding effects of aromatic systems and their surroundings. Today they have generally matured into the form of Eq. 12. In the current section we want to highlight the 3 most popular models [60], [61].

$$\Delta\delta_{ring} = i \times B \times G(r) \tag{12}$$

Here, G(r) is a geometric function dependent on r, the vector connecting the aromatic ring center and the nucleus under investigation. B are constants that represent the expected effect given a simple benzene ring and i is a scaling factor that represents the ratio between the expected effect in benzene to the effect of a specific aromatic system. I, in theory must therefore be determined separately for every aromatic system.

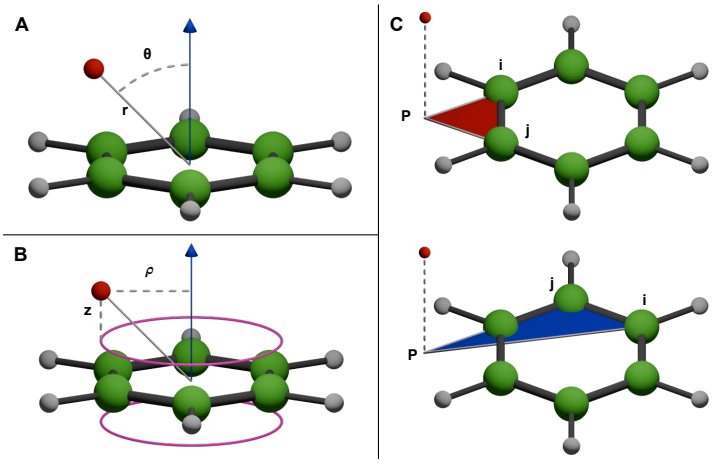


Figure 3: Comparison of different geometric functions. A: Parameters r and θ of the pople model. B: Parameters z and ρ of the WFJB model. C: Signed area S_{ij} of the triangle Pij for the HM model. Negative (red, clockwise) and positive (blue, counter-clockwise) areas depending on the order of the points in the triangle

The first and so-called classical model for the ring current effect was published by John Pople. He used Paulings model of circulating electrons and simplified into a point dipole located at the center of the aromatic ring. With this, he found that the average secondary field $\overline{H'}$ under fast rotational averaging of the ring plane can be estimated by Equ. 13 [62], [63].

$$\overline{H'} = \frac{e^2 a^2 H}{2mc^2 r^3} \tag{13}$$

Here, e is the unit charge of an electron, a is the diameter of the aromatic ring, H is the primary or applied magnetic field, m is the mass of an electron, c is the speed of light and r is the distance between the ring center and a proton. As Pople was mostly interested in protons that are directly connected to the aromatic system, he did not discuss any out-of-plane protons either intra- or intermolecular. McConnell later developed a more general model for long range shieldings based on the magnetic shielding model of Ramsey, which in the case of aromatic ring-currents includes out of plane protons. The Pople model to calculate the isotropic nuclear shielding constant in ppm as referenced in recent publications is shown in Equ. 14 [64]. n in this case is the number of electrons and θ is the angle between the vector r and the ring plane normal. All other symbols keep their previous meaning.

$$\Delta\sigma = 10^6 \times \frac{ne^2 a^2}{4\pi mc^2} \times \frac{3cos^2\theta - 1}{r^3} \tag{14}$$

Later, a model introduced by Waugh and Fessenden and further refined by Johnson and Bovey modeled the ring current not as a single

point dipole, but rather as explicit current carrying circuits (WFJB model) [65], [66]. They also converted the position of the nucleus of interest to cylindrical coordinates expressed in units of a, the diameter of the ring. As it is known that the maximum of the current density is not located in the plane of the ring, they concluded that it is necessary to consider 2 such parallel rings simultaneously which are separated by equal distances from the aromatic ring plane.

$$\Delta\sigma = 10^6 \times \frac{ne^2}{12\pi mc^2 a} \times \sum_{p=1}^{2} \left(\frac{1}{\sqrt{(1+\rho)^2 + z_p^2}} \left(K(k) + \frac{1-\rho^2-z_p^2}{(1-\rho)^2+z_p^2} E(k) \right) \right) \tag{15}$$

$$k^2 = \frac{4\rho}{(1+\rho)^2 + z_p^2} \tag{16}$$

Here, K and E are the complete elliptical integrals of the first and second kind respectively with their modulus k^2. The index p is either the first or the second current carrying ring with a certain distance from the aromatic ring. ρ and z_p are the cylindrical coordinates in units of a. In the case of z_p, the distance is measured from the ring under consideration. One downside of this model is, that it has an additional variable to be determined in the form of the optimal displacement of the current carrying rings from the aromatic ring. Johsnon and Bovey determined the optimal value for benzene as 1.28 $\mathring{A}$. Multiple claims about optimal values can be found in the literature and some perform no separation of the rings at all [67], [68].

There are a number of ring-current models that are originally based on Londons [69] quantum mechanical theory of π-electron ring currents and as an extension on the molecular orbital theory of Hückel [70]. These methods are therefore commonly termed the quantum mechanical ring-current models. The best known model among them is probably the one introduced by Haigh and Mallion (HM) [71].

$$\Delta\sigma = 10^6 \times KJ \times \sum_{ij} S_{ij} \left(\frac{1}{r_i^3} + \frac{1}{r_j^3} \right) \tag{17}$$

Here, K is a constant, J is a quantum mechanical quantity that is supposed to be calculated from Hückel wave-functions. For the geometric term, an iteration over each consecutive atom pair i,j is performed around the aromatic ring. For each pair, r_i and r_j are the distances between the respective atoms and the proton. Given a proton P, S_{ij} is the signed area of the triangle P_{ij} after projection onto the plane of the aromatic ring. The advantage of this method is, that it takes the additional pentagonal or hexagonal symmetry of the aromatic system into account, while the classical theories assume that they are perfectly circular. The factor J should ideally be calculated

using QM-methods. Most commonly, and also by Haigh and Mallion, it is simply fit from experimental data.

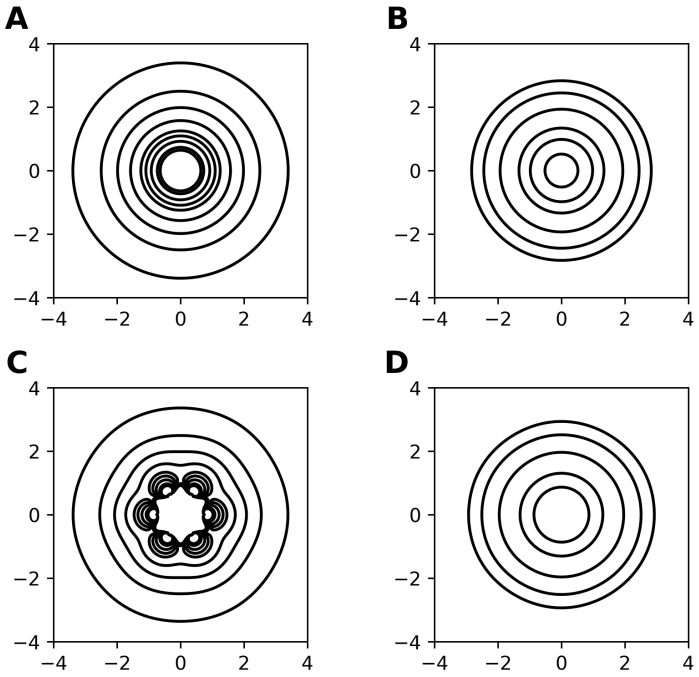


Figure 4: Comparison of the geometric function between the Pople and HM models. A+B: Pople model. C+D: HM model. In the ring plane (A+C) we see the additional 6-fold symmetry of the HM model compared to the Pople model. This additional symmetry plays no major role at a distance of 2 $\AA$ above the ring plane (B+C).

The WFJB and HM models have traditionally been more popular compared to the Pople model. The reason for this may be that the authors of these 2 models published lookup tables of precalculated values for given geometries between aromatic rings and protons. This was particularly useful, as calculation of these values with the equipment of the time was far from convenient. While the Pople model also has the reputation of being the most basic and imprecise of the three, the discrepancy between the Pople and the HM model is small. Even Haigh and Mallion themselves conclude that at longer distances, the difference between the 2 models is negligible [61]. The simplicity of the Pople model on the other hand comes with a number of advantages. It does not require any additional parameters to be fit, like the ring separation for the WFJB model and it is less dependent on the ideal conformation of the ring atoms, like the HM model. Additionally, its geometric term resembling a $P_2(cos\theta)$ Legendre polynomial is both intuitive and relatively straightforward to differentiate compared to the other models. This is especially important, if one wants to use the model to calculate the gradient of a simulated annealing simulation during structure calculation.

3.3 CH-π INTERACTIONS

Single molecular interactions are of prime importance in structural biology and effectively tune the overall interaction of many systems. In the end, whether a protein folds into its native conformation or a ligand binds to its target, depends on a surprisingly small net energy difference between states [72]. Between these states we find large differences in enthalpic and entropic energy contributions that compensate one another. While salt bridges and strong hydrogen bonds between polar groups are among the strongest non-covalent interactions, other weaker contributions should not be neglected [73], [74]. A comparison of the number of protein-ligand interactions found in the PDB, shows that hydrophobic interactions, while overall weaker than polar interactions, are by comparison ubiquitous and in sum have a major impact on the overall interaction energy of the system. Among these hydrophobic interactions, the vast majority in turn is formed between aromatic and aliphatic groups [75].

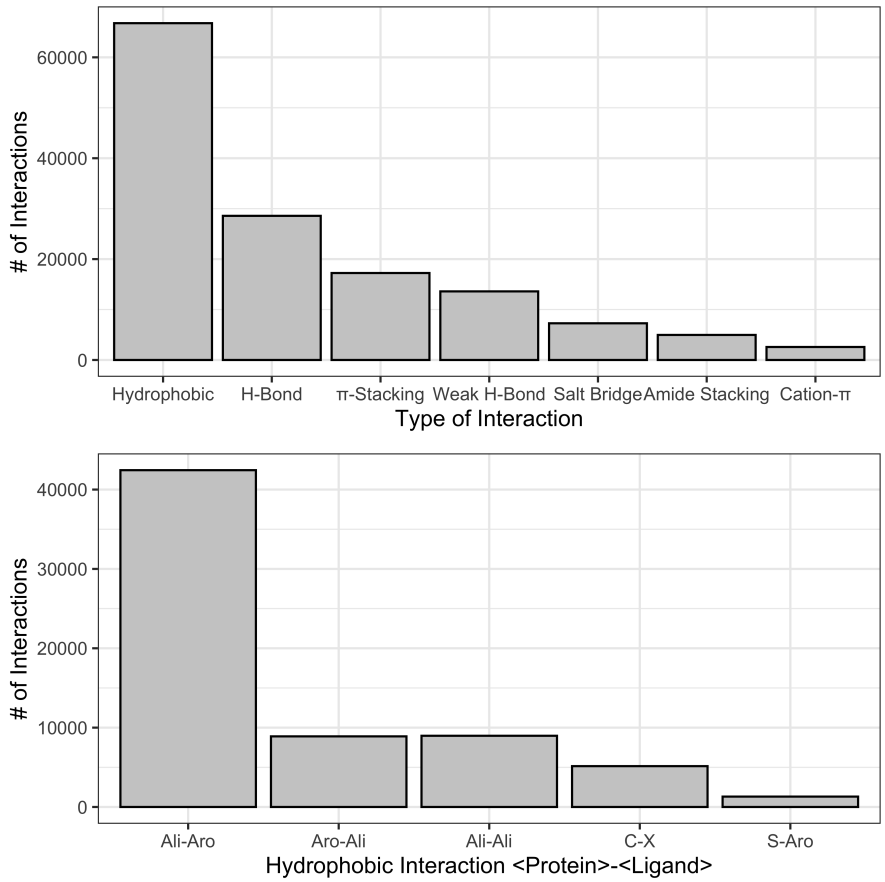


Figure 5: Protein-ligand interactions found in the PDB. Overall interactions (top) and hydrophobic interactions in detail (bottom). Data was extracted from [75]

These so called CH-π interactions are especially common in protein-ligand systems due to the large propensity of aromatic groups in drug-like compounds [76], [77]. While they are often classified as hydrophobic interactions, they do have the character of a weak non-canonical hydrogen bond in which the alipahtic CH group is the donor and the electron cloud of the aromatic ring is the acceptor [78]. As all hydrogen bonds, they are highly directional and the interaction geometries over all CH-π contacts in the PDB shows a clear preference for certain orientations (see Publication 2, Fig. 3BC) [79]. As aromatic systems induce a large anisotropic chemical shift change in surrounding nuclei, NMR is the ideal method to detect and quantify these types of interactions. Simple chemical shift perturbation (CSP) measurements via 2D HSQCs are sufficient to quantify the induced chemical shift. Calculation of the experimental values using the aforementioned ring-current models then provide direct and indirect information on the bound geometry and interaction energies. This combination of trivial experiments and quick analysis via cheap models can act as a powerful tool for medicinal chemists during drug discovery.

4

SUMMARY OF THE THESIS

This thesis and its presented papers highlight very different ways of applying NMR in the context of structural biology. In the first paper, we use paramagnetic relaxation techniques to resolve the structure of the heavily researched protein α-synuclein that is formed upon interactions with mimics of biological membranes. We combined PRE and PRI NMR techniques with mass spectrometry cross-linking experiments to solve both the monomer and the dimer structure. In the second publication, we use chemical shifts, the most fundamental observable in NMR, to compare the change induced by ligand aromatic rings in the target protein to detect CH-π interaction in protein ligand systems and subsequently fit the most probable position and orientation of aromatic ring centers based on both chemical shifts and methyl-π interaction geometries in the Protein Data Bank.

Part II

SHOWCASE OF PUBLISHED WORK

The central part of this cumulative thesis consists of two self-contained scientific articles revolving both NMR based structure determination as well as the characterization of protein-ligand systems. All articles have been published in peer-reviewed scientific journals. Aside from the continuous page numbering, the original layouts are fully reproduced and should be considered intellectual properties of the respective journals. Where possible, supporting figures and tables are included in the original format as well. Supporting data and code that could not be included in this thesis can be found in the online versions of the articles. Their respective digital object identifiers (DOIs) are provided together with a short statement of personal contribution at the beginning of each chapter.

5

PRE AND PRI BASED DETERMINATION OF A COMPACT α-SYNUCLEIN STRUCTURE

T. C. Schwarz, A. Beier, K. Ledolter, *et al.*, "High-resolution structural information of membrane-bound α-synuclein provides insight into the moa of the anti-parkinson drug ucb0599," *Proceedings of the National Academy of Sciences*, vol. 120, no. 15, e2201910120, 2023 [49]

DOI: 10.1073/pnas.2201910120

PERSONAL CONTRIBUTION

Protein production, conceptualization , data analysis, validation, co-writing, shared first author with T. C. Schwarz

PNAS
RESEARCH ARTICLE | BIOCHEMISTRY
BIOPHYSICS AND COMPUTATIONAL BIOLOGY
OPEN ACCESS

High-resolution structural information of membrane-bound α-synuclein provides insight into the MoA of the anti-Parkinson drug UCB0599

Thomas C. Schwarz[a,1,2], Andreas Beier[b,1], Karin Ledolter[a], Thomas Gossenreiter[c], Theresa Höfurthner[b], Markus Hartl[c], Terry S. Baker[d,3], Richard J. Taylor[d], and Robert Konrat[a,2]

Edited by David Baker, University of Washington, Seattle, WA; received February 4, 2022; accepted February 13, 2023

α-synuclein (αS) is an intrinsically disordered protein whose functional ambivalence and protein structural plasticity are iconic. Coordinated protein recruitment ensures proper vesicle dynamics at the synaptic cleft, while deregulated oligomerization on cellular membranes contributes to cell damage and Parkinson's disease (PD). Despite the protein's pathophysiological relevance, structural knowledge is limited. Here, we employ NMR spectroscopy and chemical cross-link mass spectrometry on $^{14}N/^{15}N$-labeled αS mixtures to provide for the first time high-resolution structural information of the membrane-bound oligomeric state of αS and demonstrate that in this state, αS samples a surprisingly small conformational space. Interestingly, the study locates familial Parkinson's disease mutants at the interface between individual αS monomers and reveals different oligomerization processes depending on whether oligomerization occurs on the same membrane surface (cis) or between αS initially attached to different membrane particles (trans). The explanatory power of the obtained high-resolution structural model is used to help determine the mode-of-actionof UCB0599. Here, it is shown that the ligand changes the ensemble of membrane-bound structures, which helps to explain the success this compound, currently being tested in Parkinson's disease patients in a phase 2 trial, has had in animal models of PD.

α-synuclein | oligomeric structure | XL-MS | paramagnetic NMR spectroscopy | UCB0599

α-synuclein (αS) plays a prominent role in neurodegenerative disorders such as Parkinson's disease and Lewy body dementia (1) by changing its oligomeric state and ultimately forming aggregates (2). The toxic αS aggregates interact with the lipid bilayer of cells and have been shown to incorporate bilayer constituents and disrupt the bilayers themselves (3–5), Lewy bodies also to a large degree consist of αS and membranes (6). In addition, binding of αS to the surface of membranes has been demonstrated to promote further aggregation in vitro (7, 8) and the mechanism of this aggregation has been investigated recently (9), although a decrease of aggregation can also be observed (10). However, regulated interaction with lipid membranes is crucial for the proper functionality (i.e., neuronal differentiation and apoptosis) (3, 4) of αS, as it is involved in SNARE complex assembly (11, 12). In addition, it regulates exocytosis by promoting dilation of vesicle fusion pores (13). Thus, detailed knowledge about the structural properties of the various oligomeric states of αS at the membrane surface is highly desirable for an understanding of its biological role and to guide drug research which aims to modify αS aggregation behavior at the membrane (14).

αS has been shown to exist in different conformations depending on the environmental context (2). Purified from mouse brain, it is a monomeric, intrinsically disordered protein (IDP) (15). In vitro studies performed on recombinantly expressed αS with circular dichroism (16) and small-angle X-ray scattering (17) also show it to be an IDP. The technique delivering the richest data on IDPs is NMR (18), and both in vitro and in-cell NMR experiments (19) have demonstrated that monomeric αS in solution is mostly disordered.

When studied in the context of membranes, monomeric αS adopts a predominantly α-helical structure. This has been demonstrated by circular dichroism, Förster resonance energy transfer, electron paramagnetic resonance, and solid-state as well as solution NMR-based studies (4, 16, 20–24). High-resolution structures were obtained for αS bound to sodium dodecyl sulfate (SDS) or sodium lauroyl sarcosinate-based micelles, although the orientation of helices as well as the existence of several different states is still a matter of discussion (24–28).

In mature fibrils, much of the αS sequence adopts a β-sheet fold as shown by both cryogenic electron microscopy and solid-state NMR studies (29, 30) with fibril growth including similar charge-based interactions as those observed for membrane binding

Significance

Although the pathophysiological relevance of α-synuclein in Parkinson's disease (PD) is clearly established, its natural function and the mechanism by which it causes toxicity are not fully understood. Recruitment of α-synuclein to membranes is required for correct vesicle dynamics, while altered oligomerization at membranes can contribute to PD. We present a high-resolution structural model of α-synuclein in its oligomeric, membrane-bound state based on solution NMR and chemical cross-linking mass spectrometry (XL-MS) restraints. The obtained structures prove useful in explaining the consequences of well-known familial PD mutations and revealed a plausible mode-of-action for UCB0599, a drug designed to interfere with membrane-bound α-synuclein showing disease-modifying potential. We expect this structural information to support future research furthering our understanding of α-synuclein's intricate functions.

This article is a PNAS Direct Submission.

[1]T.C.S. and A.B. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: t.schwarz@univie.ac.at or robert.konrat@univie.ac.at.

[3]Deceased.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2201910120/-/DCSupplemental.

Published April 7, 2023.

(31). Oligomer formation at the membrane has been studied by several different methods (32, 33) and various different relative orientations of the α-helical structures have been proposed (34, 35). In addition, some previous work has indicated that multimers assemble from dimers in vivo, (35) while other work has shown sequential accumulation of oligomers (36).

In this study, we employ an integrated structural biology approach combining NMR spectroscopy and chemical cross-linking mass spectrometry (XL-MS) using differently isotope-labeled forms of αS to provide for the first time detailed structural information for αS oligomers formed at and bound to membrane vesicles. We show that the phase 2 compound UCB0599 modulates these oligomers and provides a plausible mode-of-action (MoA) for this new drug.

Results

Determination of Oligomeric State Distribution. Here, we used XL-MS to probe the structural characteristics of αS in different states: monomeric SDS-micelle bound, monomeric bicelle bound, 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (POPG) based liposome bound, and bicelle-aggregated. The oligomer distribution of αS on membrane surface mimics (POPG-liposomes or bicelle-derived aggregates) was determined by cross-linking the protein while it is bound to the respective membrane mimic using the zero-length cross-linker 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and detecting the oligomeric state by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) (Fig. 1*A*). It should be noted that the experimental αS to membrane concentration ratios are expected to also occur in vivo (see *SI Appendix* for details). To confirm the oligomeric status seen in PAGE samples, a set of experiments was performed where the EDC cross-linked αS species were removed from the liposomes using a small-molecule compound to displace αS from POPG-based surfaces (37). The protein was separated from these liposomes by subsequent ultracentrifugation, and monomeric species present in the sample were removed by size exclusion chromatography (SEC) of the resultant supernatant. The soluble multimers obtained were subjected to intact-mass spectrometry (Fig. 1*C*). Most importantly, monomer, dimer, and trimer masses were clearly detected (Fig. 1*B*). From the observed "ladder" structure of the different oligomeric bands, it can be concluded that αS aggregation on membrane surface mimics proceeds via a stepwise oligomerization mechanism (linear growth/propagation).

Structural Probing of Membrane-Bound αS Using XL-MS. In a next step, structural investigations of liposome surface-bound αS were performed using differentially isotope-labeled (^{14}N and ^{15}N) proteins similar to the methodology demonstrated in ref. 38, enabling the use of an experimental strategy illustrated in Fig. 1*C*. Here, the analysis of individual isotopomers (^{14}N–^{14}N, ^{14}N–^{15}N, ^{15}N–^{14}N, and ^{15}N–^{15}N) allows for the discrimination of intramolecular and intermolecular cross-links. While intramolecular cross-links are exclusively observed for homoisotopomers (^{14}N–^{14}N and ^{15}N–^{15}N), intermolecular cross-links can be found for both homo- (^{14}N–^{14}N and ^{15}N–^{15}N) and hetero-isotopomeric (^{14}N–^{15}N and ^{15}N–^{14}N) forms. The prototypical MS spectrum of a cross-linked peptide shown in

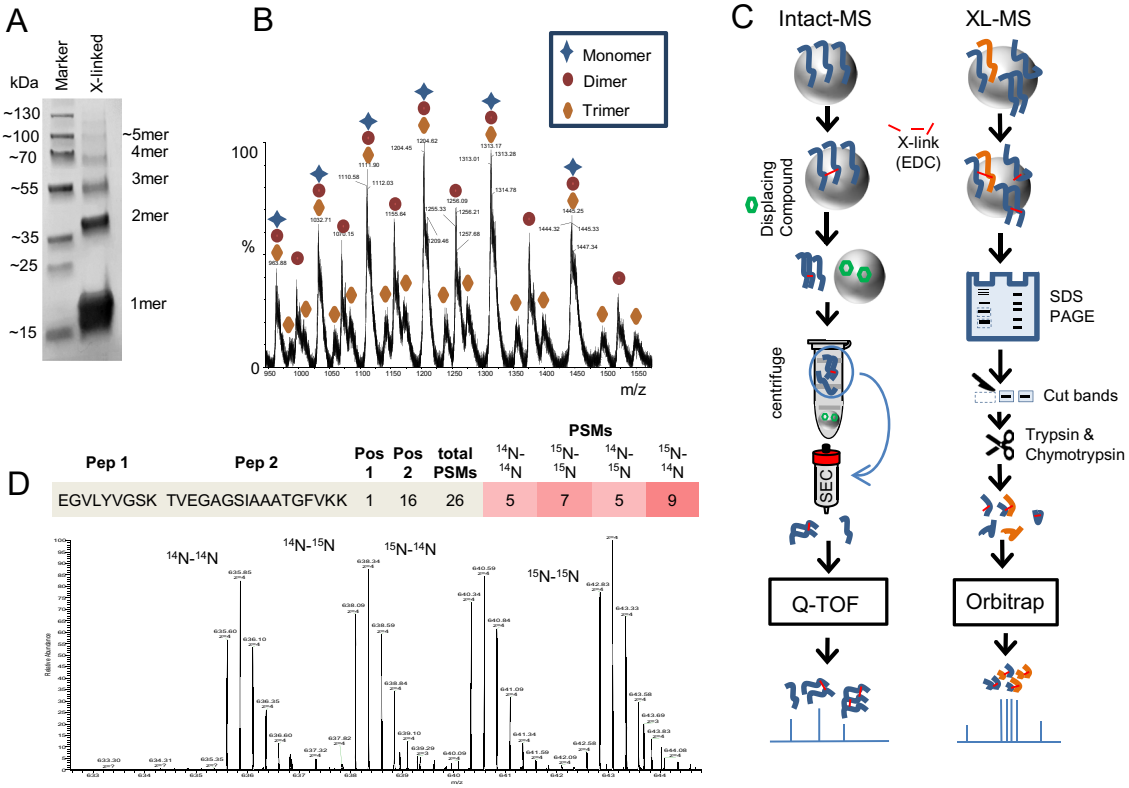


Fig. 1. Cross-linking of α-synuclein at the membrane. (*A*) SDS-PAGE (4 to 20%) of αS cross-linked in the presence of POPG-based liposomes and subsequent removal of liposomes followed by SEC. Molecular mass increases from monomers up to tetramers are clearly visible. Higher oligomers are present to a smaller degree. (*B*) Selection of m/z species showing the masses of monomeric, dimeric, and trimeric forms of cross-linked αS as detected by intact mass spectrometry post SEC. (*C*) Schemes for cross-linking and purification as used for detection of oligomers in intact mass (*Left*) and intermolecular/intramolecular cross-links of peptides (*Right*). A previously reported molecule that releases αS from membrane surfaces (37) is depicted as a green hexagon. (*D*) Spectrum demonstrating how inter- and intra-molecular cross-links can be distinguished through isotopomeric labeling. At the *Top*, the two different peptides detected are indicated with their cross-linking positions as well as the detected peptide spectrum matches for each combination of isotopomers. The mass spectrum allowing for this distinction is shown below.

Fig. 1*D* demonstrates that the different isotopomers can be resolved and discriminated by this approach. In order to facilitate data analysis, "crosslink-contact maps" of αS were generated based on the number of peptide spectrum matches (PSMs) found for specific cross-linked residue pairs in a similar fashion as published previously (28). As this approach focuses on regions with high cross-link density, it reduces the risk of unspecific encounters between αS molecules contributing significantly to the analysis (*SI Appendix*, Fig. S14*C*). Fig. 2 *A*–*D* shows intramolecular contacts found for monomeric αS obtained for different membrane-bound states (liposome, SDS-micelle, monomeric bicelle-bound, and aggregates induced by bicelles). The majority of cross-links obtained for the various monomeric membrane-bound states of αS (a-c) are close to the diagonal and clustered at the N-terminus indicating relatively extended or slightly bent conformations. Intramolecular (homoisotopomeric) links obtained from dimer bands of αS bound to liposomes (Fig. 2*E*) show an overall similar pattern as the monomer band, but are enriched in specific regions. In our statistical analysis of these maps, we see the similarity of the pattern in a low Cohen's d value, while the statistically significant change in the pattern is evaluated by the low *P* value (*SI Appendix*, Fig. S22). In the dimer band, more cross-link PSMs are observed for segments (55–65) to (90–105) and (30–40) to (90–105), indicating the higher abundance of more compact αS structures in the liposome-bound dimeric state as compared to the monomeric state. Even higher abundancies of these links were found in the trimer bands (*SI Appendix*, Fig. S13). In contrast, the intramolecular contact map for αS dimers obtained for the bicelle-aggregated state (Fig. 2*G*) is largely comparable to the bicelle-bound monomeric state (Fig. 2*D*), which differs in incubation time of the mix prior to cross-linking (*Methods*). A detailed statistical (quantitative) comparison of the obtained cross-link maps was performed by adapting a bootstrap methodology and is detailed in *SI Appendix*, *Methods* and Fig. S22.

Mapping of the Oligomerization Interface. Residue-resolved information about intermolecular contacts and the interaction interface was obtained by analyzing the heteroisotopomeric ^{14}N–^{15}N peptide cross-links. For the liposome-bound dimer state intermolecular contacts are predominantly observed for regions (90–100) to (50–65) as well as (55–65) to (10–15) and (40–50) to (25–35) (Fig. 2*F*). These interaction regions coincide with regions containing several of the point mutations involved in early-onset parkinsonism and have been used for the design of small-molecule therapeutic compounds (3, 39). The intermolecular interaction pattern of bicelle-aggregated αS is distinctly different (Fig. 2*H* and *SI Appendix*, Fig. S22). Since the bicelle-bound form of αS is monomeric, aggregation necessarily proceeds via interbicelle interactions (trans) rather than between two αS proteins bound to the same surface as is the case in liposomes (cis). Thus, and most importantly, the cross-link data provide evidence for different oligomeric species resulting from aggregation caused by intermolecular interactions between αS molecules bound to the same vesicle (intraliposome/cis) or located on different vesicles (interbicelle/trans), respectively.

NMR-Based High-Resolution Structure of Membrane-Bound αS. Signal assignment of αS was achieved for both the SDS-micelle (BMRB-ID 50996) and bicelle-bound states (BMRB-ID 51148)

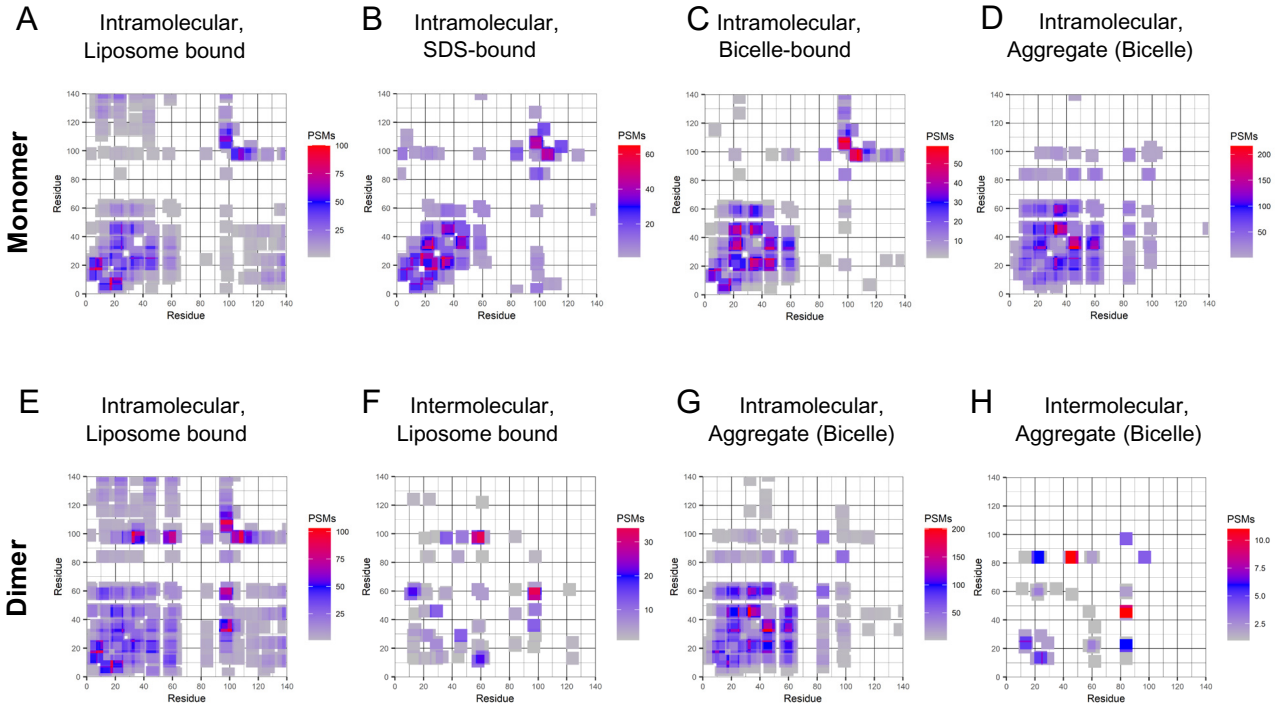


Fig. 2. Cross-linking patterns in α-synuclein are sensitive to conditions used. (*A*–*D*) Cross-links detected in the monomer-band of αS after incubation with various membrane mimics. The number of PSMs found in a four-residue window is indicated in the color scales. While some long-distance contacts (130–20) are observed in the liposome-bound monomer (*A*) and medium distances (60–35) in bicelles (*C*), SDS bound monomers do not show a significant amount of long-range interactions (*B*). Both intra- and inter-molecular contacts were detected for the bound form from the dimer band of αS (*E* and *F*). Interestingly, the intramolecular cross-linking pattern changed substantially between dimeric, bound αS (*E*) and the monomeric state (*A*) (statistical measures are available in *SI Appendix*, Fig. S22). Intermolecular contacts detected indicate two main regions for dimer formation (60–15 and 95–60). Note that loop-links (cross-links within a single peptide fragment) were omitted in our analysis, thus no diagonal peaks are expected. Additionally, residues located in the segment 60–80 are not amenable to the cross-linking reagents used and therefore no cross-links are observed for these residues. (*G* and *H*) When incubated with bicelles αS can form aggregates. These show dissimilar cross-linking patterns as compared to those on liposomes. Intramolecular contacts both in the monomer as well as the dimer bands (*D* and *G*) were more plentiful and show fewer medium range contacts (95–60 and 95–35). Intermolecular contacts (*H*), however, are less abundant and are changed drastically as compared to the liposome-bound dimer.

(28). The similarity between SDS-bound secondary structure propensity (SSP) data and that obtained for the bicelle-bound state (*SI Appendix*, Fig. S1) suggests that these structures are independent from membrane curvature and may serve as a valid model system for the membrane-bound state of αS. Due to its more favorable NMR relaxation properties, the SDS-bound state was selected for further structural characterization. NMR signal assignment of the membrane-mimic-bound state(s) in our conditions (323 K and pH = 5.5) revealed the existence of three α-helical segments largely comparable to the already reported SDS-micelle-bound state (23). In this study, we employ different NMR experiments than used previously in order to provide a larger and more comprehensive NMR data set for structural characterization of membrane-bound αS. The previous NMR study (23) largely relied on measurements of nuclear Overhauser effects (NOEs) and residual dipolar couplings. In our study, we significantly increased the number of long-range contact-dependent NMR observables and consequently the quantity of such experimental constraints in the structure calculation protocol. We employed paramagnetic relaxation enhancements (PRE) and interference (PRI) as these experiments have exquisite sensitivity and can probe even sparsely populated conformational substates (40), in contrast to conventional PREs. While the previous NMR study only involved a single PRE measurement utilizing changes in intensity, we recorded PRE-rate profiles for three independent spin label positions (*SI Appendix*, Fig. S3). Moreover, the involvement of two additional PRI datasets (*SI Appendix*, Fig. S2) substantially improves the available data, as PRI experiments also provide angular constraints. The (PRE) relaxation rates obtained were not compatible with expected rates calculated from published data (*SI Appendix*, Fig. S3) (23), implying that additional compact conformational states must exist for αS. For the structure determination of these state(s), we adapted the well-established Xplor-NIH protein structure calculation protocol to incorporate PRE-derived distance and PRI-derived angular constraints (*Methods*). Additionally, NMR protein backbone chemical shift data were converted into secondary structure constraints. In order to derive an unbiased set of representative structures, an extended chain representation was used as a starting point (for details, see *Methods*). In Fig. 3, α-carbon to α-carbon (CA-CA) rmsd-based clustering of the 10 best solutions together with representative structures of each cluster and variations within each cluster are shown. Interestingly, one of the solutions largely corresponds to the previously determined structure (23) (cluster C in Fig. 3). Most importantly, however, the two other clusters of Fig. 3 comprise structures with a different relative orientation of the three α-helices. It is important to note that the detection of these compact states was only possible due to the substantially increased number of experimental (distance and angular) constraints and the measurement of PRI data as a sensitive probe for (globally) compacted states (40) stemming from the strong distance dependency of PREs (r^{-6}). The quality of the obtained structures was assessed by back-calculation of PRE rates and comparison with experimental rates (*SI Appendix*, Fig. S19 *A–C*). The structures were further corroborated by comparison with intramolecular cross-link maps. Specifically, the distance map obtained for cluster B (Fig. 3, *Right*) shows shorter distances between the region around residue 80 to 100 and the N-terminus, which fits cross-links observed with SDS micelles (Fig. 2*B*), while the calculated distance map for cluster A (Fig. 3, *Right*) shows more similarity to the intramolecular pattern found for dimeric αS bound to liposomes (Fig. 2*E* and *SI Appendix*, Figs. S8 and S20). A thorough and quantitative assessment was performed by applying the recently developed XLmap

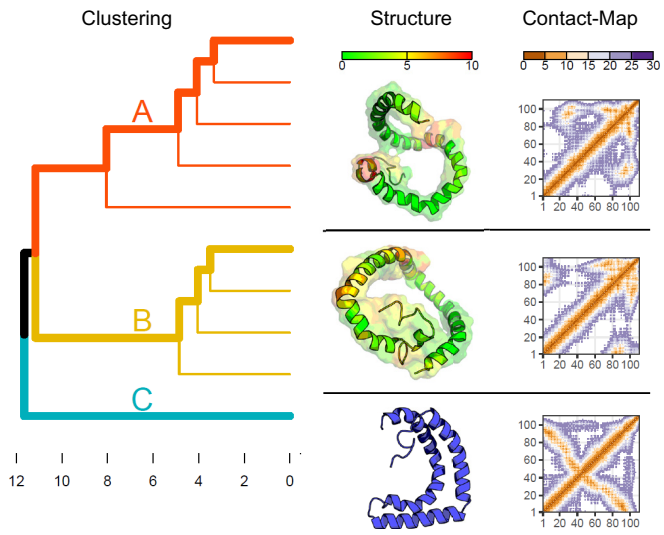


Fig. 3. The PRE/PRI/SSP derived compact ensemble of α-synuclein shares features with the multimer forming ensemble at the membrane. (*Left*) CA-CA rmsd clustering of the 10 lowest energy structures obtained from monomer calculations. Residues 101 to 140 are excluded from the analysis as they are not constrained by experimental data. Three major clusters are indicated by different colors (A: red, B: dark yellow and C: light blue), the cluster representatives are highlighted by bold lines, rmsd [Å] is indicated at the bottom. (*Middle*) Representative 3D cluster structures. The surface representation includes all structures contained in the cluster aligned to the cluster representative and is colored by the maximum CA-CA distance between structures at each position. Since cluster C contains only a single structure, no common surface representation is given. (*Right*) CA-CA based contact map for the individual clusters. The color codes for the maximum distance in surface representations (*Middle*) and the distances in contact maps (*Right*) are indicated at the *Top* of each column (both given in [Å]).

methodology (41), an R package to visualize and score protein structure models based on sites of protein cross-linking. The calculated CMScores (*SI Appendix*, Fig. S20) convincingly (and quantitatively) corroborate the qualitative fits between structure and experiment (41). Taken together, the PRE and PRI experiments clearly demonstrate the presence of (at least) three conformational substates with slightly different compaction and relative orientations of α-helices. Structures within cluster C (and partly cluster B) are in good agreement with the previously determined SDS-micelle bound structure of αS (23) and likely represent the dominant state on SDS surfaces and possibly monomers on liposomes. Cluster A, however, shows structural similarity to liposome-bound dimeric forms of αS (obtained from XL-MS data) and points to the relevance of these states for membrane-induced oligomerization processes.

Computational Modeling of Wild-Type αS Oligomers. High-resolution models for αS dimers and higher oligomers were obtained using intermolecular XL-MS data (Fig. 2) within the Xplor-NIH structure calculation framework. Given the agreement of intramolecular XL-MS data obtained from liposome-bound αS with NMR data of SDS-bound monomer structures (Figs. 2 and 3 and *SI Appendix*, Fig. S20), NMR structural constraints derived for the SDS-bound state were employed together with XL-MS intermolecular restraints in the calculation of the oligomers (the calculation protocol is outlined in *SI Appendix*, Fig. S23). Ambiguous constraints were defined based on the experimental PSMs. Specifically, the following contacts were used in the calculation: [(57, 61)–(96, 97), (10, 13)–(57, 60), (28)–(45)] (more details can be found in the *Methods* section). Possible structures of αS dimers were calculated assuming internal symmetry (two αS molecules with

identical monomer structures in the dimer). This assumption was made as the overall conservation of cross-link patterns in dimers and trimers (*SI Appendix*, Fig. S13) suggests largely similar 3D structures of the different oligomers thus pointing toward a propagation mechanism which progressively links more or less structurally equivalent αS monomers. Although cluster A (Fig. 3) best reproduces the intramolecular cross-links obtained from the liposome-bound dimeric form of αS, (*SI Appendix*, Fig. S9*A* and S20) independent calculations were also performed using representative monomer structures of clusters A, B, and C (Fig. 3) as input. While the starting structures of monomer A and B are able to interconvert during the dimer structure calculation, the starting structure C—which closely resembles the previously determined SDS-micelle–bound state (23)—is locked in its starting conformation and likely represents the most populated state. Importantly, the incorporation of the new intermolecular constraints did not change the monomer structures (A and B) and thus indicate that the SDS-bound compact monomers are largely preserved in liposome-bound oligomers.

Interestingly, independent of the starting (monomer A or B) structure, only two consistent solutions for the dimer structure are found (Fig. 4). While both dimer structures share the same interaction interface between αS molecules, the relative orientation of different αS monomers in the dimer varies. This is illustrated in Fig. 4 by depicting the first major axis of the aligned monomer unit and the respective dimer solutions. Importantly, the documented familial Parkinson mutations A30P and A53T are located at the dimer interface.

As αS forms oligomers on membranes, the obtained dimer structures were used to generate structural models for higher oligomers. Note that the chemical cross-link data (Fig. 1) indicate that αS oligomers are formed by progressive propagation on the membrane surface. Utilizing this information, oligomer structures were generated by joining αS dimers into larger chains, while retaining the inherent symmetry (*Methods*). Starting from the two

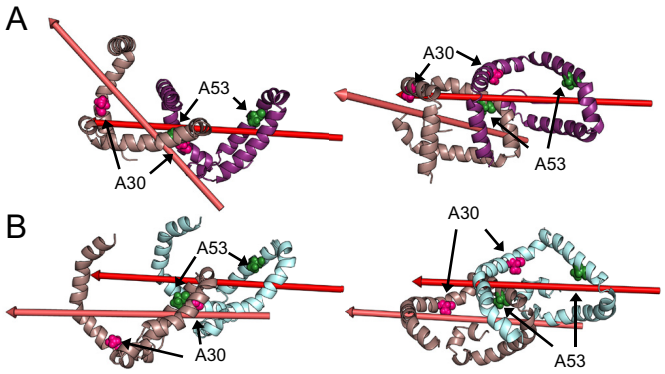


Fig. 4. Dimer structures derived from intermolecular XL-MS data. (*A* and *B*) Two prototypical solutions of the dimer calculation resembling monomer cluster A. While the interface is similar in both cases, slight deviations lead to a noticeable kink between subunits as demonstrated by the major axis of the monomers. While (*A*) shows an angle between the major axis of the two subunits (*B*) shows a near parallel offset. Mutations associated with familial forms of Parkinson's disease and studied by XL-MS here are indicated by pink (A30) and green (A53) spheres and are located at the interface.

dimer structures (Fig. 4), we obtain oligomeric species with distinct symmetry properties: circular structures (Fig. 5 *A* and *B*) or elongated structures with helical periodicity along the translational (propagation) axis (Fig. 5 *C*–*F*). It should be noted, however, that the structures shown in Fig. 5 are representative solutions and slightly different structures can result from minute changes in the orientation across the αS dimer interface. For example, for the circular species also other ring sizes can be obtained and, likewise, for the elongated structure, a range of twist angles are possible. Depending on the twist angle, linear αS oligomers can be continuously attached to the membrane surface (Fig. 5 *C* and *D*) or adjacent αS monomers expose their membrane attachment sites (N-terminus) to opposite sides (Fig. 5 *E* and *F*, up–down arrangement). Thus, growing αS aggregates could either cover the surface

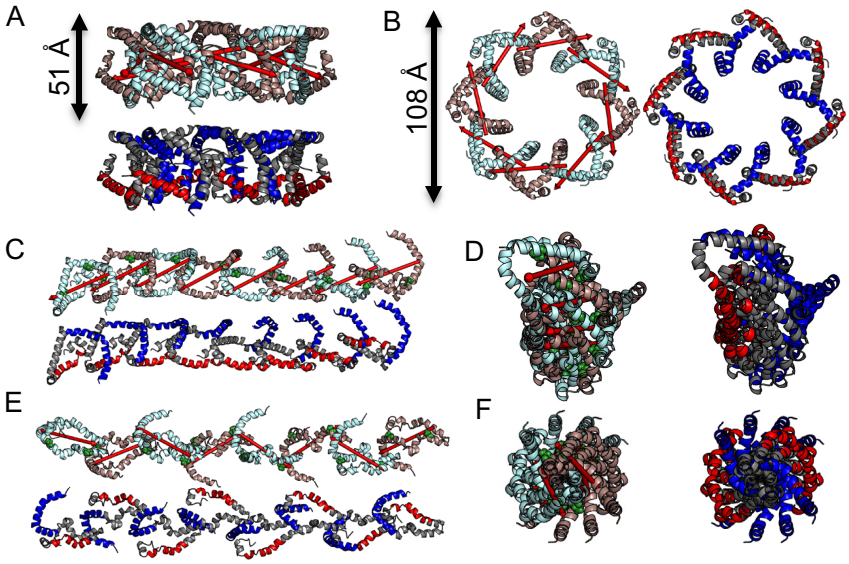


Fig. 5. Propagation of α-synuclein into higher oligomers. Deviations of the interface demonstrated in Fig. 4 lead to distinct oligomeric structures upon further addition of monomers. Dependent on the kink between the major axes, this results in either circular (*A* and *B*) or linear/helical/extended (*C*–*F*) oligomeric structures. Individual αS molecules in the oligomers are coloured alternatingly (cyan brown) in one representation for each view [*Top* (*A*, *C*, and *E*) and *Left* (*B*, *D*, and *F*)]. Regions identified to interact with the polar head groups of the membrane surface (1–36: blue) or lipid tails within the inner layer of the lipid bilayer (70–88: red) in ref. 32 are indicated (*Bottom* (*A*, *C*, and *E*) and *Right* (*B*, *D*, and *F*)]. Ring-like structures have a height and diameter of 51 Å and 108 Å, respectively, which is in good agreement with dimensions extracted from EM-particle classifications. (*C*–*F*) Two representative structures of extended oligomers. While both oligomers form extended chains upon continuous addition of monomers, the orientation of monomers is different. Small (*C* and *D*) or large (*E* and *F*) rotational angles result in distinctly different interfaces either forming an *amphiphilic-like* [surface (1–36, blue) and inner layer (70–88, red) interactions on opposite sides (*C* and *D*)] or alternating [surface and inner layer interactions are alternating (*E* and *F*)] membrane interaction interface.

of one membrane vesicle or bridge two different vesicles. Most importantly, the obtained circular αS oligomer structure (Fig. 5 *A* and *B*) is supported by previous experimental data. For example, comparable pore diameters (~5 nm as compared to ~4 nm in Fig. 5*B*) were observed in EM studies of liposome-bound αS (34). Further support comes from distance measurements based ondouble electron-electron resonance experiments which are also compatible with our structures (42) and atomic force microscopy (AFM) (43) which detects particles of about ~10 nm as compared to ~11 nm in Fig. 5*B*.

It is also instructive to relate our findings to published solid-state NMR data of oligomeric αS (32) which revealed the regions of αS either interacting with the lipid tail (residues 70 to 88) or the polar head groups (residues 1 to 36) of membranes. For example, insensitive nuclei enhancement by polarization transfer and dipolar-assisted rotational resonance spectra demonstrated differential mobilities for the two segments (32). The respective regions are depicted in Fig. 5 (rigid part: 70 to 88, red; mobile part: 1 to 36, blue). In the circular species, the segment demonstrated to interact with the lipid tail (70 to 88, red) is located on one side of αS while the N-terminal helix (1 to 36, blue), which is in contact with the lipid head-group, is positioned toward the top and center of the ring-like structure. The elongated form also shows a distinct distribution of lipid and head group interacting residues. Depending on the twist angle, the oligomeric form can display an amphiphilic character with both hydrophilic and hydrophobic surface. Alternatively, lipid tail and head group interaction motifs can alternate as the oligomeric propagation proceeds.

The Effect of UCB0599 on Membrane-Bound Oligomeric αS. The obtained detailed structural information presented here allowed us to also gain insights into the mechanism of action of the small compound UCB0599 (*SI Appendix*, Fig. S18) developed to target αS in its membrane-bound state. This compound with disease-modifying potential and an acceptable safety profile (44) has entered phase 2 clinical trials as a different therapeutic strategy to combat Parkinson's disease. The design strategy for the compound was based on the interference with and ultimate release of αS from membrane vesicles. NMR heteronuclear single quantum coherence experiments (HSQCs) were used to probe the interaction with αS and the compound-induced release of αS from the POPG-based liposomes used as membrane mimic. ^{15}N–^{1}H HSQC titration experiments of this compound with αS bound to POPG were performed as described previously (37). Although liposomes are beyond the size limit of NMR and, thus, do not allow for direct detection of the bound state, the measurable signal of αS reveals how much of the protein is still freely tumbling in solution. When referencing the signal intensities observed to those of αS without liposomes, we can identify the regions (and fractions) of αS bound to liposomes (*SI Appendix*, Fig. S5). It should be noted that while the clinical candidate is the purified enantiomer, here measurements were performed with the racemic mixture, as we have demonstrated that interactions between membrane-bound αS and the two enantiomers are identical (*SI Appendix*, Fig. S4). Titration experiments were performed to get more detailed insights into the characteristics of the interaction between UCB0599 and membrane-bound αS (Fig. 6). Here, all detected signal intensity is stemming from unbound protein regions, and a comparison between conditions without liposomes and those with liposomes present allows for a determination of the liposome-bound fraction. Inspection of the concentration dependence (Fig. 6*A*) clearly reveals a pronounced cooperative transition from the membrane-bound to the free state [as also observed for NPT100-18A, an early molecule used for proof of mechanism studies (37)] indicating a specific targeting of UCB0599 to the oligomer. Most importantly, higher αS concentrations require smaller amounts of UCB0599 for displacement (as shown in Fig. 6*A*) indicating that the relevant species for compound binding

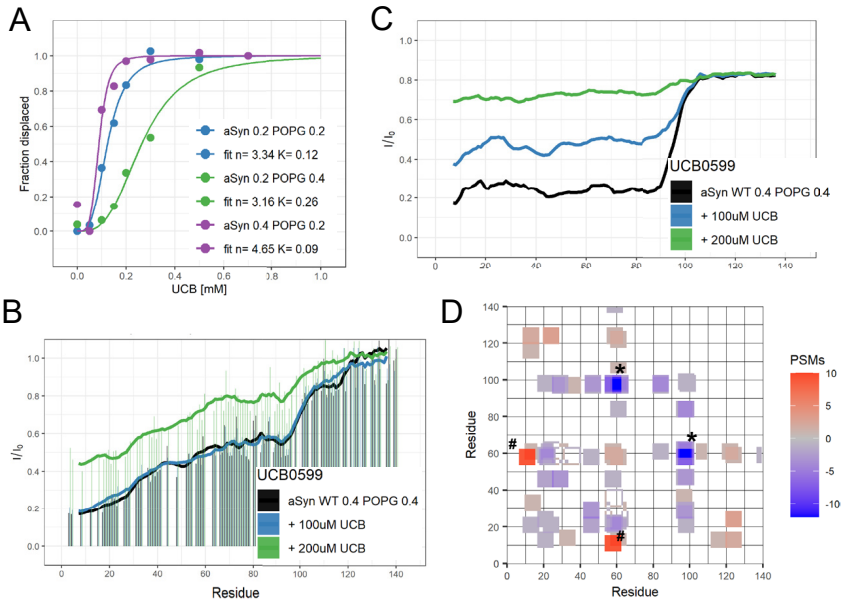


Fig. 6. The effect of UCB0599 on α-synuclein. Addition of the small-molecule UCB0599 to liposome-bound αS modifies the underlying ensemble at low concentration and displaces the protein at elevated ligand concentration. (*A*) Following the stepwise addition of the compound while referencing to the free C-terminus of the protein (*SI Appendix*, Fig. S5), concentration-dependent displacement of αS from POPG-based liposomes is observed. Here, signals of unbound αS regions are measured and their relative intensity between conditions with and without liposomes reports on the bound fraction. This dependence shows cooperative behavior and is dependent on the ratio between protein and liposomes. Best-HSQC (*B*) and DEST (*C*) measurements show that at lower concentrations no change in displaced αS is observed in HSQCs while an increase in DEST intensity already reflects an increase in flexibility of the bound form. When testing for changes in the bound form of the protein (*D*), clear changes in the cross-linking pattern are observed upon addition of the compound. Intermolecular PSMs show a reduction in links in some regions [(92–101) to (58–62), marked *] while increasing in others [(54–61) to (8–16), marked #]. These regions are found at the interface in our calculated dimer (Fig. 4*A*) and highlighted in *SI Appendix*, Fig. S7. The number of PSMs found in a four-residue window is indicated in the color scale.

and subsequent displacement is not monomeric αS, but rather the oligomeric state. This is further substantiated by the fact that the release is not reducing affinity for different regions of αS but releases all parts of the protein at once (*SI Appendix*, Fig. S5). We thus decided to analyze putative changes of the oligomeric state of αS by using low concentrations of UCB0599 insufficient to release αS from the liposome surface (Fig. 6). Most importantly, while the HSQC data (Fig. 6*B*) were nearly unchanged and thus confirmed the prevalence of the liposome-bound state of αS under these conditions, dark state electron-transfer (DEST) experiments (45) clearly revealed an alteration of the membrane-bound species (Fig. 6*C*). The decreased DEST effect (smaller signal linewidth in the bound state) suggests increased flexibility of αS in the bound state prior to displacement, presumably by weakening both (homooligomeric) interactions between individual αS units and αS interactions with the membrane.

In order to specifically probe UCB0599-induced structural changes in the oligomeric state, we again employ $^{14}N/^{15}N$ labeled 1:1 mixtures of isotopically labeled αS to probe XL-MS patterns of liposome-bound αS by using low concentrations of UCB0599.

The differences between intermolecular PSMs observed for αS in each form are shown in Fig. 6*D* (the underlying PSM maps are given in *SI Appendix*, Fig. S6). The addition of UCB0599 leads to clear changes of the intermolecular cross-links (Fig. 6*D*), while intramolecular cross-links did show relatively small fluctuations. Specifically, intermolecular interactions were less frequent between regions (94–101) and (58–62), while an increase in cross-links was observed for regions (54–61) and (8–16). These regions are located at the interface of the obtained dimer structures (*SI Appendix*, Figs. S7 and S12). Distinct changes of the intermolecular cross-link pattern were also observed for interfacial residues in αS versions carrying familial PD (A30P and A53T) mutations and indicate alterations of the interaction interface in membrane-bound oligomers for these pathogenic forms of αS (*SI Appendix*, Fig. S12). Summing up, distinct interactions exist between UCB0599 and liposome-bound αS oligomers and these interactions affect the dynamics of individual αS monomers as well as interactions across the oligomeric interface and thereby alter the oligomeric state distribution of αS, giving us a good description of the mechanism-of-action of UCB0599.

Discussion

Despite the experimental evidence linking the toxicity of αS oligomers and higher-order aggregates to the disruption of membrane integrity, little is known about the molecular-level details and how membrane permeability is modulated by αS's oligomeric state. In this work, we utilize NMR spectroscopy to probe compact states of αS's structural ensemble in the membrane-bound state. In combination with isotope-edited cross-linking mass spectrometry data generated with a zero-length cross-linker, we focus on the compact state of αS and derive high-resolution structural models of membrane-bound αS dimers and higher oligomers. Despite the well-established conformational flexibility (i.e., interconversion between extended and bent α-helical states) of membrane-bound αS, a surprisingly well-defined conformational ensemble was obtained with our approach. It consists of three compact states, one of which is in good agreement with the previously determined structure of SDS-micelle–bound αS (23). Importantly, however, the improved sensitivity of our integrated experimental approach (NMR spectroscopy and XL-MS, see also *SI Appendix*, Fig. S3) revealed the existence of an ensemble of conformational substates.

The experimental data suggest the existence of two prototypical oligomeric topologies: circular, ring-like structures and extended, twisted oligomers. Although the diameter of the circular structures and the exact twist angle of the elongated species can vary (and presumably depend on biological context and environmental conditions, i.e., additional interaction partners located at synaptic vesicles), we postulate a dynamic equilibrium between the prototypical oligomeric species. It is important to note that the oligomeric forms likely differ significantly in terms of membrane embedding and size distribution.

The circular (ring-like) structure shows features that presumably allow it to partly insert into membranes and lead to membrane distortion. Specifically, inspection of the ring-like structure in Fig. 5 suggests a plausible mechanism how αS inserts into lipid bilayers. Many amphipathic peptides can extract lipid molecules from bilayer assemblies and cause semitransmembrane defects (i.e., cathelicidin LL-37) (46) and pore disruptions (47). The distribution of hydrophobic (red) and hydrophilic (blue) residues in the ring-like structure (Fig. 5) suggests a similar insertion mechanism for αS (similarities to such a peptide are expanded upon in supplementary information *SI Appendix*, Fig. S10). Support for this possibility comes from an AFM study on the interaction between αS and membranes (43). AFM studies performed on planar lipid bilayers detected semitransmembrane defects induced by αS. These defects formed by lipid extraction are covered by a large number of αS nanoscale particles with a diameter of ~10 nm (43) matching the diameter of ring-like structures presented here (Fig. 5*B* ~11 nm). Further evidence for the plausibility of this mechanism comes from the established preference of αS for localization on liquid-disordered phases in anionic vesicles (48), where insertion into more fluid parts of the anionic vesicle is facilitated, an effect also seen in the higher affinity observed for smaller diameter vesicles (49).

In contrast, the elongated oligomeric structures can linearly propagate (without additional geometrical constraints) and form high molecular weight oligomers. Due to the inherent twist angle (Fig. 4), progressive (linear) oligomerization is not compatible with continuous membrane attachment as increasing chain length leads to a displacement of membrane-interacting residues from the surface. The proposed ultimate detachment from the membrane would, thus, lead to an oligomer with exposed interaction surfaces, which might be more prone to intermolecular interactions and subsequent aggregate formation.

We thus propose that proper regulation of the equilibrium between different oligomer topologies is crucial for the authentic functionality of αS in its biological context. The possibility to populate several oligomeric forms offers αS the possibility to interact and engage with its diverse binding partners. Conversely, misbalance or deregulation of this process might prevent proper interactions and contribute to the observed pathologies. On the other hand, interventions influencing the equilibrium or modifying the oligomeric states might alleviate pathologies that result from a misbalance of the preexisting equilibrium. Interestingly, the two well-known familial Parkinson's mutations (A53T and A30P, see Fig. 4 *B* and *C*) are located at the interface (facing each other across the interface). Analysis of XL-MS data indicates that the overall structures of the αS monomers and dimers in the fully bound state are not strongly influenced when these point mutations are introduced while intermolecular cross-links are distinctly altered (*SI Appendix*, Fig. S8). We thus postulate that upon mutation, slight adaptations occur across the interaction interface and lead to a shift in the structure or prevalence of oligomeric species, which could explain the differing surface affinities observed (likely the elongated, twisted form is preferentially formed in A53T and A30P) with subsequent alterations of overall membrane affinities, as reported previously (50).

Another important finding of this study is the difference in intermolecular αS interactions depending on whether the

interaction occurs on the same lipid vesicle (cis, i.e., liposomes) or between different membrane species (trans, i.e., bicelles). Importantly, both of these states show different interaction patterns than those observed in the fibrillar state, reflecting the remarkable in vivo structural heterogeneity of αS (Fig. 2) (3, 51). The data obtained in this study suggest that although the widely studied αS amyloids and their toxicity are of vital importance in PD pathology, they probably represent later stages of in vivo (in vitro) aggregation processes (2), while toxicity at the membrane is likely initially related to smaller oligomeric species with different conformations. The high amount of membrane structures and relatively low fibril content of Lewy Bodies also points toward membrane-bound αS as a main culprit in PD pathology (6). New therapeutic strategies in addition to ongoing approaches focussing on the development of antiamyloid agents disrupting large intracellular fibrils, targeting cell-to-cell propagation of misfolded beta-sheet rich aggregates or inhibiting fibril growth processes starting from protofibril "seeds" (50) might thus offer attractive opportunities. In a previous study, a small-molecule compound was identified that interferes with the initial steps of the toxicity-inducing process by altering the structural properties of membrane-bound αS and thereby preventing the subsequent formation of smaller oligomeric αS (37). Although no detailed structural information was available at that time, the hypothesis was formulated that there exists an equilibrium between different oligomeric states and that the compound reduces the conversion to propagating dimers and oligomers, as evidenced by NMR spectroscopy and electron microscopy (37). Additional development led to a compound demonstrating beneficial effects in mouse model systems (52). Building on the demonstrated mechanism, significant optimization efforts were undertaken to improve brain permeability and obtain a new molecule with pharmacological properties suitable for clinical development, yielding the compound UCB0599 (*SI Appendix*, Fig. S18). In contrast to the earlier compound (37), UCB0599 has much improved bioavailability and reaches the brain in sufficient quantities for clinical development. Its safety profile has also been evaluated in phase1/1b studies to support further clinical development (44). In analogy to the previous compound (37), we show a cooperative effect and increasing cooperativity at higher αS:POPG ratios, demonstrating that UCB0599 targets αS which in our conditions is present in oligomeric states. This observed cooperativity holds for all regions of the protein (*SI Appendix*, Fig. S5) and is sustained in point mutations (*SI Appendix*, Fig. S16), indicating that the oligomeric state is specifically targeted and displaced. The XL-MS and NMR data presented in our study provide strong evidence for the proposed hypothesis. Specifically, the DEST experiments (Fig. 6*C*) show that prior to displacement, additional flexibility is introduced to bound αS, followed by full displacement at higher concentrations. The XL-MS data (Fig. 6 and *SI Appendix*, Fig. S6) convincingly show that the compound leads to changes in the cross-link pattern for residues located at the interface, presumably a reflection of the same process causing increased flexibility introduced by UCB0599. Most importantly, the DEST data provided evidence that the compound affects the structural integrity (rigidity) of αS oligomers and their interactions with the membrane mimic already at low concentration. This likely diminishes membrane interactions and impairs embedding of αS oligomers into the membrane and, thus, mitigates possible semitransmembrane defects or membrane-pore formation caused by αS. Our study supports the previous hypothesis and suggests that the beneficial therapeutic effect of UCB0599 is indeed most likely due to an alteration of membrane-bound αS oligomers, the subsequent reduction in the number of membrane-bound oligomeric structures capable of seeding aggregate formation and ultimate displacement of αS from the membrane. To conclude, UCB0599 exerts its beneficial effects already in the membrane-bound state by reducing the populations and thus detrimental downstream effects of both (toxic) oligomeric forms, introducing higher flexibility to oligomers, before ultimately releasing αS in its monomeric (random-coil like) form. A model for the dynamic equilibrium of membrane-bound αS oligomers and the influence of UCB0599 is shown in Fig. 7.

Disrupting the formation of membrane-embedded αS oligomers at an early stage of aggregation before irreversible neurodegenerative processes has been initiated as a promising intervention strategy to combat Parkinson's disease, complementing efforts to reduce amyloids. The now available (high-resolution) 3D structural information

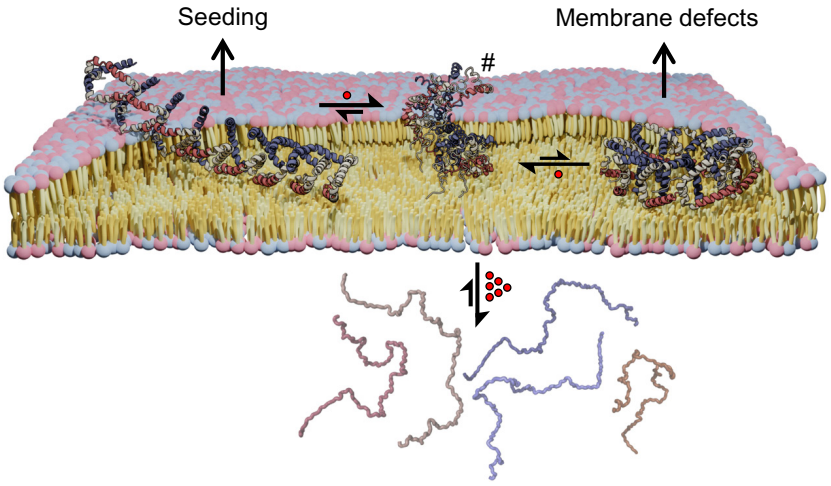


Fig. 7. Model for membrane-bound oligomeric α-synuclein and UCB0599's mode-of-action. The membrane-bound state of αS is characterized by an ensemble of different oligomer topologies (circular vs. elongated). While the elongated form (*Left*) might be involved in seeding with membranes (7) when growing too large, the circular structure (*Right*) is likely relevant for membrane defects and possibly pore formation (43). Proper functioning of αS requires a subtle balancing in order to avoid the formation of these toxic variants. The proposed MoA of UCB0599 involves interference with oligomeric αS on the membrane and thereby shifting the equilibrium away from species capable of generating toxic effects (elongated and circular) toward a conformational state (represented by the central cartoon marked #) characterized by increased flexibility and decreased membrane embedding. The resulting loosening of membrane-attachment facilitates displacement of αS from the membrane with UCB0599 (depicted as red spheres). At sufficiently high concentrations, αS is displaced from the membrane in a monomeric form (*Bottom*). Regions of αS interacting with lipid tails (red) or hydrophilic head groups (blue) are indicated (32).

about surface-bound αS oligomers together with the XL-MS interaction data provide unique insight into UCB0599's MoA, a molecule currently being tested in Parkinson's disease patients in a phase 2 clinical trial. This demonstration of the potential therapeutic relevance of αS oligomer changes on membrane surfaces for Parkinson's disease sets the stage for future intervention strategies exploiting this fundamental process. Additionally, detailed knowledge about the structural features of membrane-bound αS oligomers offers exciting possibilities to expand our understanding regarding the many facets of this still enigmatic protein.

Materials and Methods

Protein Production. The expression and purification of recombinant human αS were carried out similarly as described previously (37). In order to obtain >95% of ^{15}N labeling in samples used for XLMS, precultures were grown in a labeled medium. NMR samples were obtained by generating cell mass in an unlabeled medium. Samples used for assignment were produced using the same method but with D_2O replacing H_2O in the induced medium to partially deuterate and achieve improved TROSY efficiency (53).

Cross-Linking. Cross-linking reactions of αS were carried similarly as published (28). As required, 50 μM ^{14}N αS or 50 μM of a 1:1 isotopic ^{15}N:^{14}N mixture of αS bound to 1 mg/ml of POPG-based liposomes was incubated with 2 mM EDC and 5 mM hydroxy-2,5-dioxopyrrolidine-3-sulfonic acid for 60 min at room temperature for all cross-linking experiments containing liposomes (produced as published) (37). Despite the relatively high amount of αS relative to the liposomal surface, the majority of the protein is observed in the monomeric, bound form (*SI Appendix*, Fig. S14). In order to determine the effect of UCB0599 without displacement of the protein, 50 μM of the compound (racemic mixture in dimethyl sulfoxide (DMSO)) or 1% DMSO (in the reference condition) was added. Conditions containing bicelles or SDS-micelles were generated with 100 μM of isotopically mixed αS and 15 mM bicelles (composed of 10 mM 1,2-diheptanoyl-sn-glycero-3-phosphocholine and 5 mM 1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol)) or with isotopically mixed αS and 40 mM SDS, respectively. All cross-linking reactions were carried out in 50mM phosphate buffer, pH = 6.5 containing 50 mM NaCl. Cross-linking was carried out shortly after mixing the protein and membrane components with the exception of the aggregated sample where αS and bicelles were incubated at 50 °C o/n. All reactions were subjected to SDS-PAGE after being stopped with 50 mM Tris and 20 mM β-mercaptoethanol. Bands of the monomer and dimer forms were cut and subjected to mass spectrometric analysis.

Intact Protein Mass Spectrometry. For intact mass measurements, unlabeled (^{14}N only) αS was cross-linked in the presence of liposomes. After the reaction, 1 mM of the compound (37) was added to displace αS from the liposomal surface. The sample was then incubated at RT for 20 min before ultracentrifugation at 100.000 [g] for 1 h. The supernatant of this reaction was injected onto a SEC column (Superdex 75 Increase 5/150GL). We detected a peak of the monomer and a second, mixed peak. Protein solutions from the mixed peak were diluted with water to a final concentration of 5 to 10 ng/μL. Up to 30 ng of protein were injected on a nano high-performance liquid chromatography (HPLC) system (Thermo Scientific, Dionex) equipped with a C4 column (AERIS C4 widepore, 3.6 μm, 150 × 4.6 mm, Phenomenex) running a step gradient from 9 to 36–63% ACN in 0.08% formic acid. The LC was coupled to the Synapt G2Si mass spectrometer via the Zspray™-ion source (both Waters) operated by MassLynx 4.1 software. The mass spectrometer was run in resolution mode, and capillary voltage was set to 3 kV cone voltage to 40 V. The source was at 120 °C, desolvation at 400 °C. Lock spray correction was applied and Glu-1-fibrinopeptide was used as lock mass. Data were acquired in the m/z range of 600 to 2,000. Spectra were summed over the peak retention time window and masses were deconvoluted using the MaxEnt1 method.

XL-MS and Data Analysis. Isotopically mixed (^{14}N and ^{15}N) αS was cross-linked in the presence of liposomes as described. The resulting protein was subjected to SDS-PAGE, and bands corresponding to mono- and dimer bands were excised and destained After reduction with dithiothreitol and alkylation with iodoacetamide, trypsin and chymotrypsin were used for proteolytic cleavage. Peptides were extracted in the ultrasonication bath and desalted (54).

Peptides were analyzed on an HPLC coupled to a Q Exactive HF-X or an Orbitrap mass spectrometer, equipped with a Nanospray Flex ion source. Equal amounts of the samples were loaded on a trap column using 0.1% trifluoroacetic acid as mobile phase. After 10 min, the trap column was switched in line with the analytical C18 column and peptides were eluted applying a segmented linear gradient from 2 to 80% solvent B (80% acetonitrile, 0.1% formic acid; solvent A 0.1% formic acid) at 230 nL/min over 120 min. The mass spectrometer was operated in a data-dependent mode, survey scans were obtained in a mass range of 350 to 1,650 m/z with lock mass activated, at a resolution of 120,000 at 200 m/z and an AGC target value of 3E6 (HF-X) or 4E5 (Lumos). The 10 most intense ions (HF-X) were selected with an isolation width of 1.6 Thomson for a max. of 250 ms, fragmented in the higher-energy C-trap dissociation cell at 28% (HF-X) or 30% (Lumos), and the spectra were recorded at a target value of 1E5 and a resolution of 60,000. Peptides with a charge of +1, +2, or >+7 were excluded from fragmentation, peptide match was set to preferred, the exclude isotope feature was enabled, and selected precursors were dynamically excluded from repeated sampling for 20 s within a mass tolerance of 8 ppm.

Raw data were processed using the MaxQuant software package (55) (v. 1.6.0.13), and spectra were searched against a combined database of the αS construct sequence, the *E. coli* K12 reference proteome (Uniprot), and a database containing common contaminants.

To identify cross-linked peptides, the spectra were searched using Kojak (56) (v. 1.6.1). Raw files were converted to mzML format using msconvert in ProteoWizard (57) and searched against a database containing the 8 most abundant protein hits identified in the MaxQuant search. Carbamidomethylation of cysteine was set as a fixed modification and oxidation of methionine as variable. Trypsin/chymotrypsin was set as enzyme specificity; EDC was used for cross-linking chemistry, and the 15N mode was activated. Search results were filtered for 1% FDR (q-value < 0.01) at the PSM level using Percolator (v. 2.08). An additional PEP cutoff of < 0.05 was applied.

Due to the flexibility in our system, a large set of cross-links was observed. As the information on cross-links is too sparse to calculate reliable ensembles, we chose to make a visual representation of the average distribution of cross-links the focus of our analysis. The resulting cross-link pattern is shown as a matrix depicting contacts between positions in the primary sequence.

^{14}N–^{15}N and ^{15}N–^{14}N contacts were observed in the dimer bands of SDS-PAGE gels. PSMs detected within ^{14}N–^{14}N and ^{15}N–^{15}N isotope-labeled versions of the protein stem from both inter- and intra-molecular interactions of the protein. We, therefore, subtracted mixed isotopic contacts from those observed within each isotopic species in order to obtain the intramolecular contacts within a single αS molecule of a dimer band.

More details regarding the construction of the "PSM-matrix" and statistical treatment of the observed differences are available in supplementary methods.

NMR. Assignment of SDS-bound αS was carried out using best versions of 3D backbone assignment pulse sequences on an AVANCE NEO 600 MHz instrument equipped with a triple resonance probe (58). Measurements were carried out on a 0.15 mM sample at 323 K with 40 mM SDS in a 20 mM phosphate buffer pH = 5.5 with 0.1 mM EDTA and 0.1 mM azide, and the resulting data along with detailed measurement parameters was deposited on the BMRB database (ID:50996) (59). Best-HSQC and DEST measurements (Fig. 6) were carried out with 0.4 mM αS and 0.4 mg/mL POPG liposomes at 35 °C, in a 50 mM phosphate buffer pH = 6.5 with 50 mM NaCl using dark state energy-transfer irradiation at +/−2 kHz offset in the nitrogen dimension with referencing to 30 kHz offset and 0.5 s of irradiation. Best-HSQC measurements in Fig. 6a also used 0.2 mM αS and 0.2 mg/ml POPG where indicated.

Molecular Structure Determination and PRI Potential. An energy term for Xplor-NIH was developed to allow the use of experimental PRI values as restraints during structure calculation. This was implemented as a long-range angle potential that penalizes angle segments depending on the sign of the measured PRI (*SI Appendix*). Structures were calculated using the molecular structure determination package Xplor-NIH (60, 61). Initial monomer structure calculations were based mainly on PRE, PRI, and torsion angle restraints. In a follow up, the best monomers were used to calculate dimeric complexes by

incorporating intermolecular MS-cross-links as pseudo NOEs (*SI Appendix*). A short graphical summary of the model generation workflow is shown in *SI Appendix*, Fig. S23.

Data, Materials, and Software Availability. All data used to demonstrate our findings in this paper are available within the article, its *SI Appendix* files or in public repositories. Chemical shift assignments for SDS micelle and bicelle bound αS are deposited in the BMRB database (59) with the accession numbers 50996 and 51148. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (62) with the dataset identifier PXD027349. Structural models of oligomeric αS are deposited on the PDB-Dev database (63) with the accession number PDBDEV_00000089.

ACKNOWLEDGMENTS. Terry Baker passed away during finalization of this manuscript. He was an inspiring scientist and dedicated researcher who made countless, invaluable contributions to the discovery of therapeutic agents and will always be remembered as a great colleague. T.C.S. is grateful for funding provided by UCB Biopharma. We gratefully acknowledge the financial support to A.B. and T.H. by the Austrian Federal Ministry for Digital and Economic Affairs, the National Foundation for Research, Technology and Development and the Christian Doppler Research Association. We thank Rachel Angers for fruitful discussions and her input in finalising the manuscript, Anja-Leona Biere for continuous support of the project and Dorothea Anrather for conducting intact protein mass spectrometry. All mass spectrometry measurements were performed using the instrument pool of the Vienna BioCenter Core Facilities.

Author affiliations: [a]Department of Structural and Computational Biology, Max Perutz Labs, University of Vienna, Vienna 1030, Austria; [b]Department of Structural and Computational Biology, Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, University of Vienna, Vienna 1030, Austria; [c]Mass Spectrometry Facility, Max Perutz Labs, Vienna BioCenter, Vienna 1030, Austria; and [d]UCB Pharma, Slough SL1 3WE, United Kingdom

Author contributions: T.S.B., R.J.T., and R.K. designed research; T.C.S., A.B., K.L., T.G., and T.H. performed research; T.C.S., A.B., T.G., and M.H. analyzed data; and T.C.S., A.B., M.H., and R.K. wrote the paper.

Competing interest statement: T.S.B. and R.J.T. were employed by UCB Pharma and held stock options and shares of the same during the writing of this paper. T.C.S. and R.K. received funding from UCB Pharma and Neuropore Therapies. T.C.S., K.L., and R.K. are co-authors on a publication describing an early molecule showing a "proof-of mechanism" used by UCB0599.

1. M. Goedert, R. Jakes, M. G. Spillantini, The synucleinopathies: Twenty years on. *J. Parkinsons Dis.* **7**, S51–S69 (2017).
2. P. Alam, L. Bousset, R. Melki, D. E. Otzen, α-synuclein oligomers and fibrils: A spectrum of species, a spectrum of toxicities. *J. Neurochem.* **150**, 522–534 (2019).
3. Y. C. Wong, D. Krainc, alpha-synuclein toxicity in neurodegeneration: Mechanism and therapeutic strategies. *Nat. Med.* **23**, 1–13 (2017).
4. A. van Maarschalkerweerd, V. Vetri, A. E. Langkilde, V. Fodera, B. Vestergaard, Protein/lipid coaggregates are formed during alpha-synuclein-induced disruption of lipid bilayers. *Biomacromolecules* **15**, 3643–3654 (2014).
5. P. H. Weinreb, W. Zhen, A. W. Poon, K. A. Conway, P. T. Lansbury Jr., NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry* **35**, 13709–13715 (1996).
6. S. H. Shahmoradian *et al.*, Lewy pathology in Parkinson's disease consists of crowded organelles and lipid membranes. *Nat. Neurosci.* **22**, 1099–1109 (2019).
7. H. J. Lee, C. Choi, S. J. Lee, Membrane-bound alpha-synuclein has a high aggregation propensity and the ability to seed the aggregation of the cytosolic form. *J. Biol. Chem.* **277**, 671–678 (2002).
8. M. Perni *et al.*, A natural product inhibits the initiation of alpha-synuclein aggregation and suppresses its toxicity. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E1009–E1017 (2017).
9. L. Antonschmidt *et al.*, Insights into the molecular mechanism of amyloid filament formation: Segmental folding of alpha-synuclein on lipid membranes. *Sci. Adv.* **7**, eabg2174 (2021).
10. J. Burre, M. Sharma, T. C. Sudhof, Definition of a molecular pathway mediating alpha-synuclein neurotoxicity. *J. Neurosci.* **35**, 5221–5232 (2015).
11. J. Burre *et al.*, Alpha-synuclein promotes SNARE-complex assembly in vivo and in vitro. *Science* **329**, 1663–1667 (2010).
12. J. Lautenschlager *et al.*, C-terminal calcium binding of alpha-synuclein modulates synaptic vesicle interaction. *Nat. Commun.* **9**, 712 (2018).
13. D. Sulzer, R. H. Edwards, The physiological role of alpha-synuclein and its relationship to Parkinson's Disease. *J. Neurochem.* **150**, 475–486 (2019).
14. G. Musteikytė *et al.*, Interactions of α-synuclein oligomers with lipid membranes. *Biochim. et Biophys. Acta. Biomembr.* **1863**, 183536 (2020), 10.1016/j.bbamem.2020.183536.
15. J. Burre *et al.*, Properties of native brain alpha-synuclein. *Nature* **498**, E4–E6 (2013), discussion E6–E7.
16. T. Bartels *et al.*, The N-terminus of the intrinsically disordered protein alpha-synuclein triggers membrane binding and helix folding. *Biophys. J.* **99**, 2116–2124 (2010).
17. K. Araki *et al.*, A small-angle X-ray scattering study of alpha-synuclein from human red blood cells. *Sci. Rep.* **6**, 30473 (2016).
18. H. J. Dyson, P. E. Wright, NMR illuminates intrinsic disorder. *Curr. Opin. Struct. Biol.* **70**, 44–52 (2021).
19. F. X. Theillet *et al.*, Structural disorder of monomeric alpha-synuclein persists in mammalian cells. *Nature* **530**, 45–50 (2016).
20. C. C. Jao, A. Der-Sarkissian, J. Chen, R. Langen, Structure of membrane-bound alpha-synuclein studied by site-directed spin labeling. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8331–8336 (2004).
21. B. Uluca *et al.*, DNP-enhanced MAS NMR: A tool to snapshot conformational ensembles of alpha-synuclein in different states. *Biophys. J.* **114**, 1614–1623 (2018).
22. M. Drescher *et al.*, Antiparallel arrangement of the helices of vesicle-bound alpha-synuclein. *J. Am. Chem. Soc.* **130**, 7796–7797 (2008).
23. T. S. Ulmer, A. Bax, N. B. Cole, R. L. Nussbaum, Structure and dynamics of micelle-bound human alpha-synuclein. *J. Biol. Chem.* **280**, 9595–9603 (2005).
24. A. J. Trexler, E. Rhoades, Alpha-synuclein binds large unilamellar vesicles as an extended helix. *Biochemistry* **48**, 2304–2306 (2009).
25. M. Bortolus *et al.*, Broken helix in vesicle and micelle-bound alpha-synuclein: Insights from site-directed spin labeling-EPR experiments and MD simulations. *J. Am. Chem. Soc.* **130**, 6690–6691 (2008).
26. S. B. Lokappa, T. S. Ulmer, Alpha-synuclein populates both elongated and broken helix states on small unilamellar vesicles. *J. Biol. Chem.* **286**, 21450–21457 (2011).
27. Y. H. Sung, D. Eliezer, Structure and dynamics of the extended-helix state of alpha-synuclein: Intrinsic lability of the linker region. *Prot. Sci. Publ. Prot. Soc.* **27**, 1314–1324 (2018).
28. S. T. A. Amos *et al.*, Membrane interactions of α-Synuclein revealed by multiscale molecular dynamics simulations, markov state models, and NMR. *J. Phys. Chem. B* **125**, 2929–2941 (2021), 10.1021/acs.jpcb.1c01281.
29. B. Li *et al.*, Cryo-EM of full-length alpha-synuclein reveals fibril polymorphs with a common structural kernel. *Nat. Commun.* **9**, 3609 (2018).
30. M. D. Tuttle *et al.*, Solid-state NMR structure of a pathogenic fibril of full-length human alpha-synuclein. *Nat. Struct. Mol. Biol.* **23**, 409–415 (2016).
31. P. Kumari *et al.*, Structural insights into alpha-synuclein monomer-fibril interactions. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2012171118 (2021).
32. G. Fusco *et al.*, Structural basis of membrane disruption and cellular toxicity by alpha-synuclein oligomers. *Science* **358**, 1440–1443 (2017).
33. Z. Lv *et al.*, Assembly of alpha-synuclein aggregates on phospholipid bilayers. *Biochim. Biophys. Acta Prot. Proteom.* **1867**, 802–812 (2019).
34. I. F. Tsigelny *et al.*, Dynamics of alpha-synuclein aggregation and inhibition of pore-like oligomer development by beta-synuclein. *FEBS J.* **274**, 1862–1877 (2007).
35. J. Burre, M. Sharma, T. C. Sudhof, alpha-Synuclein assembles into higher-order multimers upon membrane binding to promote SNARE complex formation. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E4274–4283 (2014).
36. N. B. Cole *et al.*, Lipid droplet binding and oligomerization properties of the Parkinson's disease protein alpha-synuclein. *J. Biol. Chem.* **277**, 6344–6352 (2002).
37. W. Wrasidlo *et al.*, A de novo compound targeting alpha-synuclein improves deficits in models of Parkinson's disease. *Brain A J. Neurol.* **139**, 3217–3236 (2016).
38. C. Arlt *et al.*, An integrated mass spectrometry based approach to probe the structure of the full-length wild-type tetrameric p53 tumor suppressor. *Angew Chem. Int. Ed. Engl.* **56**, 275–279 (2017).
39. M. H. Polymeropoulos *et al.*, Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045–2047 (1997).
40. D. Kurzbach *et al.*, Detection of correlated conformational fluctuations in intrinsically disordered proteins through paramagnetic relaxation interference. *Phys. Chem. Chem. Phys.* **18**, 5753–5758 (2016).
41. D. K. Schweppe, J. D. Chavez, J. E. Bruce, XLmap: An R package to visualize and score protein structure models based on sites of protein cross-linking. *Bioinformatics* **32**, 306–308 (2016).
42. J. Varkey *et al.*, α-Synuclein oligomers with broken helical conformation form lipoprotein nanoparticles. *J. Biol. Chem.* **288**, 17620–17630 (2013).
43. J. Pan, A. Dalzini, N. K. Khadka, C. M. Aryal, L. Song, Lipid extraction by α-synuclein generates semi-transmembrane defects and lipoprotein nanoparticles. *ACS Omega* **3**, 9586–9597 (2018).
44. J. W. Smit *et al.*, Phase 1/1b studies of UCB0599, an Oral inhibitor of alpha-synuclein misfolding, including a randomized study in Parkinson's disease. *Mov. Disord.* **37**, 2045–2056 (2022), 10.1002/mds.29170.
45. N. L. Fawzi, J. Ying, D. A. Torchia, G. M. Clore, Probing exchange kinetics and atomic resolution dynamics in high-molecular-weight complexes using dark-state exchange saturation transfer NMR spectroscopy. *Nat. Protoc.* **7**, 1523–1533 (2012).
46. E. Sancho-Vaello *et al.*, The structure of the antimicrobial human cathelicidin LL-37 shows oligomerization and channel formation in the presence of membrane mimics. *Sci. Rep.* **10**, 17356 (2020).
47. P. Kumar, J. N. Kizhakkedathu, S. K. Straus, Antimicrobial peptides: Diversity, mechanism of action and strategies to improve the activity and biocompatibility in vivo. *Biomolecules* **8**, 4 (2018).
48. M. Stöckl, P. Fischer, E. Wanker, A. Herrmann, Alpha-synuclein selectively binds to anionic phospholipids embedded in liquid-disordered domains. *J. Mol. Biol.* **375**, 1394–1404 (2008).
49. S. M. McClain, A. M. Ojoawo, W. Lin, C. M. Rienstra, C. J. Murphy, Interaction of alpha-synuclein and its mutants with rigid lipid vesicle mimics of varying surface curvature. *ACS Nano.* **14**, 10153–10167 (2020).
50. H. A. Lashuel, C. R. Overk, A. Oueslati, E. Masliah, The many faces of α-synuclein: From structure and toxicity to therapeutic target. *Nature Rev. Neurosci.* **14**, 38–48 (2013).
51. I. F. Tsigelny *et al.*, Role of alpha-synuclein penetration into the membrane in the mechanisms of oligomer pore formation. *FEBS J.* **279**, 1000–1013 (2012).
52. D. L. Price *et al.*, The small molecule alpha-synuclein misfolding inhibitor, NPT200-11, produces multiple benefits in an animal model of Parkinson's disease. *Sci. Rep.* **8**, 16165 (2018).
53. A. Eletsky, A. Kienhöfer, K. Pervushin, TROSY NMR with partially deuterated proteins. *J. Biomolecular NMR* **20**, 177–180 (2001).
54. J. Rappsilber, M. Mann, Y. Ishihama, Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).

55. S. Tyanova, T. Temu, J. Cox, The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).
56. M. R. Hoopmann *et al.*, Kojak: Efficient analysis of chemically cross-linked protein complexes. *J. Proteome Res.* **14**, 2190–2198 (2015).
57. M. C. Chambers *et al.*, A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**, 918–920 (2012).
58. E. Lescop, P. Schanda, B. Brutscher, A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *J. Magn. Reson.* **187**, 163–169 (2007).
59. E. L. Ulrich *et al.*, BioMagResBank. *Nucleic Acids Res.* **36**, D402–D408 (2008).
60. C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73 (2003).
61. C. D. Schwieters, J. J. Kuszewski, G. Marius Clore, Using Xplor–NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **48**, 47–62 (2006).
62. Y. Perez-Riverol *et al.*, The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).
63. B. Vallat, B. Webb, J. D. Westbrook, A. Sali, H. M. Berman, Development of a prototype system for archiving integrative/hybrid structure models of biological macromolecules. *Structure* **26**, 894–904. e892 (2018).

PNAS
www.pnas.org

Supplementary Information for

High-Resolution Structural Information of Membrane-bound α-Synuclein provides Insight into the MoA of the Anti-Parkinson Drug UCB0599

Thomas C. Schwarz[1], Andreas Beier[1], Karin Ledolter, Thomas Gossenreiter, Theresa Höfurthner, Markus Hartl, Terry S. Baker, Richard J. Taylor, and Robert Konrat[*]*

[1]A.B. and T.C.S. contributed equally to this work

[*]to whom correspondence may be addressed

Email: robert.konrat@univie.ac.at; t.schwarz@univie.ac.at

This PDF file includes:

Other supplementary materials for this manuscript include the following:

Dataset S1, code PRI Potential

SUPPLEMENTARY INFORMATION TEXT

The influence of mutations on the oligomeric state of membrane bound αS

In addition to crosslink maps constructed for the wildtype protein, the point mutants A30P and A53T were also studied. As can be observed in Figure S8 the intramolecular crosslinking pattern does show mostly similar results for the monomers of mutants and wildtype, with changes being mostly due to a globally observed slight change in PSMs. The monomeric structure of the bound state of αS, thus, does not seem to be dramatically changed by the two point mutations tested. Intramolecular links observed in the dimer, however do show changes. A decreased amount of crosslinks between regions (50-65) to (90-100) as well as (30-40) and (90-105) in the mutant A30P indicates a change in the monomers making up dimers (and multimers) of this species. As these longer range crosslinks span the NAC region, it is likely that this region is more exposed in the A30P mutation than observed for WT and A53T. Even more pronounced are the changes observed in intermolecular crosslink maps. Here we can see that the point mutations, despite causing only small changes in the monomeric membrane bound form, do alter the contacts between monomers extensively. Especially prominent are increases in crosslinks between regions (110-140) to (55-65) in A30P and (15-25) to (115-125) in A53T. These links are not absent in the wildtype protein but are much more prominent in the mutants, likely indicating that a different oligomeric state is populated predominantly in these mutants. To exclude that local a concentration effect impacts the PSM maps derived, we subjected dimeric αS obtained by high concentration (200μM in this case) rather than association with liposomes to the same XL-MS procedure. The resulting maps show very small amounts of detected PSMs both in the inter- and intramolecular cases (Fig.S14). As the same amount of digested peptide was injected into the mass spectrometer in each case and the protein intensity and MSMS counts are fluctuating by no more than 10%, this change clearly indicates the

emergence of specific crosslinks by the introduction of liposomes. The reduced PSMs in the intramolecular case are due to the lack of a stable structure. While in the intermolecular case they are explained either by a broad and divergent set of crosslinked peptide fragments, leading to a lack of signal intensity for PSM detection, or less likely, by a specific link that is enriched in this case but not detected due to poor XL-MS properties. In either case the detected PSMs with liposomes clearly represent a novel, specific and distinct pattern.

The biological relevance of XL-MS in high protein to surface ratio conditions

In the conditions chosen for XL-MS experiments demonstrated in this paper, we expect a large degree of saturation for the liposomal surface with αS. As the protein is one of the most abundant proteins in the brain and naturally found both in its membrane bound as well as its free state (1), a relatively high saturation of membrane surface with αS is expected. The formation of oligomers at membranes has been shown to be modulated (reduced) by β-Synuclein and γ-Synuclein incorporation into oligomers upon their overexpression, further indicating a high amount of αS at the membrane (2). In addition, oligomerisation of αS has been demonstrated at much lower concentrations, demonstrating that oligomerisation is not an artefact of the high concentrations used (3). The used conditions are also still yielding large amounts of bound, monomeric αS, indicating that a full saturation of the membrane that leads to unwanted crosslinks by forced proximity was not observed (see Fig.S14e). In addition, the use of PSM maps, rather than specific single crosslinks, reduces the risk of overinterpretation of single contacts. As encounters between αS molecules at the membrane that are driven only by charge and/or chance would yield a wide distribution of crosslinks it would introduce noise in the background, rather than specific crosslink locations. The possibility of some regions of any protein being more prone to crosslink formation is always present, but is addressed here as

we can see differing crosslink patterns in varying conditions (Fig.2, Fig.S22) as well as a lack of focused crosslink regions being detected in high concentration conditions (Fig.S14c). The observation that crosslinks analysed in the trimeric state isolated in the SDS-PAGE (Fig.S13) further confirms our interpretation of hot spot regions in the crosslink maps.

Our experimental data is not fully explained by previously available structural data

When carrying out relaxation measurements of αS bound to SDS micelles, we were anticipating only minor variations from the expected rate patterns as calculated from published structural data (PDB: 1XQ8). Although measured at different temperature and pH, we still were measuring monomeric membrane bound αS with a very similar secondary structure propensity, so the changes observed were surprising (see Fig.S3). We expect especially the higher temperatures to be contributing to the possibility of forming more compact states in our ensemble than previously observed. It is these compact states that contribute highly to PRE and PRI measurements and are, therefore, the structures we can calculate from the available data although they were not captured in previous efforts. Although deemphasized in our calculation protocol, structures very similar to the "horseshoe-type" were also obtained (see Fig.3). Intermolecular contributions to PRE measurements were excluded by measuring 1:1 mixtures of the ^{14}N MTSL labelled mutant αS19C and ^{15}N wildtype protein. Here we did not observe meaningful deviations from the noise both in bicelle- and micelle-bound measurements (see Fig.S15).

Using the same PDB file (1XQ8) we also calculated CA-CA distances to demonstrate where crosslinking contacts would be most expected for this structure. The resulting distances (Fig.S3d) are relatively large and although the pattern does not fit the monomeric liposome crosslinks too well, it is a better match than for the dimeric state (see Fig.2a vs. 2e), once again

indicating that the calculated more compact structures are an improvement for the explanation of the oligomeric state.

This is also demonstrated in Figure S9, where CA-CA distance plots for the calculated dimer (in its ring forming conformation) are overlaid with crosslinks obtained for dimers (i.e. the dimer band in SDS-PAGE). The pattern of crosslinks fit these CA-CA distances better than is the case for the horseshoe-type structure (see Fig.S3d). A comparison of crosslinks obtained from aggregates that were formed in the presence of bicelles and a fibrillar published structure is also presented in Figure S9. As obtained crosslinks and calculated CA-CA differences do not match in this case, we conclude that the aggregates formed in this case do display substantially different structural characteristics than those obtained in standard fibril formation assays.

To obtain a better quantitative estimate of the improvement in realising experimental crosslinking data, we have used the XLmap (4) R-package. Here, the capability of a given structure to fulfil experimentally detected XL-MS crosslinks is tested for given CA-CA distance windows. In this setup fewer unfulfilled contacts remaining at a given distance indicate a better match between the structure and obtained data. In Figure S20 we show how this approach was successfully used to obtain a quantitative readout on the improvement the novel structures offer over published structures in terms of their correspondence to XL-MS crosslinking data. The calculated more compact states are better in capturing measured contacts in all our datasets as can be seen in Figure S20. Although the methodology cannot account for “expected links” that are missing, we are still able to reproduce the qualitative comparisons seen in Figure S9. The improvement is especially evident in the intermolecular case (Fig.S20d) as the publicly available oligomeric structures stem from fibrils, which would lead to very different crosslinking events.

NMR secondary shift and PRE data show comparability of bicelle and micelle conditions

As the micelle bound state of αS might introduce a strong curvature in the protein that is not expected in the liposome bound state, we undertook efforts to characterize the protein in its bicelle bound state as well. After finding bicelle conditions that allowed for the measurement of triple resonance experiments, we assigned αS in its bicelle bound state (BMRB-ID 51148). In these conditions only monomeric αS was detected on bicelles. Comparing the secondary structure propensity (SSP) of the bicelle bound and micelle bound states (BMRB-ID 50996) we see a very similar pattern of an unstructured C-terminus and α-helical tendencies in the first 100 residues. Three distinct regions of higher helicity are seen in both cases (see Fig.S1). As bicelles should allow for a more elongated structure to form, this indicates that the protein does not adopt a single extended helix in these conditions although slightly higher SSP values are observed. SDS micelle bound αS was used for the measurement of ^{1}H relaxation experiments as it gave sufficient signal to noise to obtain PRI data. In Figure S19d a PRE measurement of bicelle-bound αS mutant 19C is provided to demonstrate the comparison of PRE effects in the SDS- and bicelle-bound states. Although the bicelle-bound state did not allow for relaxation rate measurements, the relative intensities in the vicinity of the label match well with the rates observed in the micelle-bound state. The proximity of the label to the region around residue 100 is also seen as a dip in the intensity measurement. Here the effect is smaller than observed in the SDS-based measurements, an effect likely indicating that although present, smaller amounts of the compacted state are forming in the bicellar state. A likely explanation for the reduced prevalence of the compacted substate is the lack of strong curvature in the bicelle, which had allowed for interactions with hydrophobic tails in the SDS bound state. In liposomes this interaction is facilitated by oligomer formation.

The circular oligomers derived share similarities with pore forming peptides

The circular structure shown in Figure 5 has an intriguing distribution of hydrophobic (red) and hydrophilic (blue) residues (colours according to published data (5)). If one imagines the ring like structure positioned on top of a membrane the hydrophobic region of the protein faces towards the membrane and to the outside of the ring. The hydrophilic region, however, faces the inside and away from the membrane (see also Figure 7). These features could allow for a relatively stable position on the membrane with a hydrophobic anchor, while still providing possibilities for membrane disruption due to the hydrophilic "channel" formed in the centre of the ring. Interestingly, this process would be similar to the accumulation of amphipathic peptides on membrane surfaces which, beyond a critical threshold concentration, can lead to transmembrane pores via two structural arrangements, barrel-stave pore or toroidal pore structures. In this context it is enlightening to compare the primary sequence of αS with sequences from antimicrobial peptides known to induce carpet or toroidal-pore membrane disruptions (6), particularly in view of the antimicrobial activity of αS itself. Sequence alignment (Fig.S10) with the antimicrobial peptide of the cathelicidin family LL-37 revealed that the repeating sequence motif in αS displays a pronounced primary sequence similarity to the antimicrobial peptide. Interestingly, mapping the homologue's sequence onto the oligomer structure (Fig.S10b,c), shows a striking structural similarity of the intermolecular interface in αS oligomers to recent LL-37 pore structures that suggest the formation of a four-helix bundle of the membrane-bound antimicrobial peptide (7).

The elongated oligomers might contribute to transmission

As the elongated oligomeric structure derived can linearly propagate without being "capped" by forming a circular structure, it allows for the formation of high molecular weight oligomers.

However, the observed twist angle (Fig.4) does not allow for a continuous membrane attachment after addition of several monomer units. This is due to the increasing chain length causing a cumulatively larger twist and a displacement of membrane-interacting residues from the surface. As this would ultimately lead to a detachment from the membrane surface, this means the elongated structure might be more prone to intermolecular interactions, i.e. as primary elementary steps in aggregate formation. Depending on the twist angle, large spiral-shaped aggregates similar to already known macromolecular aggregates (8) might form and the hypothesis that these species could be relevant for cell-to-cell transmission via binding to Heparansulfate-proteoglycan containing receptors (9) seems plausible.

The effect of UCB0599 can be seen directly in bicelle bound αS

As we have established conditions in which αS can be measured bound to bicelles (Fig.S1) and these conditions mimic the membrane bound state, we were able to test UCB0599 has an effect on this monomeric state as well. Although in these conditions only a portion of the proteins' states are expected to be compact and affected by UCB0599, addition of UCB0599 to bicelle bound αS did show changes in the positions of NH amide peaks. Figure S11 shows that the strongest changes are in the region 75-95 which is also the region of αS transitioning from bicelle bound to freely tumbling in solution. While optimizing bicelle conditions, we noticed that a higher content of PiP_2 in the bicelle lead to a stronger interaction (evident in better S/N of bound residues and a stronger shift away from the free form). On the other hand, addition of the compound at high concentration is expected to weaken the interaction of synuclein with the bicelle, causing shifts in addition to those stemming from an interaction of the compound itself with the protein. We, therefore, chose to compare the directionality of shifts between addition

of the compound and a decrease in PiP_2 concentration (as both are expected to weaken the interaction with bicelles). A different directionality of the shift indicates a dissimilar effect of the compound as compared to a reduction in PiP2 concentration. Thus, a change in directionality given as an angle between the two shifts mostly indicates changes caused by factors other than displacement from the bicelle. These are caused either by the compound directly or indirectly by changes to the ensemble. As seen in Figure S11 the two regions where the strongest directional effect is observed are 35-40 and 65-75. In our model of compact αS the region 35-40 is in close proximity to 90-100 and, thus, impacted by changes to the ensemble that affect this region (either by release or directly). Changes in the region 35-40 are interestingly close to disease causing mutations and are located at a helical break in the compact structure that would be expected to straighten in the more typical conformers of the ensemble.

Displacement of αS by UCB0599 from liposomal membranes is not sequential

In Figure 6 we have shown the effect of UCB0599 by the relative average displacement achieved. The titration data obtained do, however, contain even more detailed information. As each amino acid is tracked separately, we can study the effect of UCB0599 at higher resolution. In Figure S5, we show that different areas of the protein all show very similar behaviour in response to UCB0599. If the compound would be targeting monomeric αS bound along the surface of liposomes, we would expect the effect of UCB0599 to be more pronounced in the region targeted by the compound, as a weakening of membrane interaction in a specific region would likely lead to an increase in signal intensity for this specific region prior to the effect being detected at other positions. Even if the compound caused a weakening of overall membrane affinity, we would expect the N-terminal regions of αS to stay bound longer as they show higher membrane affinity (10). The lack of strong changes upon addition of UCB0599 to

free αS as well as the observation that UCB0599 does not impact liposome integrity demonstrates that we are not measuring a secondary effect of micelle formation by UCB0599 or its impact on liposome morphology (Fig.S17). The observation in Figure S5 that all membrane bound residues are affected identically fits well with UCB0599 targeting an oligomeric compact state that, upon interaction with UCB0599 at sufficient concentration, is disaggregated. The more gradual increase in signal intensity between residues 20 and 100 observed at higher concentrations is, thus, a reflection of differing non-oligomeric bound states present in the ensemble. This interpretation is further bolstered by the observation that the mutants A30P and E83R as well as a construct missing the first 30 residues (dNter) also show sigmoidal displacement (Fig.S16). The earlier displacement of both dNter as well as A30P could be caused by either reduced membrane affinity of the oligomer or reduced stability to UCB0599. As synuclein is N-terminally anchored to the membrane (10) , the similarity in the effect seen for this truncated version of the protein speaks against a simple reduction in affinity for a non-oligomeric form of αS. As the conditions used for these experiments likely induce a large degree of oligomer formation at the membrane, the effect UCB0599 has on monomeric POPG bound αS cannot be delineated in detail. Full displacement is achieved by UCB0599, but we cannot state if this displacement starts at a specific region of the protein. Bicelle based data, however, does indicate first effects in the region 75-95 as stated above.

Enantiomers of UCB0599 are identical in their action on αS liposome binding

The *invitro* experiments described here utilised the racemic mixture of UCB0599. The compound used in phase 1 and 2 clinical trials in Parkinson's disease patients, however, is the pure enantiomer. The chemical structure of the compound is shown in Figure S18. This compound resulted from lead optimization efforts improving brain permeability and

pharmacological properties to yield a compound suitable for clinical development. To demonstrate that the effect observed in our *invitro* experiments does not depend on utilisation of the pure enantiomer, we tested each enantiomer individually for its capability to influence the membrane bound form oligomeric αS. As we have concentrated on the effect UCB0599 has on oligomeric structures of αS prior to displacing it from the membrane, we also tested the enantiomers of UCB0599 at lower relative concentrations, before full displacement is observed. In Figure S4 we show that, as observed with the racemic mixture, only a very slight increase in signal intensity is observed in HSQC measurements at such low concentrations. The change in intensity is, however, not meaningfully dissimilar between the two enantiomers. As Dobson, Bax and co-workers have shown, αS binding to phospholipids shows slow chemical exchange on the NMR timescale. (11) Therefore, all signal detected stems from free regions of the protein and signal reduction compared to the free form indicates binding. The lack of a plateau indicating distinct separate binding modes as seen in different conditions (10) is likely due to a larger ensemble of binding modes being realized in our conditions. Varying plateaus depending on phospholipid conditions and relative concentrations have also been observed previously. (12) A stronger effect is observed in dark-state exchange saturation transfer (DEST) experiments (13), where the effect of irradiating the protein in the nitrogen frequency while it is bound to liposomes is transferred to the observable free form. DEST efficiencies are affected by both exchange and dynamics and, thus, are sensitive to the changes the ligand introduces in the system. As seen in Figure S4 both enantiomers again show the same effect in our measurement, indicating that the two enantiomers are acting on liposome bound αS in the same way.

Low concentration effects of UCB0599

The capability of UCB0599 and its progenitor compounds to displace αS from membranes has been shown in various different conditions with a number of different techniques (shown here are NMR and SDS-PAGE with ultracentrifugation, progenitors were also tested with EM and CD-spectroscopy). Notably a sigmoidal action profile is observed in titration experiments, indicating that the compound needs to cooperatively act on the membrane bound oligomer. Prior to displacement, HSQC based NMR measurements mostly showed flat or even slightly reduced signal intensity when UCB0599 is added to a liposome αS mixture (Fig. 6a). The fact that DEST measurements (13) demonstrated a markedly different effect at a concentration of the compound which shows very similar signal intensities in HSQC measurements (Fig. 6b, c) clearly indicates a change in the structural dynamics of the oligomeric membrane bound ensemble of the protein. Importantly, XLMS measurements of liposome bound oligomers with UCB0599 being present were performed with similar (low) UCB0599 concentrations and thus report on the structural dynamics of the (still) membrane-bound oligomeric state (Figures 6, S6 and S12). To demonstrate that variations in the PSM maps are not due to experimental error, we also show that very similar differential patterns could be obtained on separate samples with a different mass spectrometer (Fig. S21).

SUPPLEMENTARY INFORMATION METHODS

UCB0599

The compound used in this study was manufactured and provided by UCB Biopharma SRL.

NMR

Assignment of Bicelle-bound αS was carried out using BEST-TROSY versions of 3D backbone assignment pulse sequences on an AVANCE III 700 MHz instrument equipped with a triple resonance cryo-probe (14). Measurements were carried out on a 0.1 mM partially deuterated sample at 323 K with Bicelles containing 10.6 mM DHPC, 5.1mM DMPG and 1.1mM PiP2 in 8 mM Phosphate buffer pH=5.5 with 0.1 mM EDTA and 0.1 mM Azide. The production of bicelles started by mixing stock solutions of dissolved lipids acquired from Avanti® Polar Lipids in a glass vial to obtain a specified quantity of the lipids desired. The solvents used were Chlorophorm, Methanol and H_2O. After all components were fully dissolved, the solution was evaporated under a stream of dried nitrogen gas. The glass vial was further dried under vacuum (30 millibar) for at least one hour. The desired buffer was added to the dried lipids to generate a 10x concentrated bicelle stock (i.e. 168mM total lipid content). This mix was incubated under very slight shaking at 37C for 1hr. During this time they were pipetted up and down as well as subjected to a sonication bath (2 minutes) twice. Finished bicelles were either used directly or frozen with liquid nitrogen and stored at -80C in aliquots, bicelle formation was confirmed by DLS measurements. Information from selectively leucine and alanine labeled samples was used for confirmation in some cases. The assignment and detailed measurement conditions are published on the BMRB-Database (ID 51148).

Bicelle containing chemical shift experiments were measured on an AVANCE NEO 800 MHz instrument equipped with a triple resonance probe using BEST-TROSY type experiments (15). Comparison of the R and S enantiomers were carried out on the same instrument using best-HSQC (16) and DEST (13) experiments. To measure the dark state energy transfer irradiation at +/- 3kHz offset in the nitrogen dimension was referenced to 30kHz offset, with 0.5s of irradiation. Displacements series of αS from POPG based liposomes were measured using SOFAST-HSQC measurements on an Avance NEO 600MHz with a triple resonance probe (17).

Best-HSQC and DEST measurements were carried out with 0.4mM αS and 0.4mg/ml POPG liposomes at 35C using dark state energy transfer irradiation at +/- 2kHz offset in the nitrogen dimension with referencing to 30kHz offset, and 0.5s of irradiation.

XL-MS and data analysis

In detail, the Coomasie stained gel band was destained with a 1:1 mixture of acetonitrile (Chromasolv®, Sigma-Aldrich) and 50 mM ammonium bicarbonate (Sigma-Aldrich). Reduction was carried out with 10 mM dithiothreitol (Roche) and alkylation with 50 mM iodoacetamide. Digestion was carried out with trypsin (Promega; Trypsin Gold, Mass Spectrometry Grade) at 37 °C overnight and subsequently with chymotrypsin (Promega sequencing grade) at 25 °C for 5 h. Formic acid was used to stop the digestion. Peptides were extracted in the ultrasonication bath and desalted using C18 Stagetips (18).

Peptide quantities were matched by measuring HPLC-UV absorption at 214nm on a small fraction of the sample. Peptide analysis was performed on an UltiMate 3000 HPLC RSLC nano system (Thermo Fisher Scientific) coupled to a Q Exactive HF-X or an Orbitrap Fusion

Lumos mass spectrometer (Thermo Fisher Scientific), equipped with a Nanospray Flex ion source (Thermo Fisher Scientific). The trap column used was from Thermo Fisher Scientific (PepMap C18, 5 mm × 300 µm ID, 5 µm particles, 100 Å pore size at a flow rate of 25 µL/min (using 0.1% TFA as mobile phase). The analytical C18 column used after 10 min was also from Thermo Fisher Scientific (PepMap C18, 500 mm × 75 µm ID, 2 µm, 100 Å). Peptides were eluted applying a segmented linear gradient from 2% to 80% solvent B (80% acetonitrile, 0.1% formic acid; solvent A 0.1% formic acid) at a flow rate of 230 nL/min over 120 min.

The MaxQuant search was performed with full trypsin and chymotrypsin specificity and a maximum of three missed cleavages at a protein and peptide spectrum match false discovery rate of 1%. Carbamidomethylation of cysteine residues were set as fixed, oxidation of methionine and N-terminal acetylation as variable modifications. All other parameters were left at default.

To calculate the matrices used to demonstrate the crosslink-patterns observed, we chose an averaging window of n=4 and visualized the result with the "geom_tile" function of the R-package ggplot2 (19). Loop-links were not included, as short-range contacts would not yield additional information on the overall structure of αS. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (20) partner repository with the dataset identifier PXD027349.

In order to quantify the difference in the crosslink maps, a bootstrapping approach was applied (21). As we are interested in differences of longer-range contacts, the diagonal is eliminated by removing all PSMs that connect positions less than 6 residues apart in the sequence. In addition, all crosslinks beyond residue 120 are removed as these highly flexible residues are neither included in our models, nor do they show crosslinking hotspots and likely stem from the free form of the protein. On this remaining set we define regions that are

connected within a sequence window of three residues. To compare two maps we take the set of regions that is found in both maps and normalize the sum of crosslinks for each region in each map by the total number of unfiltered crosslinks in the respective map. An example of the region definition and resulting areas is shown in Figure S22. We then pull 5000 bootstrap samples from the corresponding regions. In this way, we get a normally distributed sample of the mean relative PSM count we can expect from both maps and can compare them as we always pull from the same set of regions in the same order for each. Based on these normal distributions we can then perform a student's t-test and calculate a Cohen's d measure (22, 23). The full set of pairwise comparisons for Figure 2 is given in Figure S22. This Figure also includes examples for the obtained distributions and the regions used for pulling bootstrap samples. It is interesting to note that the p-values of nearly all combinations indicates that the maps are very clearly distinct form one another, with the exception of the SDS- and Bicelle-bound soluble states that match each other to some degree (p=0.48). The Cohen's d, however, indicates that maps obtained from the mono- and dimer-bands of the liposome bound state also share greater similarity than other combinations. This similarity is due to the fact that both bands stem from the same sample and are differentiated in the SDS-PAGE only by the presence of intermolecular crosslinks. As the monomer band is likely to contain some protein that adopted the compact state and shared in an oligomer, but did not obtain an intermolecular link, the change between the two crosslink maps reflects an enrichment of the compact state going from the mono- to the dimeric form, rather than a fully changed underlying ensemble. We chose to focus on data from the dimer band both, because it is enriched in the compact state and because it is the dataset that also provides us with intermolecular links used in the modelling of the oligomeric state.

Bicelle production for crosslinking was carried out as described in the "Methods" part under the "NMR"-section of this document, but to a final mix of 10mM DHPC and 5mM DMPG and in the buffer pH=6.5 as stated in the paper.

NMR relaxation

Relaxation measurements for PREs were obtained with delay times of 1, 5, 10, 20, 30 and 50 milliseconds in and interleaved fashion and read out using HN correlation spectra. These experiments were measured on an AVANCE NEO 800 MHz instrument equipped with a triple resonance probe and data was processed with the NMR-Pipe suite and Sparky3 for relaxation fits. Cysteine point mutants (19C, 42C, 81C) were generated using the QuickChangeII site directed mutagenesis kit (Stratagene) and spin labelling (labelling with MTSL (S-(2,2,5,5-tetramethyl-2,5-dihydro-^{1}H-pyrrol-3-yl)methyl methane-sulfonothioate) was carried out as described in (24).

Structural Computations, PRI Potential

A novel energy term for Xplor-NIH was developed to allow the use of experimental PRI values as restraints during structure calculation. As PRIs ultimately rely on Dipole-Dipole cross-correlation effects, their sign is dependent on Eq pri.1 with theta being the angle spanned between the measured nucleus and both paramagnetic spin labels.

$$PRI \approx 3cos^2\theta - 1 \qquad \text{Eq pri.1}$$

Independent of the distance between the nucleus and the respective labels, it is possible to differentiate between 2 broad angle segments that lead to either positive or negative PRI

values. Accordingly, experimental values (ePRI) are separated into 2 classes and translated into a converted PRI (cPRI, ePRI<0 → cPRI=-1, ePRI>0→cPRI=1). For each constraint, Eq pri.1 is calculated and in the case of multiple label conformers averaged over all combinations and further ensemble averaged with the respective weights if necessary. The energy term takes the form of Eq pri.2 with S being a scaling factor.

$$E_{PRI} = S * \left\{ \begin{matrix} \frac{3}{4}|3cos^2\theta - 1| - \frac{3cos^2\theta - 1}{4}, if\ cPRI * ePRI < 0 \\ 0, otherwise \end{matrix} \right\}$$

Molecular Structure Determination

Molecular structure calculations were carried out using the molecular structure determination package Xplor-NIH (25, 26). A 4-stage procedure based on previous examples was used (27). Starting from an extended structure, 150 ps of high temperature torsion angle dynamics at 3000 K were carried out followed by simulated annealing to 25 K in steps of 12.5 K and 0.4 ps each. Apart from standard potentials, the prePot (28) for paramagnetic relaxation enhancement data, the novel priPot potential, as well as the intrinsic solvent potential eefx were used (27, 29, 30). The code of the priPot potential as well as instructions for its use and application can be found in a supplementary data file (priPotential.txt). In order to prevent unresolvable atomic overlap in the high temperature stage, it was separated into 3 blocks of 50 ps each. In the initial block, only CA-CA interactions based on the repel potential (31) were turned on with a force constant of 0.004 kcal $mol^{-1}A^{-4}$. In a second block an all atom repel potential with a force constant of 0.1 kcal $mol^{-1}A^{-4}$ was used. In the final high temperature block, the eefx potential was used with a force constant of 0.1 kcal mol^{-1} A^{-4}. This scaling factor was ramped to a final value of 1 kcal mol^{-1} A^{-4} during simulated annealing. prePot and priPot scaling factors were set to 0.1 and 15

respectively during the high temperature phase and ramped to a final value of 1 and 150 respectively. To further restrain secondary structures, Xplor's CDIH potential (25) was incorporated into the calculation. Helical constraints were generated based on representative helical phi and psi values of -57.8 and -47.0 respectively and an allowed deviation of +/- 5 degrees with a weighting factor derived from chemical shift based SSP calculations. The scaling factor in this case was increased from 2 in the high temperature phase to a final value of 20 during simulated annealing. The simulated annealing phase in turn is succeeded by a final powell minimization in torsion and cartesian space. The best structures derived from an extended conformation based on the total energy of the system, are further refined using a shortened protocol using only a single 50 ps high temperature phase similar to the final phase in the initial protocol. For the dimer calculations, two chains of the cluster representatives were initialized and their position was randomized. In an initial 50 ps high temperature phase, the structures were fixed and in the subsequent annealing phase, while they were allowed to change, the posDiff potential based on distsympot (32) enforced the same relative configuration on both. XL constraints as described above were incorporated as a pseudo NOE potential, with an initial scaling factor 1.0 during the high temperature phase that was ramped to a final value of 400 during simulated annealing. In each of the 3 groups of crosslinks used, only a single one must be fulfilled in an any-to-any form. A graphical summary of the model generation workflow is provided in Figure S23.

CMScore Calculations

To quantify the agreement between our calculated structures with the XL-MS measurements we applied an approach that is described as CMScore (4) and is available as an R package (XLmap). In short it calculates the sum of the ratios between the total number of detected XL

residue pairs and the number of these pairs covered given a certain distance cutoff, over a range of integer distance bins. We adopted this approach and extended it to also cover intermolecular cases. In all our CMScore calculations we excluded XL-sites with less than 3 reported PSMs. In the intermolecular cases we also employed a diagonal filter in which we excluded XL-sites that are separated by less than 5 residues along the sequence. For the intermolecular cases, as we are working with homodimers, the distance maps were mirrored by taking the minimum value for a given residue pair form either the upper or lower triangle of the distance matrix.

SUPPLEMENTARY INFORMATION REFERENCES

1. D. Sulzer, R. H. Edwards, The physiological role of alpha-synuclein and its relationship to Parkinson's Disease. *Journal of neurochemistry* **150**, 475-486 (2019).
2. K. E. Carnazza *et al.*, Synaptic vesicle binding of alpha-synuclein is modulated by beta- and gamma-synucleins. *Cell Rep* **39**, 110675 (2022).
3. J. Pan, A. Dalzini, N. K. Khadka, C. M. Aryal, L. Song, Lipid Extraction by α-Synuclein Generates Semi-Transmembrane Defects and Lipoprotein Nanoparticles. *ACS omega* **3**, 9586-9597 (2018).
4. D. K. Schweppe, J. D. Chavez, J. E. Bruce, XLmap: an R package to visualize and score protein structure models based on sites of protein cross-linking. *Bioinformatics* **32**, 306-308 (2016).
5. G. Fusco *et al.*, Structural basis of membrane disruption and cellular toxicity by alpha-synuclein oligomers. *Science* **358**, 1440-1443 (2017).
6. P. Kumar, J. N. Kizhakkedathu, S. K. Straus, Antimicrobial Peptides: Diversity, Mechanism of Action and Strategies to Improve the Activity and Biocompatibility In Vivo. *Biomolecules* **8** (2018).
7. E. Sancho-Vaello *et al.*, The structure of the antimicrobial human cathelicidin LL-37 shows oligomerization and channel formation in the presence of membrane mimics. *Scientific reports* **10**, 17356 (2020).
8. P. Alam, L. Bousset, R. Melki, D. E. Otzen, α-synuclein oligomers and fibrils: a spectrum of species, a spectrum of toxicities. *Journal of neurochemistry* **150**, 522-534 (2019).
9. B. B. Holmes *et al.*, Heparan sulfate proteoglycans mediate internalization and propagation of specific proteopathic seeds. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E3138-3147 (2013).
10. G. Fusco *et al.*, Direct observation of the three regions in alpha-synuclein that determine its membrane-bound behaviour. *Nature communications* **5**, 3827 (2014).
11. C. R. Bodner, C. M. Dobson, A. Bax, Multiple tight phospholipid-binding modes of alpha-synuclein revealed by solution NMR spectroscopy. *Journal of molecular biology* **390**, 775-790 (2009).
12. T. Viennet *et al.*, Structural insights from lipid-bilayer nanodiscs link α-Synuclein membrane-binding modes to amyloid fibril formation. *Communications biology* **1**, 44 (2018).
13. N. L. Fawzi, J. Ying, D. A. Torchia, G. M. Clore, Probing exchange kinetics and atomic resolution dynamics in high-molecular-weight complexes using dark-state exchange saturation transfer NMR spectroscopy. *Nat Protoc* **7**, 1523-1533 (2012).
14. Z. Solyom *et al.*, BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. *Journal of biomolecular NMR* **55**, 311-321 (2013).
15. A. Favier, B. Brutscher, Recovering lost magnetization: polarization enhancement in biomolecular NMR. *Journal of biomolecular NMR* **49**, 9-15 (2011).
16. J. Schleucher *et al.*, A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *Journal of biomolecular NMR* **4**, 301-306 (1994).

17. P. Schanda, B. Brutscher, Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *Journal of the American Chemical Society* **127**, 8014-8015 (2005).
18. J. Rappsilber, M. Mann, Y. Ishihama, Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2**, 1896-1906 (2007).
19. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*, Ggplot2: Elegant Graphics for Data Analysis (2009), 10.1007/978-0-387-98141-3, pp. 1-212.
20. Y. Perez-Riverol *et al.*, The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **47**, D442-D450 (2019).
21. B. Efron, R. J. Tibshirani, *An Introduction to the Bootstrap* (Chapman and Hall / CRC, New York, ed. 1, 1994), https://doi.org/10.1201/9780429246593.
22. Student, The Probable Error of a Mean. *Biometrika* **6**, 1-25 (1908).
23. J. Cohen, *Statistical Power analysis for the Behavioral Sciences* (Academic Press, Inc., New York, 1977), https://doi.org/10.1016/C2013-0-10517-X.
24. C. C. Jao, A. Der-Sarkissian, J. Chen, R. Langen, Structure of membrane-bound alpha-synuclein studied by site-directed spin labeling. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 8331-8336 (2004).
25. C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, The Xplor-NIH NMR molecular structure determination package. *Journal of magnetic resonance (San Diego, Calif. : 1997)* **160**, 65-73 (2003).
26. C. D. Schwieters, J. J. Kuszewski, G. Marius Clore, Using Xplor–NIH for NMR molecular structure determination. *Progress in Nuclear Magnetic Resonance Spectroscopy* **48**, 47-62 (2006).
27. Y. Tian, C. D. Schwieters, S. J. Opella, F. M. Marassi, A Practical Implicit Membrane Potential for NMR Structure Calculations of Membrane Proteins. *Biophysical journal* **109**, 574-585 (2015).
28. J. Iwahara, C. D. Schwieters, G. M. Clore, Ensemble approach for NMR structure refinement against (1)H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. *Journal of the American Chemical Society* **126**, 5879-5896 (2004).
29. Y. Tian, C. D. Schwieters, S. J. Opella, F. M. Marassi, A practical implicit solvent potential for NMR structure calculation. *Journal of magnetic resonance (San Diego, Calif. : 1997)* **243**, 54-64 (2014).
30. Y. Tian, C. D. Schwieters, S. J. Opella, F. M. Marassi, High quality NMR structures: a new force field with implicit water and membrane solvation for Xplor-NIH. *Journal of biomolecular NMR* **67**, 35-49 (2017).
31. M. Nilges, G. M. Clore, A. M. Gronenborn, Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS letters* **229**, 317-324 (1988).
32. M. Nilges, A calculation strategy for the structure determination of symmetric dimers by 1H NMR. *Proteins* **17**, 297-309 (1993).
33. F. Madeira *et al.*, The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic acids research* **47**, W636-w641 (2019).
34. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671-675 (2012).

Dataset S1 (separate file). An excel data file including data used for generation of the figures shown in this publication, both main and supplementary material. Names of data sheets indicate the figure the data was used for. Structural data, chemical shift data and additional XLMS information was deposited in the respective online repositories as indicated.

PriPotential (separate file). Contains the code used with Xplor-NIH to use the experimental PRI-Data as restraints in structure calculations. Instructions for use are given at the top followed by the code tested with Xplor-NIH 2.49 and 3.3.

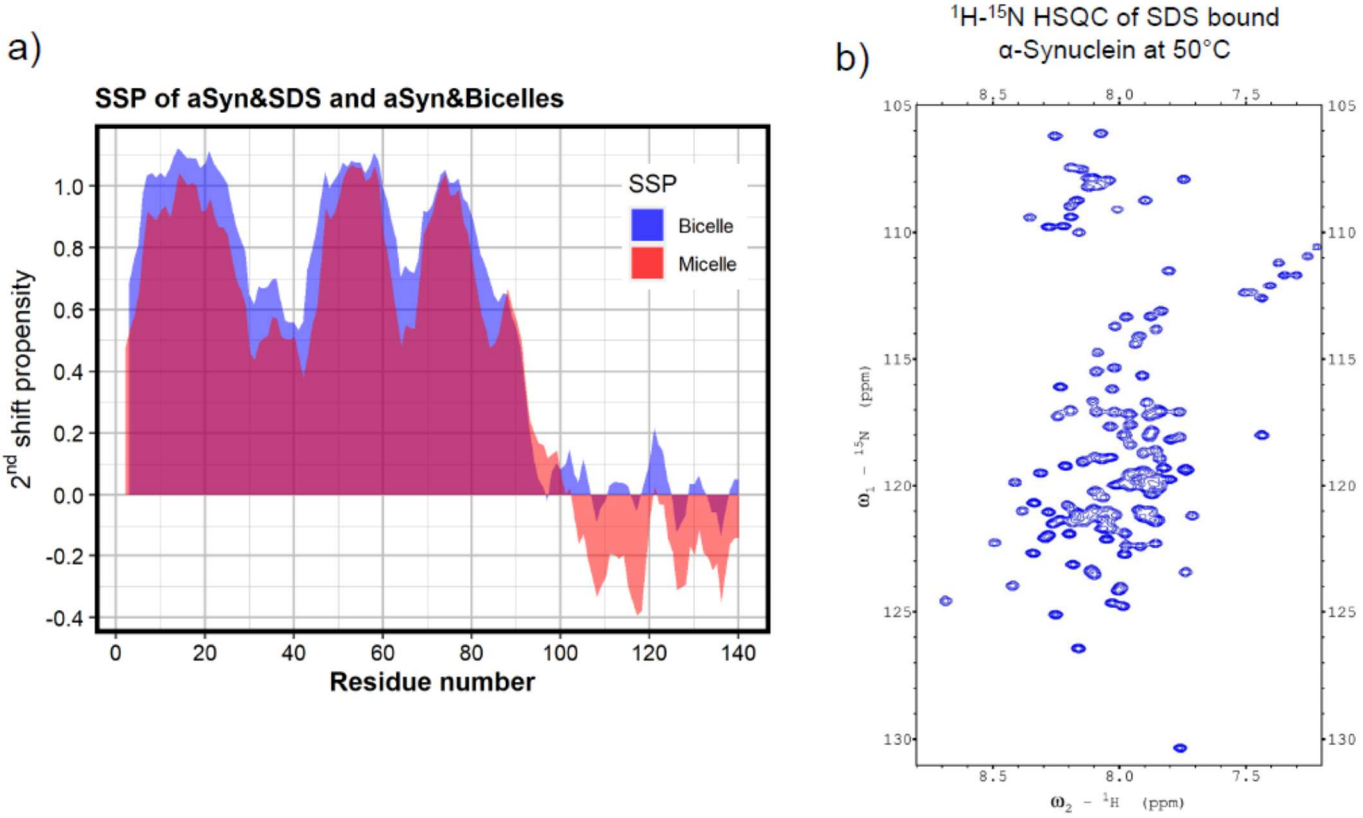


Figure S1. NMR secondary shift data of αS in varying conditions. (a) The strong overall helical tendencies of αS in its membrane bound state are observed for the SDS-micelle bound (BMRB-ID 50996) as well as for the bicelle bound form (BMRB-ID 51148). **(b)** The ^{1}H-^{15}N HSQC of αS in the conditions used for PRE-measurements shows clear signals and good peak separation, allowing us to use HSQC based relaxation measurements.

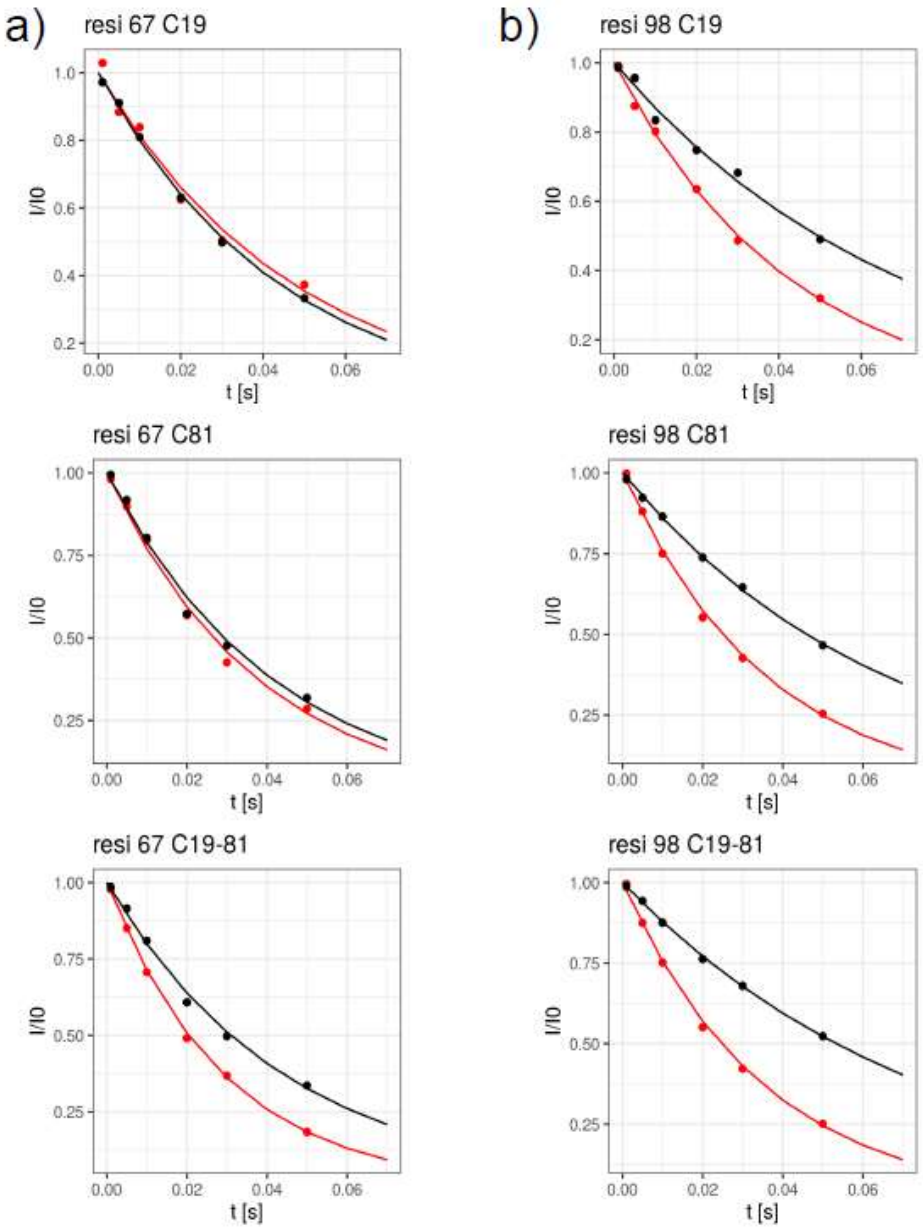


Figure S2. Signal decay in ^{1}H-T2 measurements.

Exemplary signal decay rates of two residues in various measurements. The top two rows are measured with single cysteine mutants. The bottom row is measured with the corresponding doubly labelled cysteine mutant. Red curves are measured with an active spin label, while black curves were measured after quenching of MTSL. Thus, the difference between "red" and "black" rates gives the PRE-effect. The difference between the sum of the two single mutants' PRE and the double mutant's PRE yields the PRI-effect, which can be positive as in **(a)** or negative as in **(b)**.

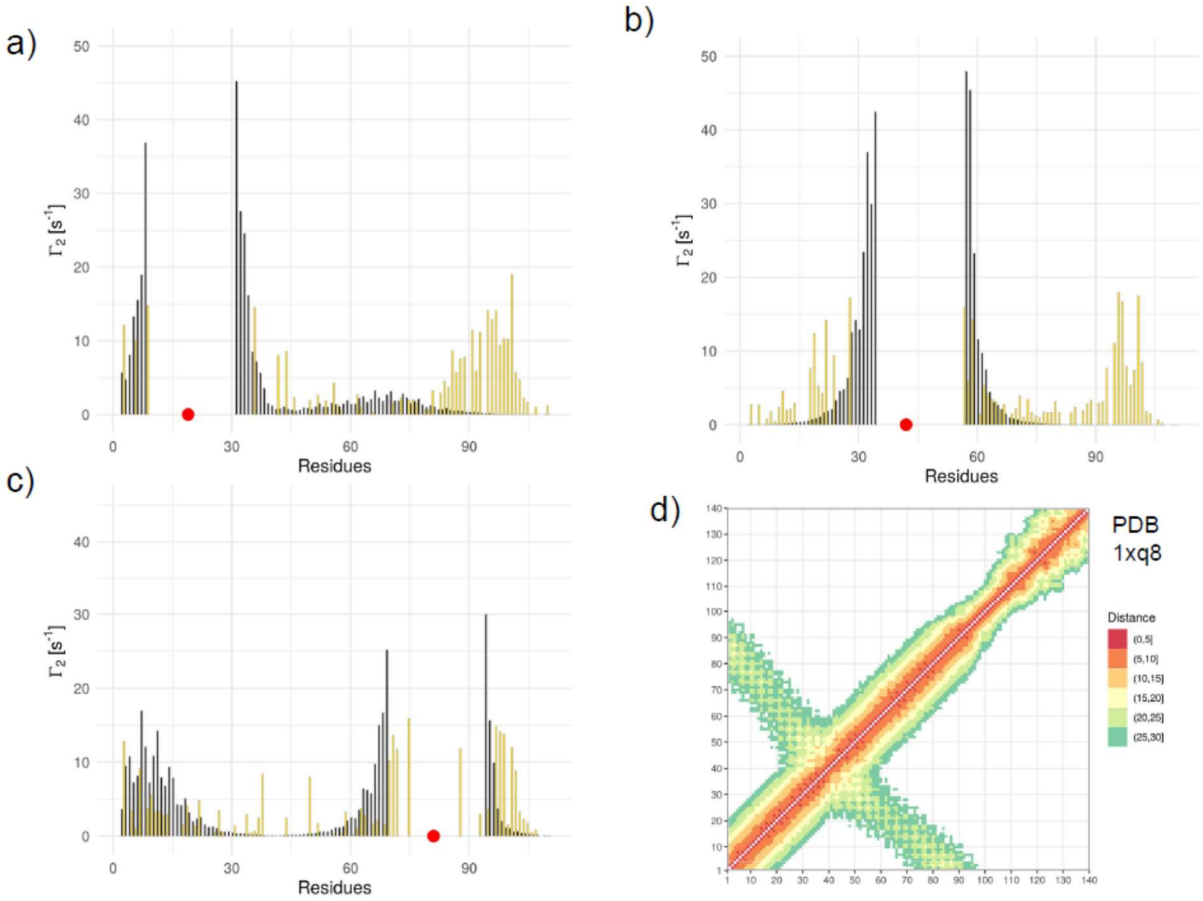


Figure S3. Comparing expected relaxation rates with experimental data. Black bars indicate the relaxation rates expected when calculated using PDB: 1XQ8 as input structure for MTSL spin labels at positions 19 **(a)**, 42 **(b)** and 87 **(c)**. The observed rates are labelled in yellow. **(d)** Intramolecular CA-CA distance map based on pdb ID 1XQ8.

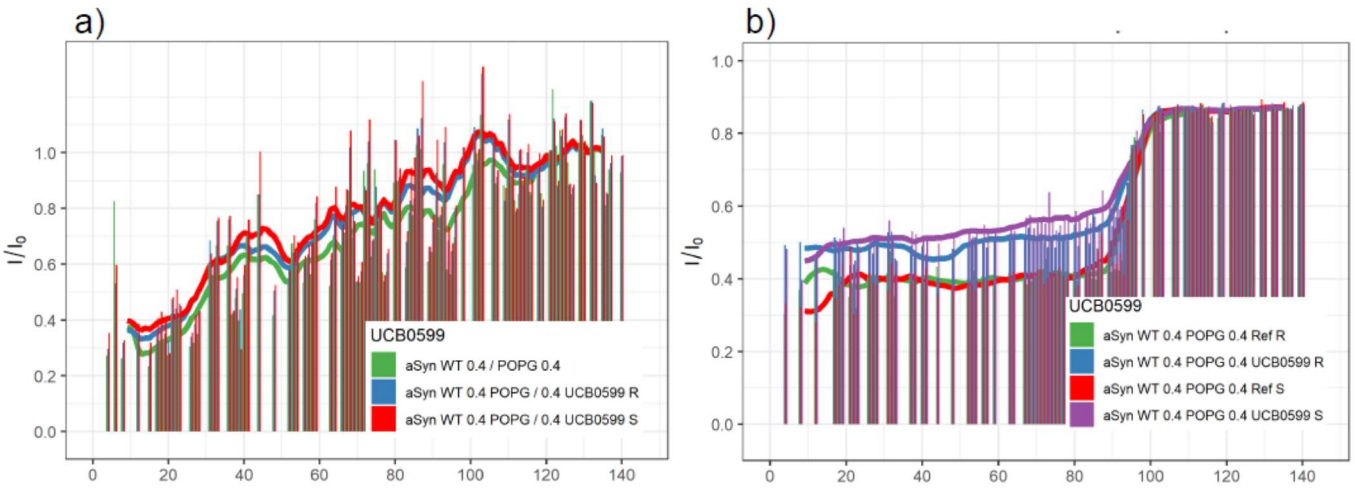


Figure S4. Enantiomers of UCB0599 are not distinguished by their action on αS liposome binding. ^{1}H-^{15}N HSQC intensities of liposome bound αS with (R blue, S red) and without (green) 0.1mM of the two enantiomers of UCB0599 are compared in **(a)** with a nine residue window average shown as a line. All detected signal stems from unbound protein regions and reductions relative to conditions without liposomes (I_0) indicate protein binding and, thus, signal loss. **(b)** The intensity of DEST signals (shown as ratio between irradiations at 30.000 Hz and +/-3.000 Hz) is plotted. For both R and S enantiomers a measurement of the sample prior to addition of the compound was carried out and is shown for reference.

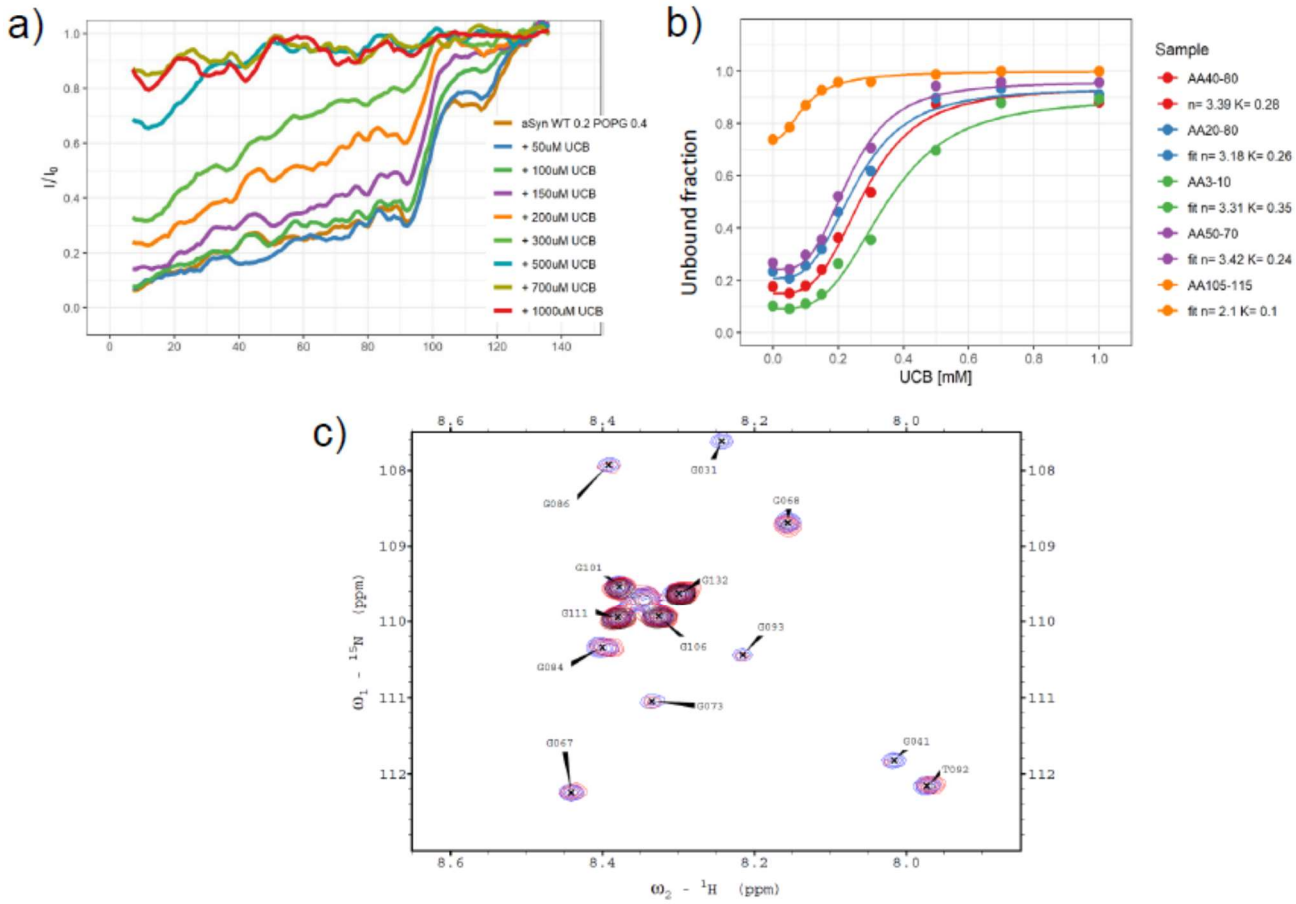


Figure S5. Tracking uniform UCB0599 initiated displacement along the primary sequence of αS. Tracking the effect of UCB0599 in the main Figure 6 is carried out through the relative displacement achieved. Here we show an example of the data these values were calculated from. **(a)** Shows the relative intensity of peaks when comparing the recoded value with αS measured without any liposomes present in solution. All detected signal stems from the free protein regions and a loss in signal indicates binding. Shown is the titration up to 1mM UCB0599 with observed intensities averaged over a nine residue sliding window. Data in **(a)** is from the titration of 0.2mM αS bound to 0.4mg/ml of POPG based liposomes. To show the influence of fitting the displacement effect within various regions of the protein to the Hill-function **(b)** depicts fits of the areas indicated (by amino acid range). Here the uniformity of the displacement can be clearly appreciated, a good indication for full displacement, rather than weakening affinity. An example of peak intensities as observed by ^{15}N-^{1}H HSQC measurements is given in **(c)**, which shows the glycine region of the titration with the condition without liposomes depicted in blue, while αS signal with liposomes present is shown in red.

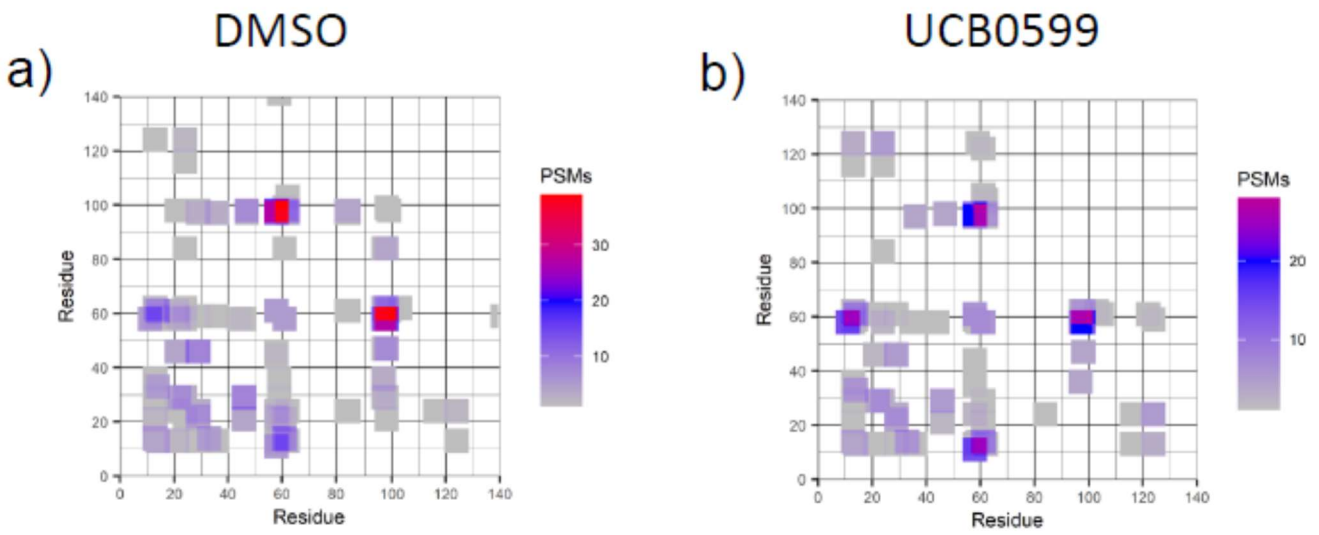


Figure S6. PSM maps of UCB0599 and DMSO containing αS and liposome preparations.

In order to generate the difference map containing the UCB0599 effect shown in Figure 6, PSM correlation maps of the individual conditions were constructed as in Figure 2. The number of PSMs found in a four-residue window is indicated in the colour scales. **(a)** shows the pattern observed in conditions containing DMSO, while **(b)** shows the data recorded when UCB0599 was added in non-displacing quantities. The changes in the intermolecular PSMs can be appreciated even prior to constructing a difference plot, with (b) showing an increase in the links from regions around 60 to15 relative to (a) despite an overall reduction in crosslinks, while links from region 95 to 60 are strongly reduced.

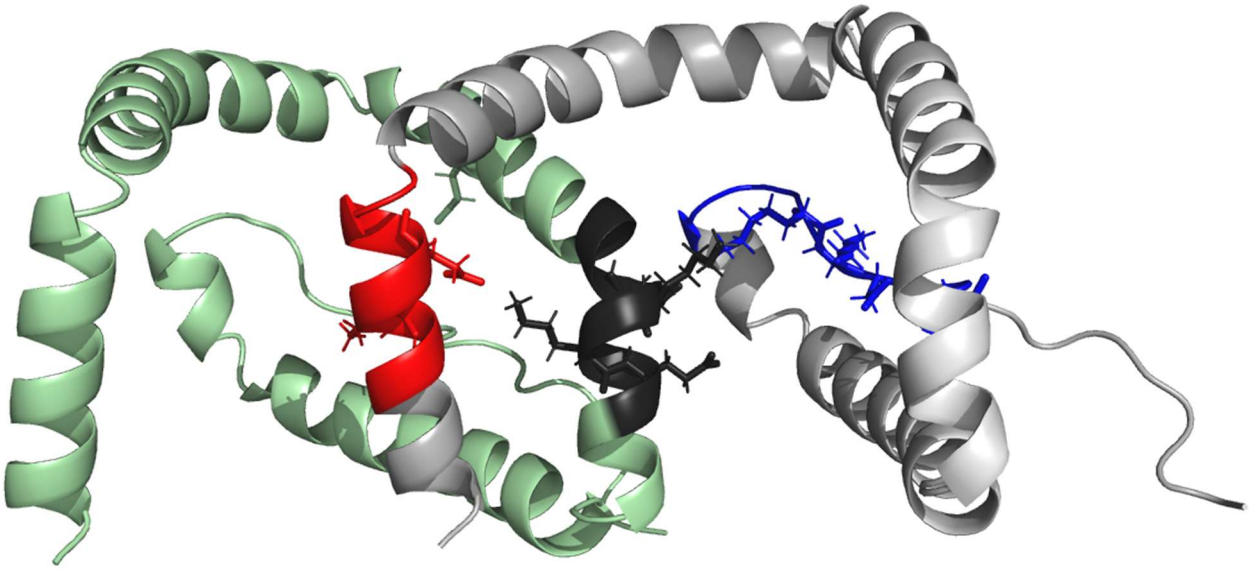


Figure S7. Localisation of the UCB0599 effect on αS dimer structure.

The influence of UCB0599 is seen in PSM maps generated from the difference in PSMs obtained with and without the compound locates to specific regions within αS (Figure 6d). These regions were mapped onto the ring-like dimer structure with a similar color code as indicated in the PSM map. Here we show the protein in the same arrangement as in Figure 4a (right) which depicts two full units of the protein. Interacting regions are colored as follows: blue (91-101), purple (55-62) and orange (8-16) with sidechains of lysine and glutamate residues indicated.

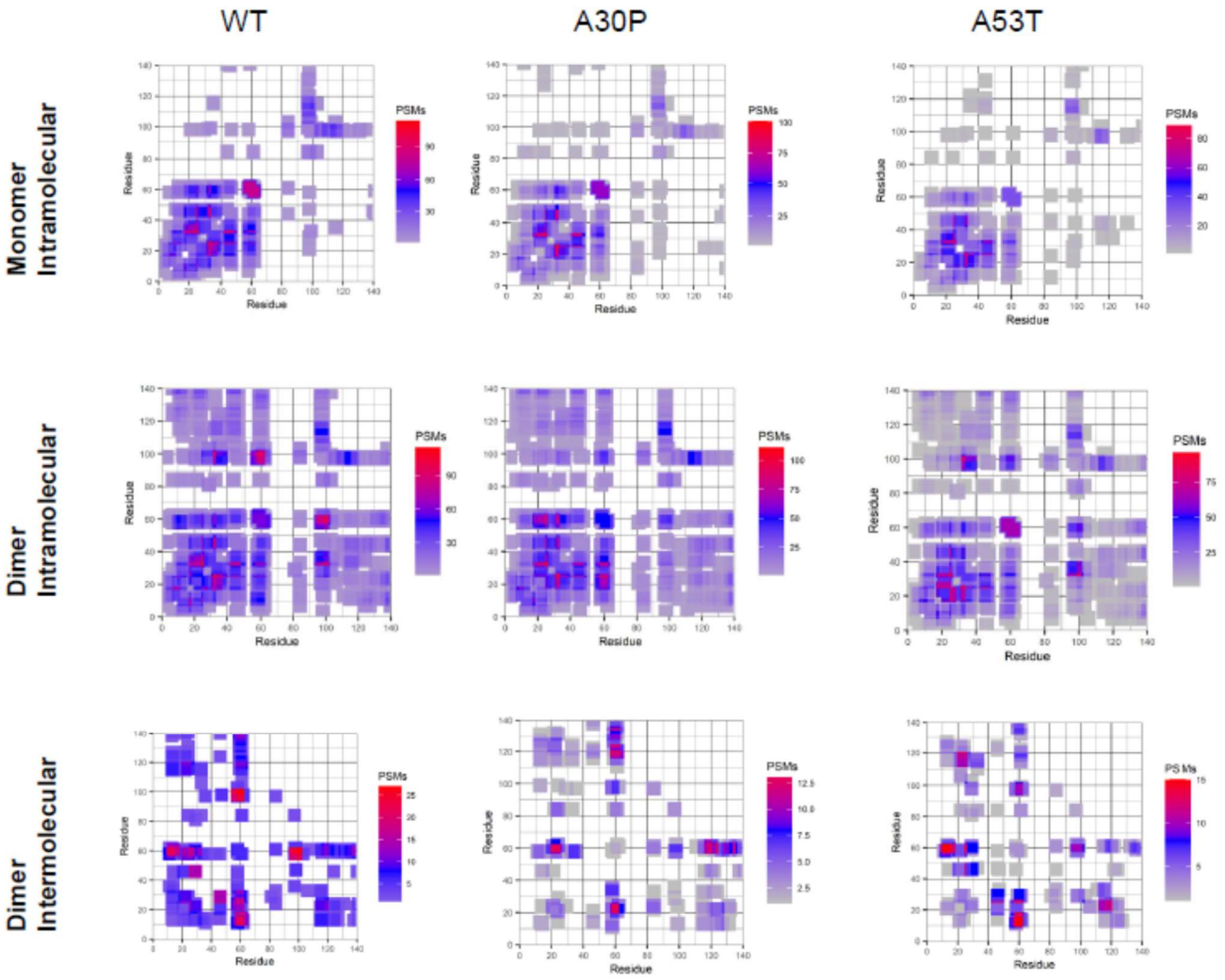


Figure S8. The effect of αS point mutations on its crosslinking patterns. The columns show the crosslink maps for WT, A30P und A53T, with rows indicating the monomeric intramolecular, dimeric intramolecular and dimeric intermolecular crosslinks. The number of PSMs found in a four-residue window is indicated in the colour scales. Changes between the monomers (top row) are global in nature, indicating similarity in structure for membrane bound monomers. More variation is observed in intramolecular links of dimers (middle row). Here A30P shows a decrease in PSMs of regions (50-65) to (90-100) as well as (30-40) and (90-105). Intermolecular crosslink maps (bottom row) show the most variation. Increases in crosslinks between regions (110-140) to (55-65) in A30P and (15-25) to (115-125) in A53T are clearly observed, indicating changes in the oligomeric states.

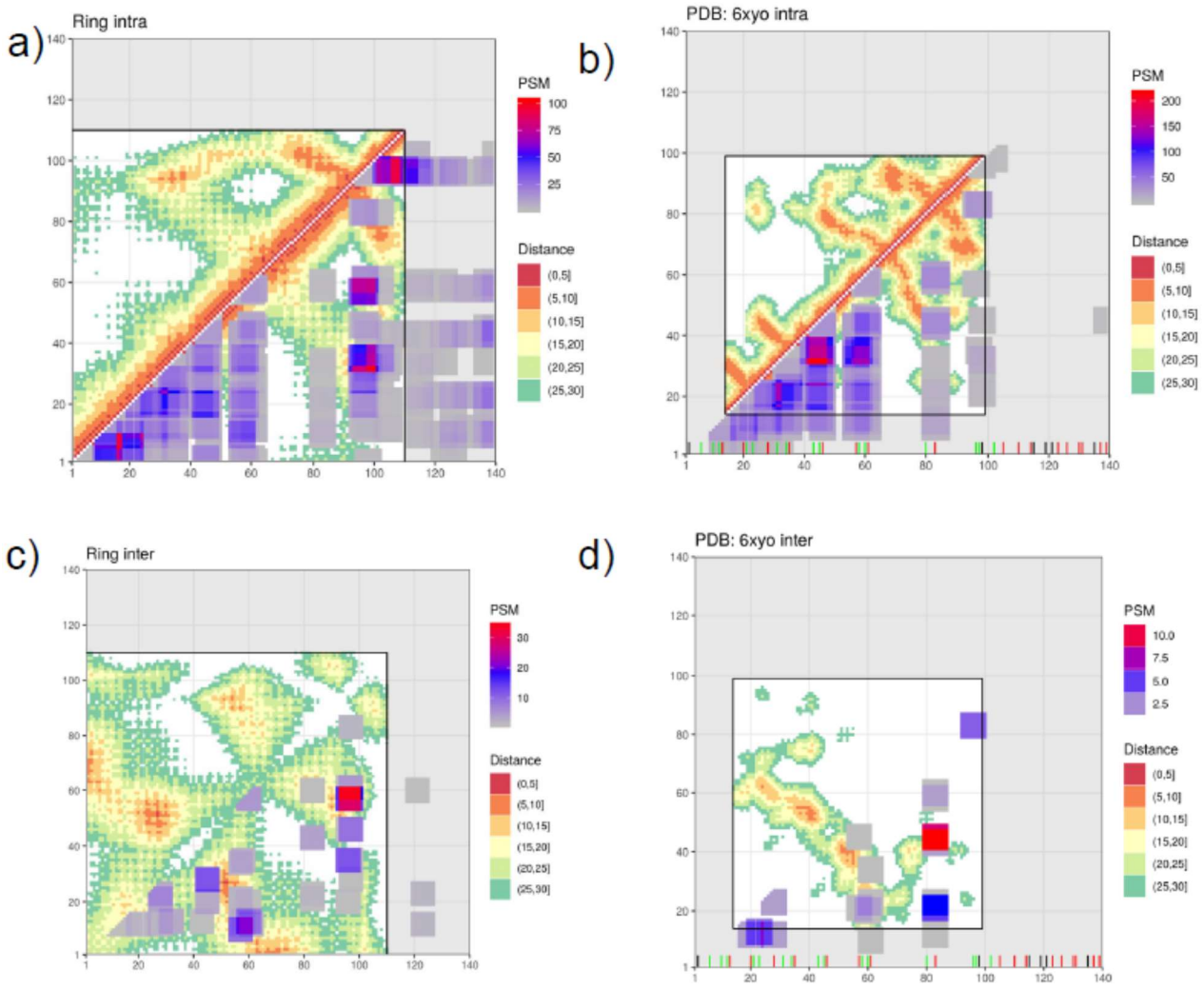


Figure S9. Comparison of CA-CA distance maps and PSMs from XL-MS. Overlay of intra- **(a)** and intermolecular **(c)** crosslink PSMs of the dimer liposome preparation with the respective CA-CA distance maps of the structure shown in Figure 4a. The PSM colour scales indicate PSMs found in a four-residue window, while the distance colour scale indicates CA-CA distance windows in nm. Overlay of intra- **(b)** and intermolecular **(d)** crosslink PSMs identified for bicelle derived aggregates and respective CA-CA distance maps for the pdb ID 6XYO. Intramolecular distances (b) are calculated as the minimal distance inside a single chain in either layer/leaflet of the fibril. Intermolecular distances (d) are calculated as the distance between CA atoms of chains directly next to one another in opposite leaflets of the fibril. Positions or residues capable of crosslink formation are indicated at the bottom with green ticks representing lysines, while red and black ticks show the position of glutamate and aspartate residues, respectively.

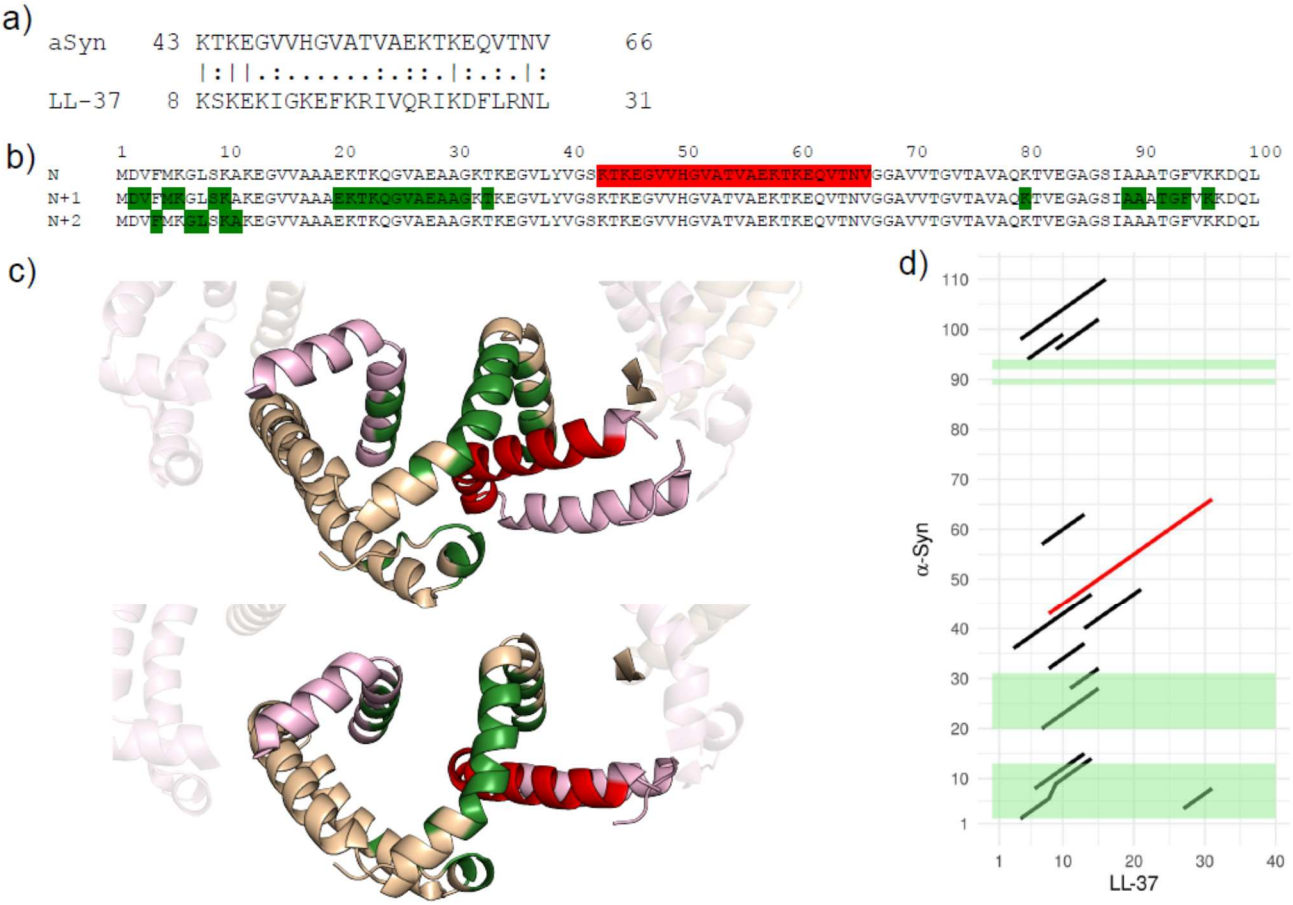


Figure S10. Sequence similarities between αS and the predominantly α-helical antimicrobial peptide LL-37. (a) Longest detected local sequence alignment between LL-37 and αS generated by lalign (33). **(b, c)** Sequence view (b) and structure (c) where the longest local alignment is highlighted in red for the ring like oligomer structure (Fig.5a). Regions in neighbouring chains closer than 0.7 nm to this stretch are highlighted in green. **(d)** Overview of all local alignments reported by lalign (33). The longest stretch is highlighted in red. The areas close in space for the ring like structures are highlighted by green shading.

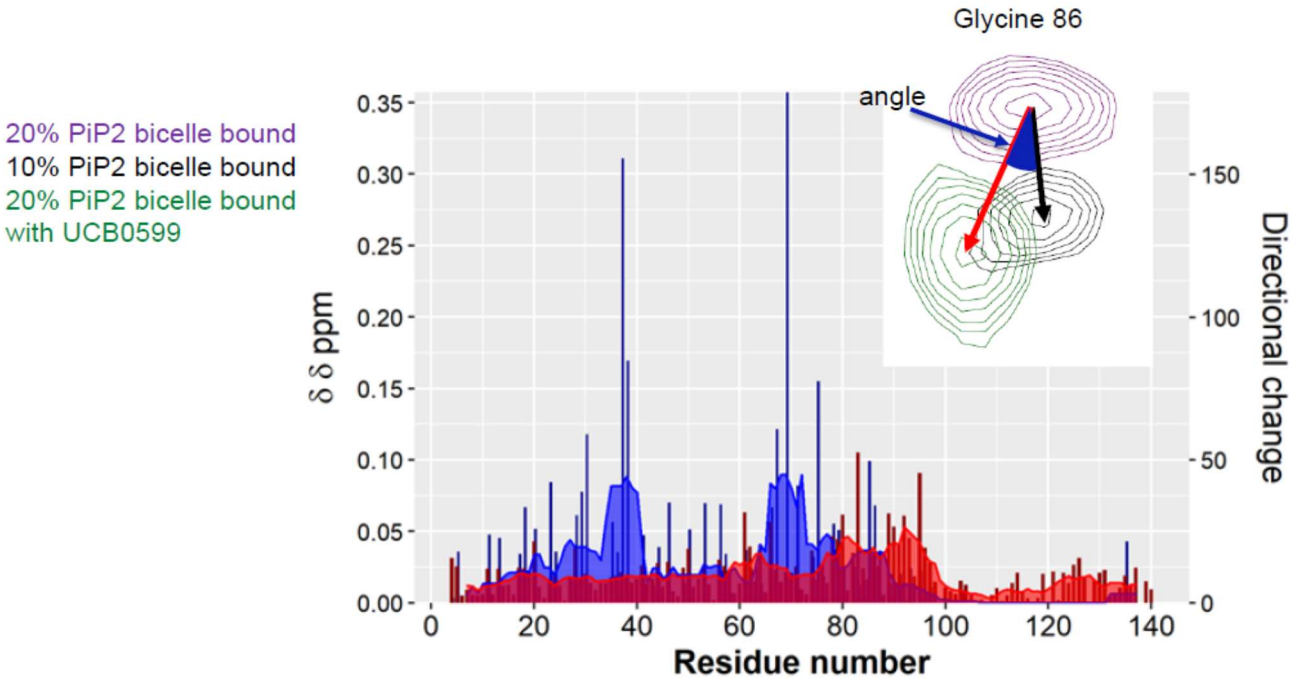


Figure S11. Effects of UCB0599 on bicelle bound αS as observed in ^{1}H-^{15}N correlations. The inset on the top right shows peaks for three conditions given in the same color code as their descriptions in the top left. The red arrow indicates the shift tracked in the residue plot with red colors. Corresponding shift values are indicated on the left (primary) y-axis. The angle between shifts caused by UCB0599 and those caused by a reduction in PiP_2 concentration is indicated in blue both in the inset and in the residue plot. Changes in direction are given in degrees and shown at the right hand along the (secondary) y-axis. For both indicators a rolling average over a nine-residue window is also given as red and blue shaded areas.

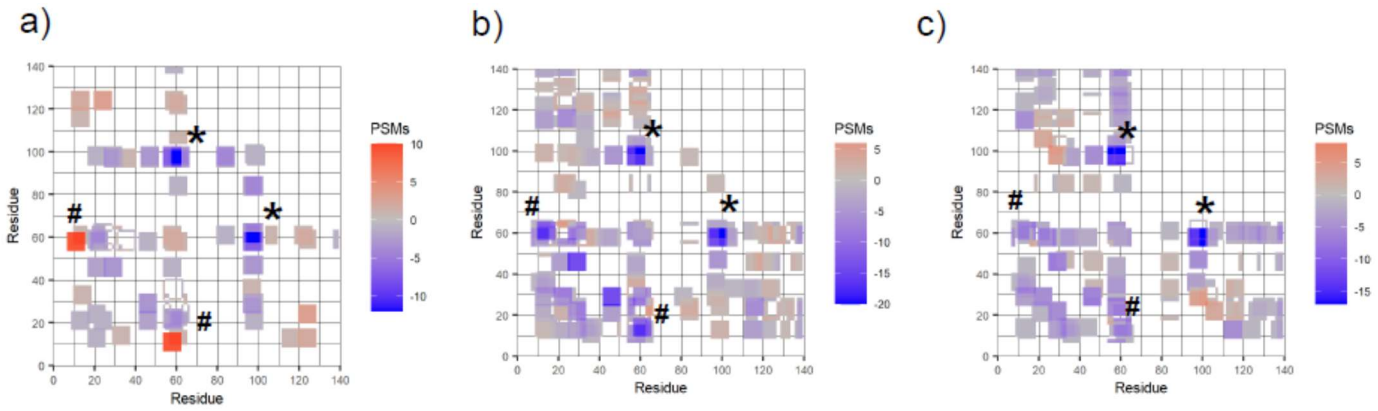


Figure S12. Comparison of point mutation and UCB0599 induced effects. (a) A reproduction of Figure 6d, where the changes in the intermolecular crosslinking pattern of liposome bound αS are shown. Here we also present the same analysis for the comparison of point mutations A30P **(b)** and A53T **(c)** to the wildtype protein. The number of PSMs found in a four-residue window is indicated in the colour scales. Regions most strongly influenced by UCB0599 are marked as in Fig.6 [(92-101) to (58-62), with *] and [(54-61) to (8-16), with #]. These same regions are among the most strongly influenced, as would be expected for shifts in the bound oligomer ensemble. The dissimilar pattern of increasing and decreasing PSMs, however, indicates that the ensemble is shifted by mutations in a different way than observed for the compound. (Data used for the difference plots in (b) and (c) is the same as shown in Figure S8.)

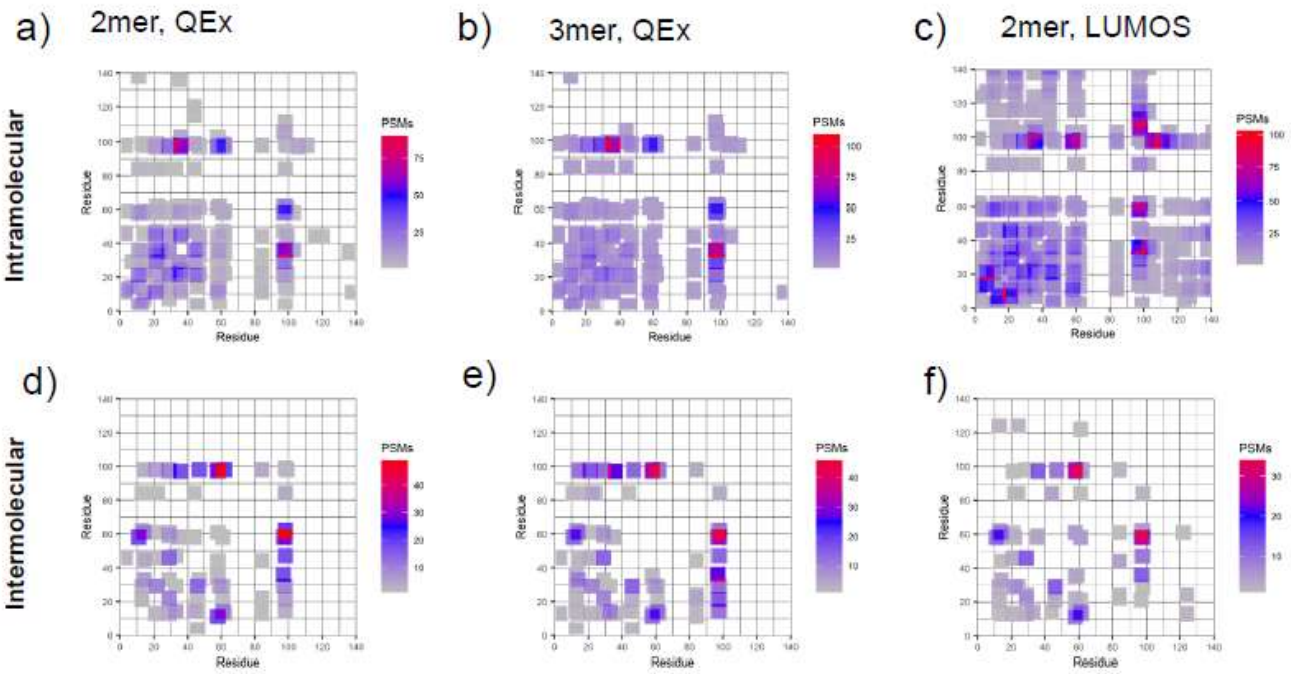


Figure S13. Comparison of PSM maps generated with different instruments. The rows indicate whether intramolecular (top) or intermolecular (bottom) PSMs were processed. Columns indicate measurement of samples either cut from dimer or trimer bands and on what instrument the measurement was carried out (QEx: Q Exactive HF-X mass spectrometer; LUMOS: Orbitrap Fusion Lumos mass spectrometer). The colour scale indicates how many links were found in a four-residue window. Although differences between the measurements are apparent, the most dominant regions of interactions are highly reproducible within a single gel, and also correlate very nicely even between different instruments. As measurements carried out on the two instruments used samples from differing crosslinking reactions, variation between different samples is also shown to be small. In addition, the similarity between dimeric and trimeric bands also provides further evidence that oligomers are formed by addition of more or less structurally equivalent monomers.

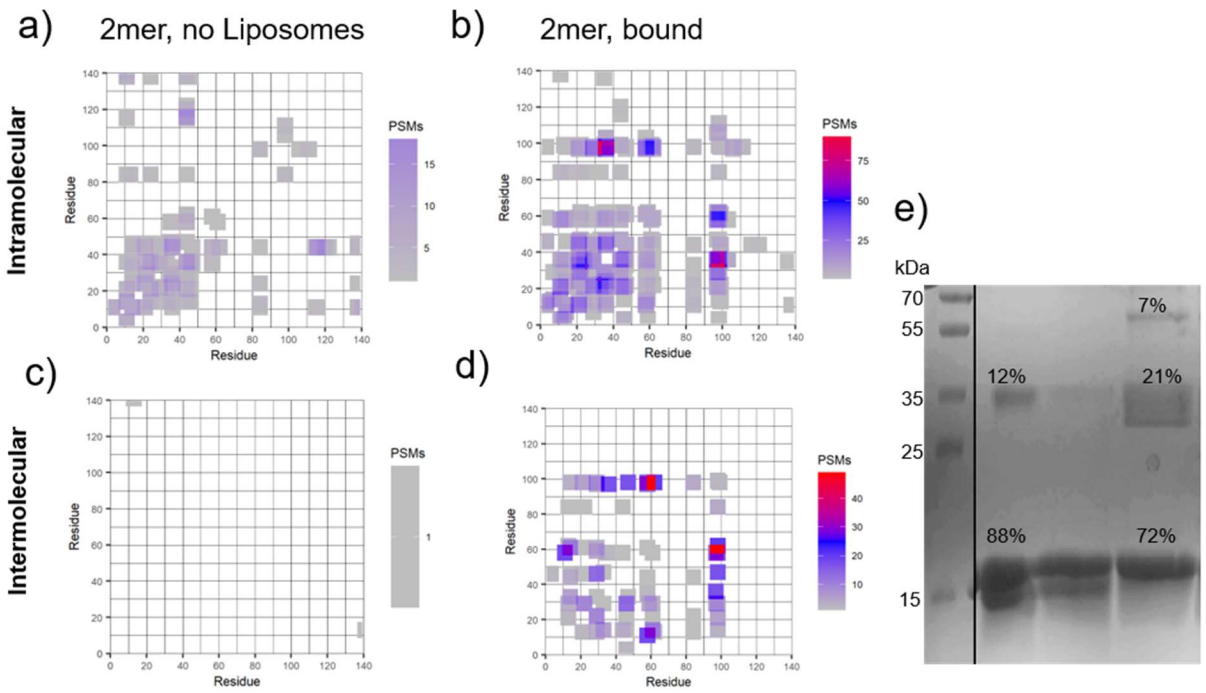


Figure S14. PSM maps indicating a possible background for inter and intramolecular crosslinking reactions and corresponding PAGE gel. PSM maps show intramolecular **(a, b)** and intermolecular **(c, d)** measurements. Here, columns show measurements of unbound αS (a, c) and liposome bound αS (b, d). When crosslinking αS at elevated concentration (200µM, without liposomes) as compared to standard measurements (50µM, with liposomes), a dimer band can be induced without liposomes (e, lane 2). Only a single PSM is detected in the intermolecular crosslinks of this species. Intramolecular crosslinks are also heavily reduced, indicating that only a very low background of crosslinks stemming from the unbound species is to be expected in our measurements. The PAGE gel in **(e)** shows the marker in lane 1. The black line indicates omitted lanes (full gel in supplementary data file). In lane 2 we can see the effect of crosslinking 200uM of αS in the absence of liposomes (the quantity loaded is the same in all lanes). Lane 3 shows the result of a crosslinking reaction at standard concentration (50µM) without liposomes, while in lane 4 we see crosslinks obtained in our standard XL-MS conditions. Analysis with ImageJ (34) indicates that crosslinks obtained by high concentration only lead to ~12% of the protein being detect in the dimer, while crosslinking with POPG lead to ~28% in the dimer and trimer. In both cases the majority of detected protein is still monomeric, despite nearly all protein being bound in the liposome containing conditions.

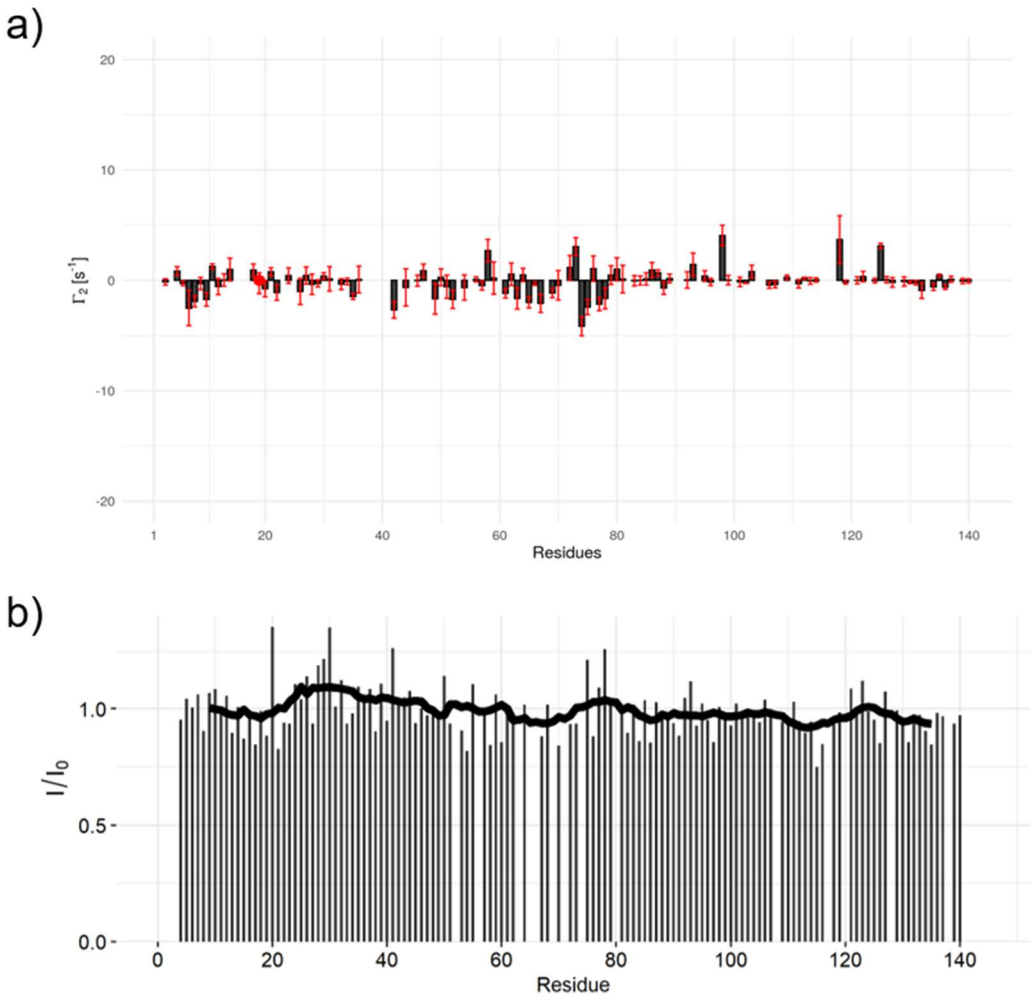


Figure S15. Intermolecular PREs as detected on SDS- as well as bicelle-bound αS. The αS mutant 19C was tagged with MTSL and measured in a 1:1 mixture with unlabeled WT protein. In **(a)** relaxation rates of the SDS-bound sample are depicted along the primary sequence of the protein. Although higher noise levels are observed as compared to those measured for intramolecular PREs, this is expected since only half the protein is NMR active in this 1:1 mixture. The PRE rates are low and outliers are observed in both directions without a clear pattern, indicating that no intermolecular PREs contribute to the measurements used in this paper. In **(b)** we show a measurement of the same setup and conditions (pH, temperature) when bound to bicelles. Due to the larger size a faster relaxation and lower S/N ratio is obtained. Therefore, we measured the relative intensity, rather than relaxation rates, in this case. We do, however, still clearly observe that reduction of the radical does not appreciably change the measured signal intensity and, thus, no intermolecular PRE is detected.

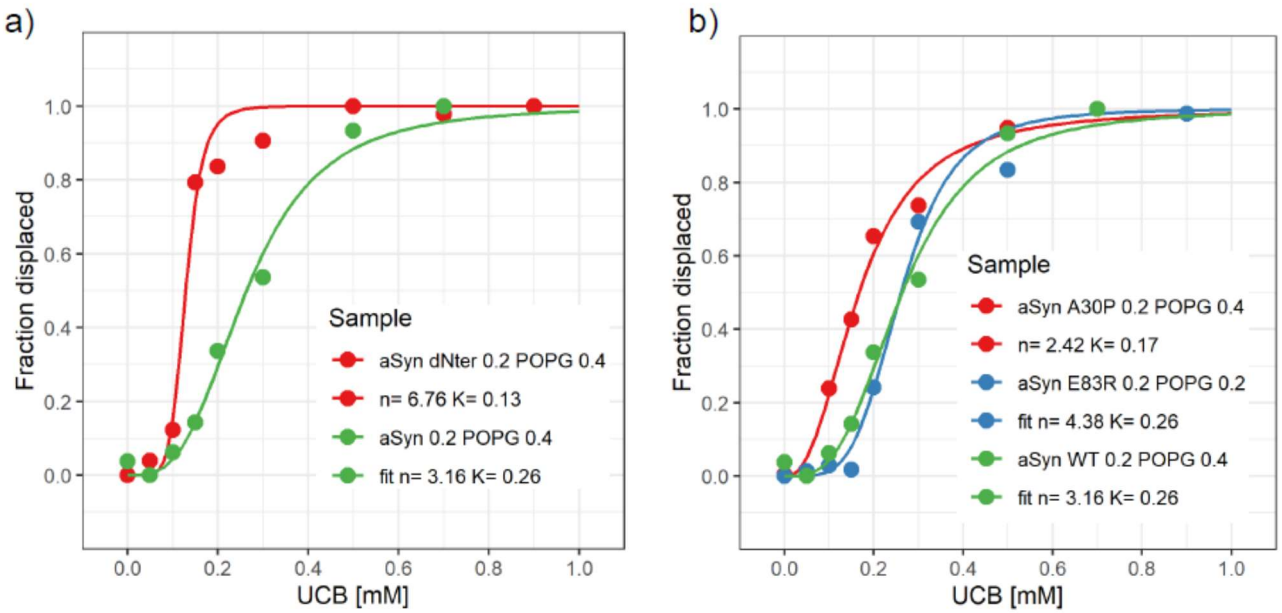
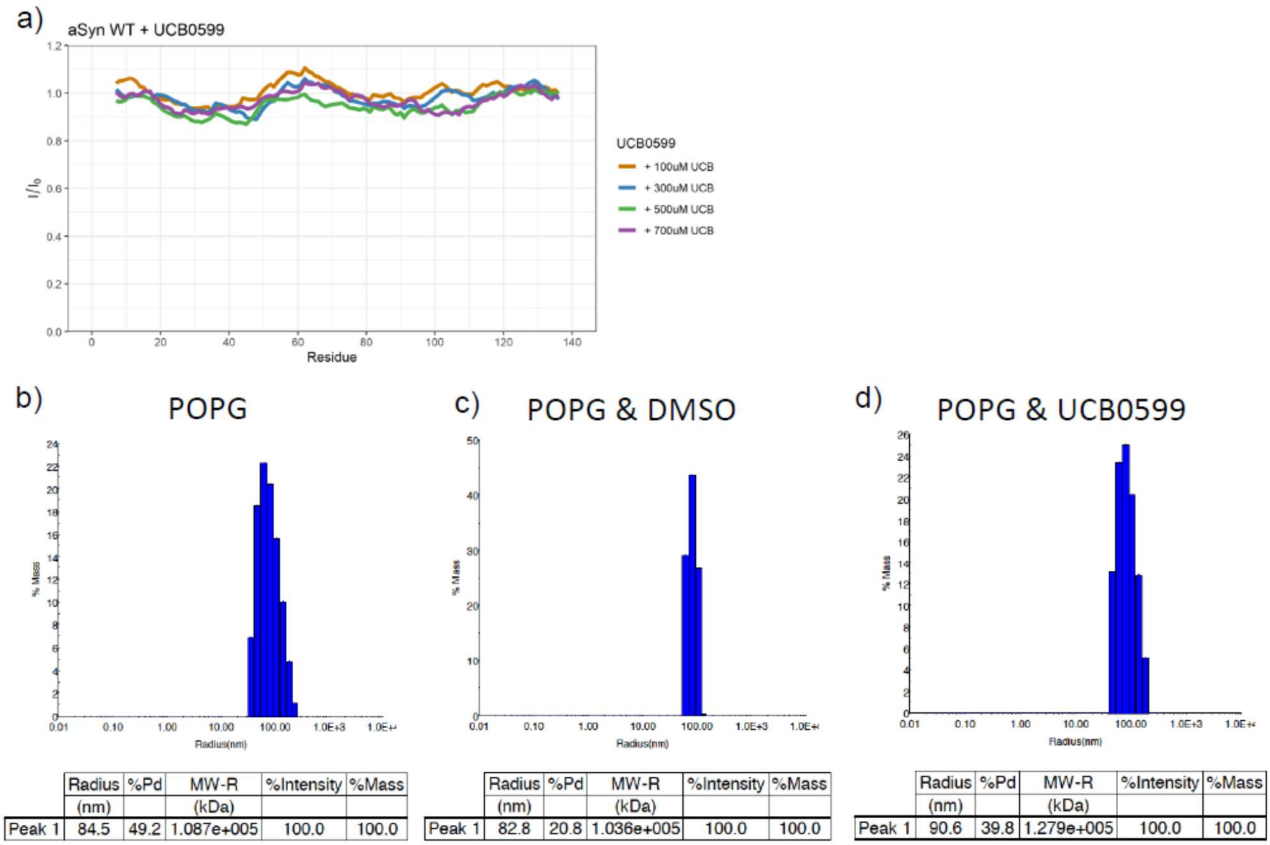


Figure S16. Comparison the effect of UCB0599 in different αS constructs. (a) Shows a comparison of wildtype αS (aSyn, green) and a construct where the first 30 residues were removed (aSyn dNter, red). Although influencing the exact point of release from the liposomal membrane, the mutant still responds to the titration with UCB0599 in a sigmoidal manner. The same is true for **(b)** where we compare the WT protein (green) to A30P (red) and E83R (blue) point mutations. In all cases a sigmoidal displacement is observed which is best explained by the compound acting on the oligomeric species of αS and does not fit with the assumption of a simple reduction of membrane affinity caused by UCB0599. Hill fits of the displacement functions are meant to guide the eyes.



	Radius (nm)	%Pd	MW-R (kDa)	%Intensity	%Mass
Peak 1	84.5	49.2	1.087e+005	100.0	100.0

	Radius (nm)	%Pd	MW-R (kDa)	%Intensity	%Mass
Peak 1	82.8	20.8	1.036e+005	100.0	100.0

	Radius (nm)	%Pd	MW-R (kDa)	%Intensity	%Mass
Peak 1	90.6	39.8	1.279e+005	100.0	100.0

Figure S17. UCB0599 does not disrupt POPG based liposomes, nor does it form micelles. **(a)** HSQC based measurements of αS signal intensities upon titration with UCB0599 do not show a strong influence of the compound. Slight variations in signal intensities are due to the addition of DMSO. **(b-d)** DLS measurements of POPG based liposomes show that the addition of the compound does not disrupt the liposomes. The size determination of liposomes by DLS has wide diameter ranges associated with it and all measurements fall within the expected range, showing that UCB0599 does not induce a measurable change.

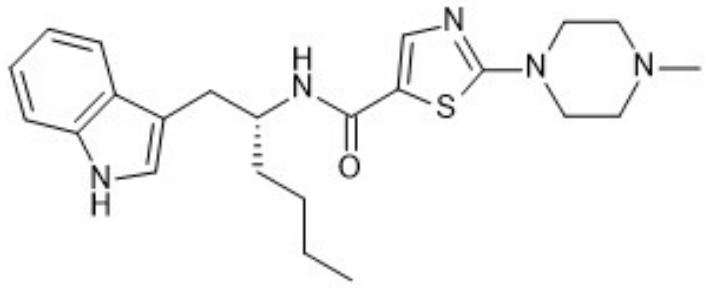
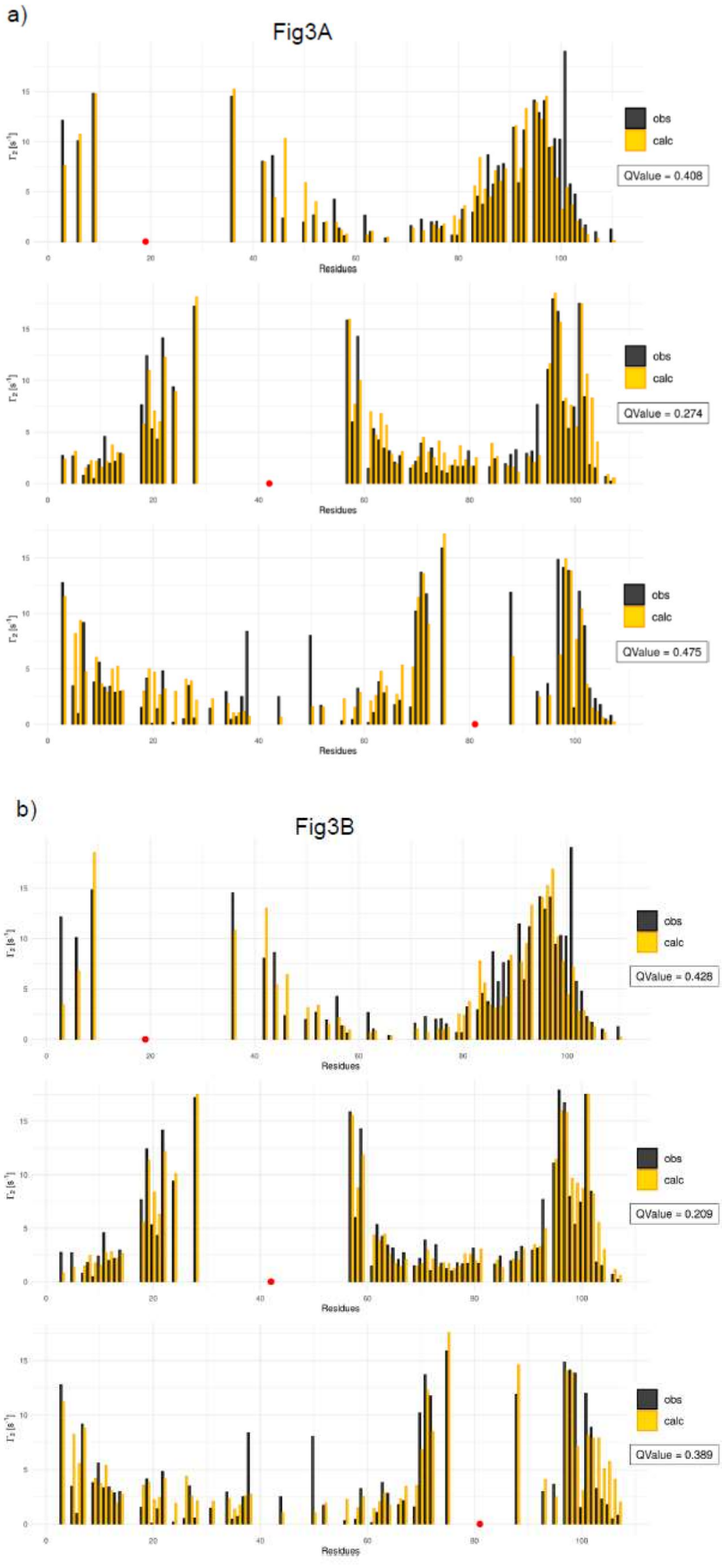


Figure S18. The structure of UCB0599. We show the structural formula of UCB0599, the compound not only used in this study but also currently in phase 2 clinical trials in Parkinson's disease patients.


a)
Fig3A
obs
calc
QValue = 0.408
QValue = 0.274
QValue = 0.475
Residues
b)
Fig3B
QValue = 0.428
QValue = 0.209
QValue = 0.389

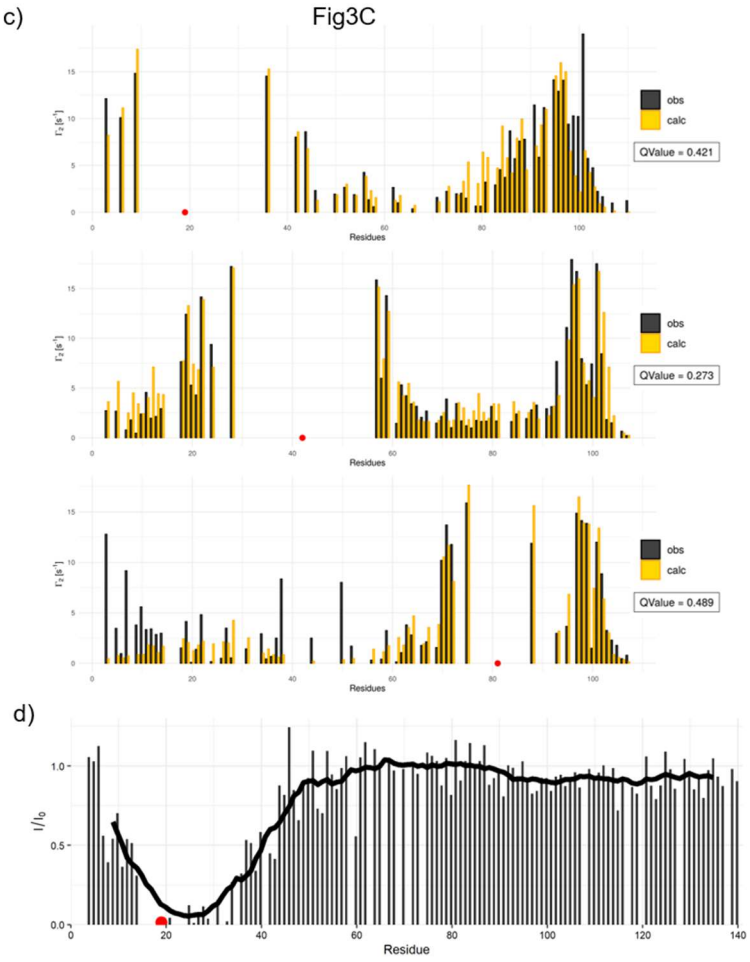


Figure S19. Experimental PRE-rates compared to those back-calculated from model structures and those observed on bicelle-bound protein. The models shown in Figure 3 were tested in terms of their capacity to fulfill the experimentally derived PRE-rates. The observed rates in the SDS-bound state are similar to intensity ratios measured in the bicelle bound state. **(a)** Corresponds to the model structure depicted in Figure 3A, while **(b)** and **(c)** correspond to structures in Figures 3B and 3C, respectively. Black bars indicate the expected rates for the model structures, while the observed rates are labelled yellow. Calculations were carried out for all datasets using the MTSL spin labels at positions 19, 42 and 87. The good obtained agreement can be appreciated when comparing the matches seen here to those in Figure S3. The average Q-Value over all 3 mutants ranges from 0.34 **(b)** through 0.38 **(c)** to 0.39 **(a)**. In panel **(d)** we show the intensity ratio between para- and diamagnetic bicelle bound 19C MTSL tagged αS. Both the range of reduction around the labelled position as well as well as the additional reduction in the range 90-100 match observations in the SDS-bound state, although reductions around 100 are smaller and extend out further then in the SDS-bound state.

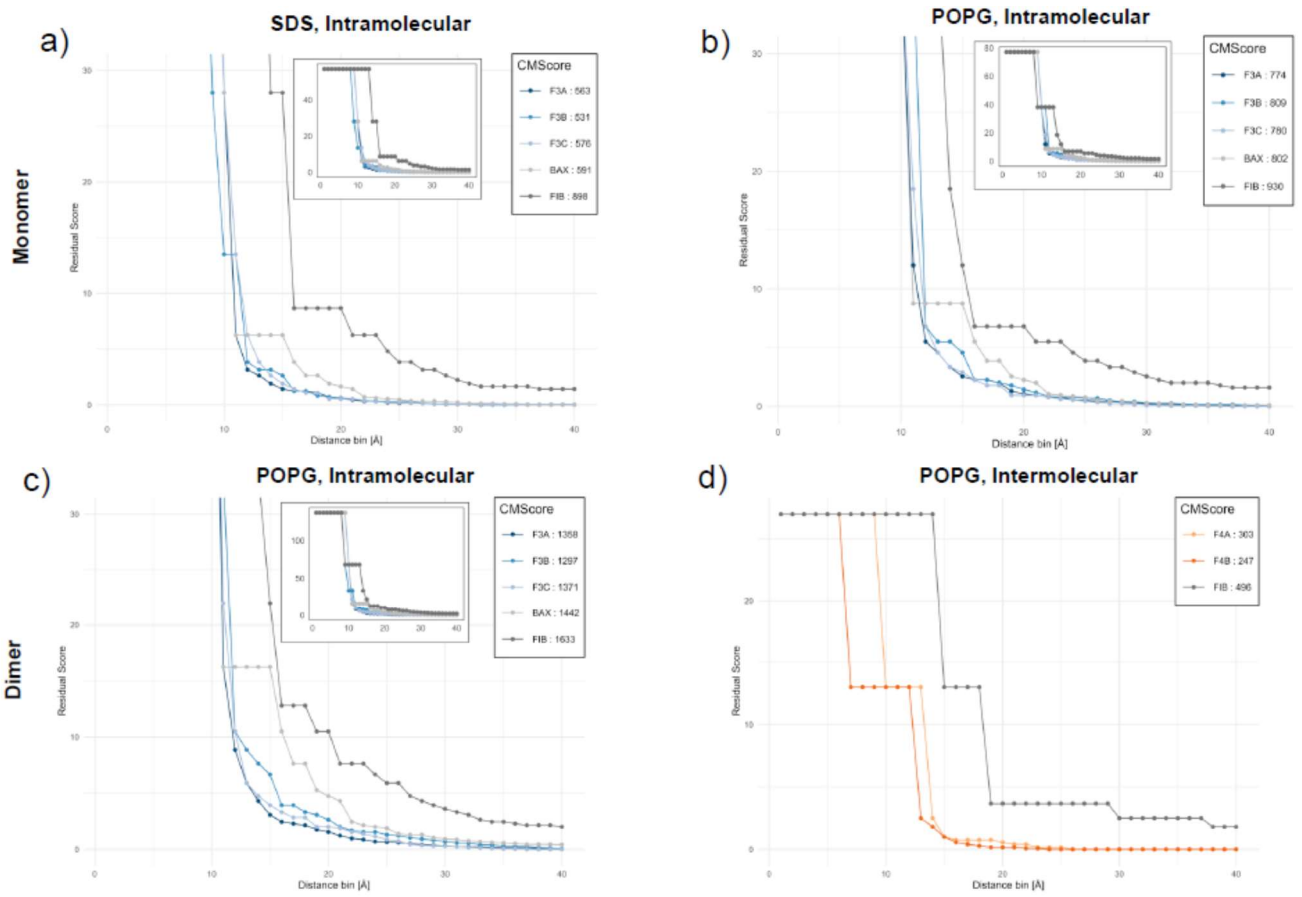


Figure S20. Quantification of the match between model structure and XL-MS data. Here we show the calculation of CMScores in order to quantify the fit between the proposed model structures and observed crosslink data. In each plot various models are matched to the same XL-MS dataset. As the curves indicate how many crosslinks cannot be fulfilled at a given CA-CA distance (y-Axis), lower values indicate a better fit. Insets show the full curve up to the maximum crosslinks, while the main plot is zoomed to highlight the differences between individual structures. CMScores indicate the area under each curve and are given in the top right panels. Structures compared to intramolecular links stem from Figure 3a,b,c (F3A, F3B, F3C) the published SDS-bound structure PDB-ID: 1xq8 (BAX), and the published fibrillar structure PDB-ID: 6xyo (FIB). Structures compared to intermolecular links stem from Figure 4a,b and the fibrillar structure PDB-ID: 6xyo (FIB). **(a)** Shows how well each structure matches to intramolecular crosslinks detected in SDS-micelle bound αS. While all our structures are matching the observed links slightly better than the published SDS-bound one, the clearest difference is seen in comparison to the fibrillar structure. In **(b)** we see again, that our structures (although not utilising and XLMS data) fit the observed crosslinks better than available structures and much better than the fibrillar structure used as a decoy. **(c)** Shows a good match

of our structurs with Figure 3b matching best, while the difference to the published SDS-bound structure (BAX) grew. As mentioned in the main text, this is due to the higher share of compact states stemming from oligomeric αS, present int the crosslinked sample (the much higher CMScores are due to a larger amount of crosslinks detected and do not indicate a decreased fit). **(d)** Shows matching of our structures to intermolecular crosslink data, and demonstrates a drastic improvement (half the CMScore) especially for the structure shown in Figure 4b, as compared to the fibrillar structure.

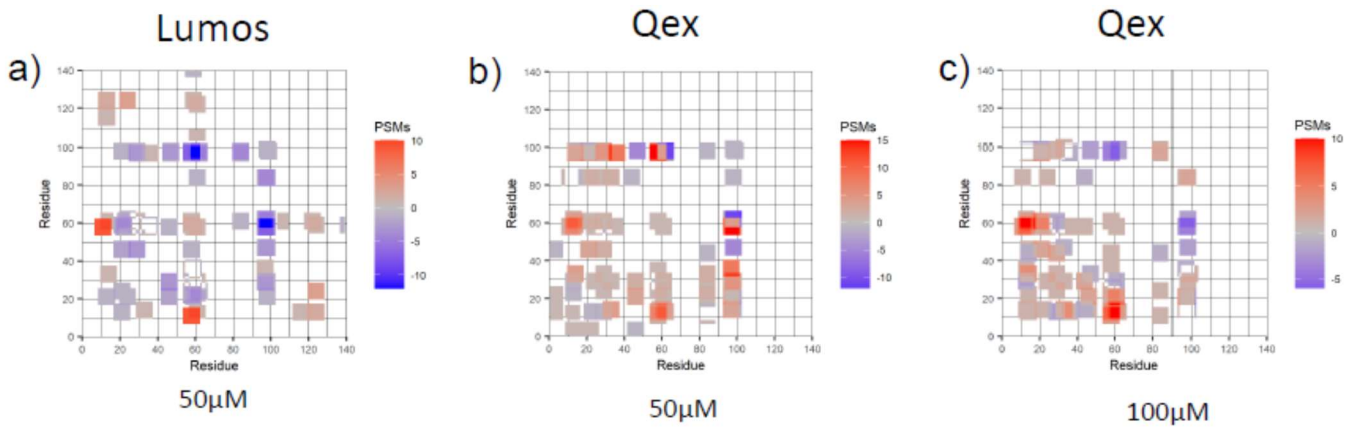


Figure S21. Comparison of UCB0599 induced differences in PSM maps generated with differing instruments. Although the measurements depicted here stem from different instruments, include a variation in the amount of ligand added and were carried out with separate samples, the influence of UCB0599 is clearly detected in similar regions of the protein. In addition, the direction of change is also maintained between different samples and machines. **(a)** shows the measurement as depicted in Figure 6d. This measurement was carried out on a Orbitrap Fusion Lumos mass spectrometer and depicts the difference induced by the addition 50µM UCB0599. **(b)** and **(c)** were recorded on a Q Exactive HF-X mass spectrometer and show the induced difference by addition of 50µM **(b)** and 100µM **(c)**. The reproducibility observed between different samples and machines demonstrates the feasibility of the approach used. The colour scale indicates how many links were found in a four-residue window.

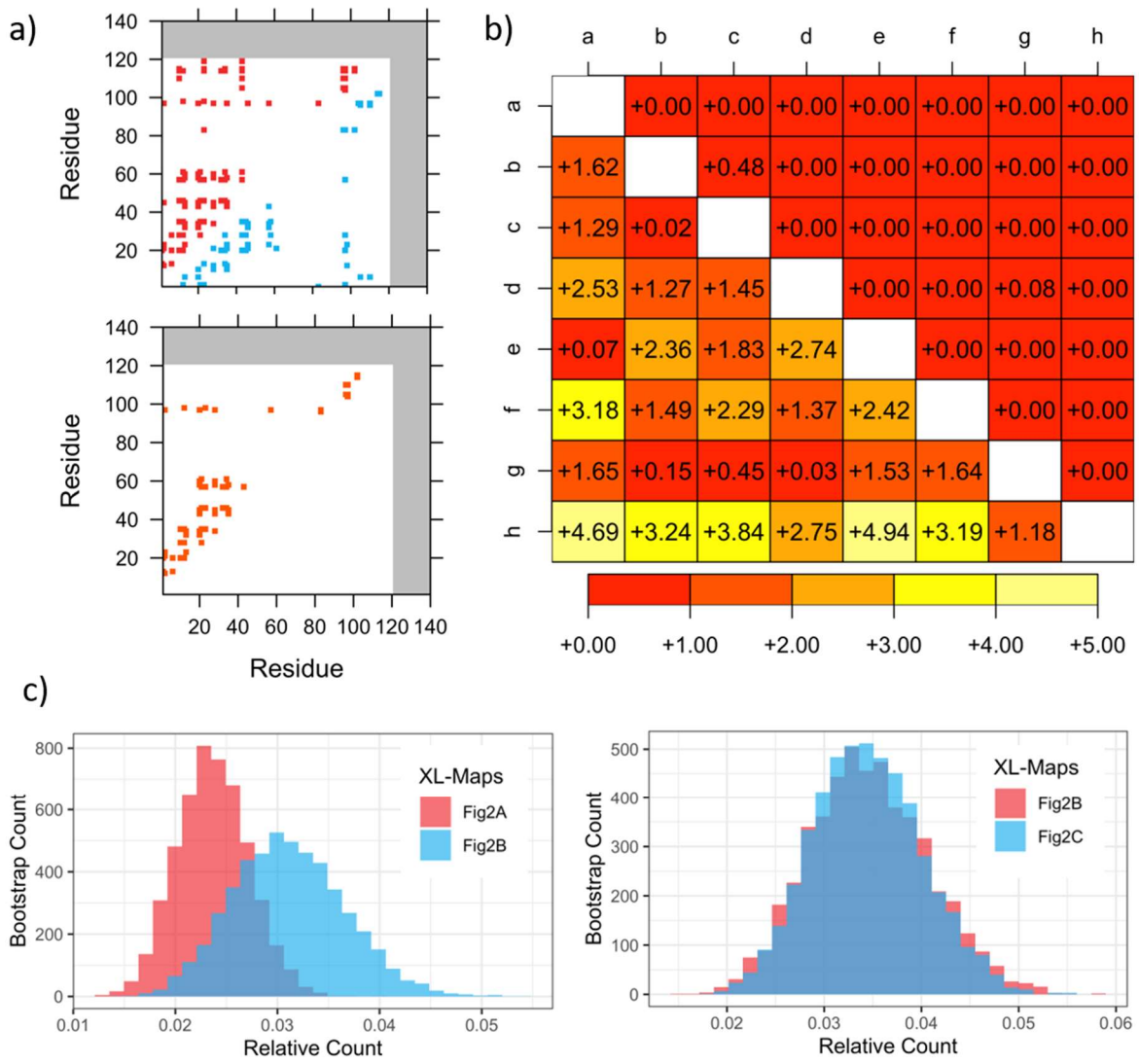


Figure S22. Statistical comparison of Crosslink-maps shown in Figure 2. In **(a)** XL-map regions averaged with a windows size of 3 are shown. Residues for Fig.2a are given in red while those in Fig.2b are shown in blue on either side of the diagonal (top). The set of overlapping regions for both is shown in orange in the bottom graph. **(b)** Shows a pairwise comparisons of bootstrapped relative counts among corresponding regions for the maps in figure 2. While values above the diagonal show p-values, those below it show Cohen's d. In **(c)** we show the example for the bootstrapped distributions of the counts in corresponding regions of maps for Fig.2a and Fig.2b (left) as well as for Fig.2b and Fig.2c (right), visually demonstrating how the values were obtained.

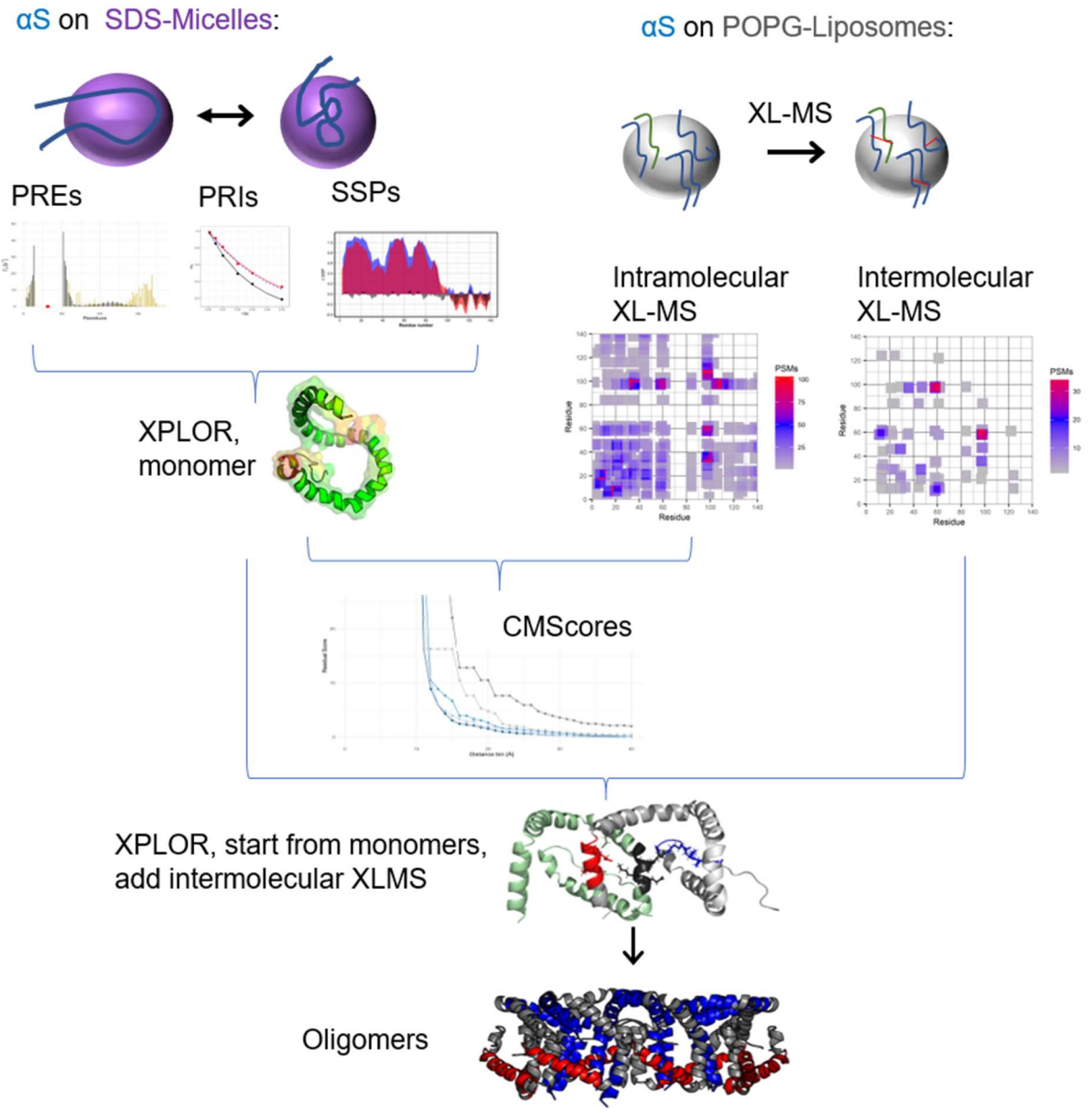


Figure S23. Graphical summary of model generation workflow. This simplified summary shows our procedure for model generation. We use NMR-derived PRE-, PRI- and SSP-data from SDS-bound αS (top left) to generate models of compact membrane bound αS. These are checked against XL-MS derived data obtained from liposome bound αS (top right). First, intramolecular crosslinks were used as a check for comparability of the fold (CMScores, center). As the obtained models fit the XL-MS data well, we use intermolecular crosslinks for the generation of dimeric models, which then are used to generate oligomeric models (bottom).

6

CHARACTERIZATION OF METHYL-π INTERACTIONS THROUGH CHEMICAL SHIFTS

Reprinted with permission from
A. Beier, G. Platzer, T. Ho¨furthner, *et al.*, "Probing protein–ligand methyl- π interaction geometries through chemical shift measurements of selectively labeled methyl groups," *Journal of Medicinal Chemistry*, vol. 67, no. 15, pp. 13 187–13 196, 2024 [79]

DOI: 10.1021/acs.jmedchem.4c01128

PERSONAL CONTRIBUTION

Protein production, conceptualization, data analysis, validation, writing

Journal of Medicinal Chemistry

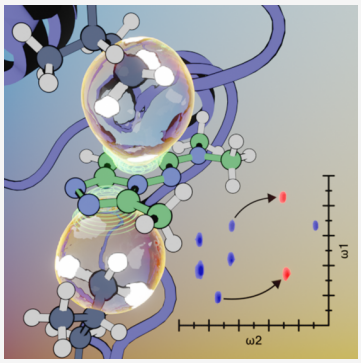
pubs.acs.org/jmc
Article

Probing Protein−Ligand Methyl−π Interaction Geometries through Chemical Shift Measurements of Selectively Labeled Methyl Groups

Andreas Beier,* Gerald Platzer, Theresa Höfurthner, Aleksandra L. Ptaszek, Roman J. Lichtenecker, Leonhard Geist, Julian E. Fuchs, Darryl B. McConnell, Moriz Mayer, and Robert Konrat

Cite This: J. Med. Chem. 2024, 67, 13187−13196

ABSTRACT: Fragment-based drug design is heavily dependent on the optimization of initial low-affinity binders. Herein we introduce an approach that uses selective labeling of methyl groups in leucine and isoleucine side chains to directly probe methyl−π contacts, one of the most prominent forms of interaction between proteins and small molecules. Using simple NMR chemical shift perturbation experiments with selected BRD4-BD1 binders, we find good agreement with a commonly used model of the ring-current effect as well as the overall interaction geometries extracted from the Protein Data Bank. By combining both interaction geometries and chemical shift calculations as fit quality criteria, we can position dummy aromatic rings into an AlphaFold model of the protein of interest. The proposed method can therefore provide medicinal chemists with important information about binding geometries of small molecules in fast and iterative matter, even in the absence of high-resolution experimental structures.



INTRODUCTION

Modern drug discovery campaigns often start with the screening of small molecule libraries in order to identify an initial starting point for the subsequent optimization of binding affinities to a target of interest. Fragment-based drug discovery (FBDD), in which simple molecules are used in the screening process, has become especially popular as they offer significant advantages.[1] For one, fragments cover the chemical space more efficiently than larger compounds, especially when care is taken to ensure chemical diversity of the used library.[2] While this also means that fragments generally show fewer interactions and lower affinities, these interactions are often fundamental for target recognition and could otherwise be masked by unfavorable contacts in larger lead compounds.[3] While interactions of polar groups like the formation of strong hydrogen bonds and electrostatic interactions dominate energetically favorable binding events for fragments, the majority of small molecule ligands also show other directional interactions like arene contacts or hydrogen bonds involving CH-groups, and their influence tends to increase with the size of the ligand.[4] The most prominent of these is the formation of weak hydrogen bonds between aromatic π-systems, commonly of a small molecule, as the acceptor and aromatic or aliphatic CH-groups as the donor.[5,6]

Nuclear magnetic resonance (NMR) has been an integral method for FBDD from the very start.[7] Simple protein detected 2D NMR experiments can identify low-affinity binders and readily distinguish them from false positive hits.[8] NMR is especially well suited to detecting the CH−π interactions mentioned above, as the additional local magnetic field induced by aromatic ring currents leads to a large anisotropic chemical shift change in surrounding nuclei that can be readily measured by standard NMR experiments. This serves not only as a reporter of CH−π interactions, but as we showed in a previous work, the magnitude of the measured shift change also correlates well with the binding energy of the respective interaction.[9] By this, NMR provides indirect information about the binding contribution of individual interactions as well as direct structural information at atomic resolution that aids medicinal chemists in their efforts to refine initial fragment hits into high-affinity binders. The potential impact is further emphasized by the relative frequency of aliphatic and aromatic amino acids in binding pockets as well as the abundance of aromatic ring systems in small molecule ligands.[6,10] NMR does not come without its own set of limitations, namely spectral overlap and limited experimental sensitivity.[10,11] To overcome these challenges, selective labeling strategies for specific residues and side chain atoms have been established which reduce spectral crowding and allow the use of optimized measurement schemes to utilize the favorable relaxation characteristics of certain chemical groups.[12,13] In combination, this has opened up NMR-based

Received: May 13, 2024
Revised: July 11, 2024
Accepted: July 15, 2024
Published: July 29, 2024

https://doi.org/10.1021/acs.jmedchem.4c01128
J. Med. Chem. 2024, 67, 13187−13196

Journal of Medicinal Chemistry pubs.acs.org/jmc Article

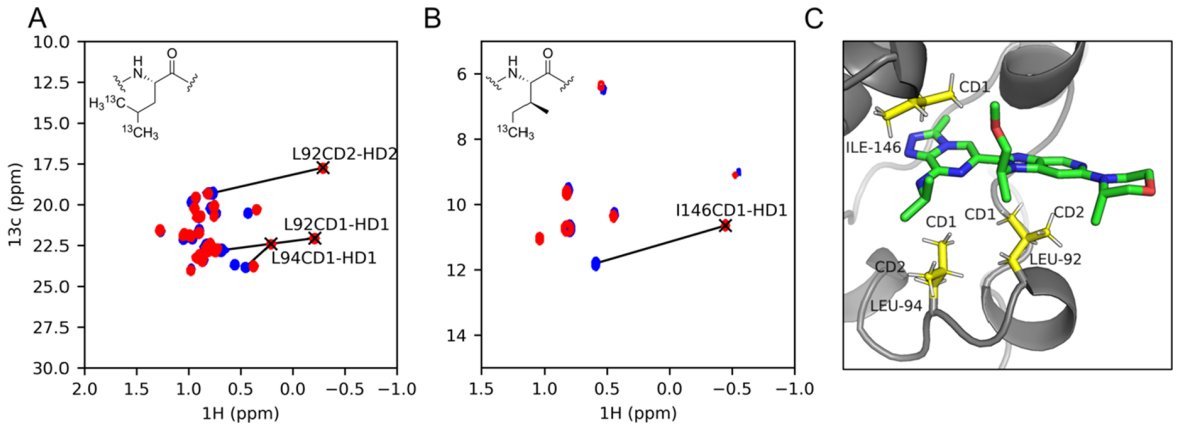


Figure 1. Overlay of $^{1}H-^{13}C$ HSQC spectra of selectively leucine (A) and isoleucine (B) labeled apo BRD4-BD1 (blue) and in complex with ligand 1 (red). (C) Position of ligand 1 and surrounding leucine and isoleucine residues (PDB: 6XUZ).

FBDD approaches to high molecular weight proteins and complexes.[14]

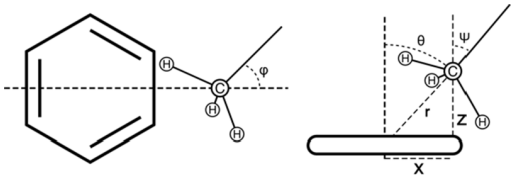
In this work, we aim to extend the work we have previously conducted on tryptophane CH-groups toward the more accessible methyl groups of aliphatic amino acids, here namely, leucine and isoleucine. These are highly abundant in the hydrophobic cores of proteins as well as in the binding pockets for small molecules which are commonly found in less exposed sites of the protein which makes them valuable reporters of ligand binding events.[4] The ease of interpretation of NMR experiments using selectively labeled side chains and the exquisite sensitivity of the chemical shift for changes in methyl−π interaction geometry allows for direct probing of these interactions.

To demonstrate the method, we use bromodomain 1 of bromodomain-containing protein 4 (BRD4-BD1) as our protein target. BRD4-BD1 acts as a chromatin reading protein by binding to acetylated histones and is therefore crucially involved in the regulation of protein expression.[15] The highly conserved substrate binding site of BRD4-BD1 selectively recognizes acetylated Lysines of H4 histones and multiple potent inhibitors of this interaction have been developed in the past.[16,17] Additionally, the binding pocket contains multiple leucine and isoleucine residues. Therefore, it acts as an ideal model system for our approach.

RESULTS AND DISCUSSION

Selective Labeling. Selective labeling of the δ methyl groups of leucine and isoleucine residues was achieved by the addition of selectively labeled alpha-keto-acids as metabolic precursors to the *Escherichia coli* expression medium. In detail, 2-ketoisocaproate and 2-ketobutyrate were used to label leucine and isoleucine residues, respectively.[18,19] In both cases, the δ position methyl carbons were selectively $^{13}C-^{1}H$ labeled and all other nonexchangeable protons in the side chain were deuterated to eliminate through bond coupling effects. In addition to reducing spectral crowding, the use of these precursors also improves signal-to-noise ratio by preventing intraresidue dipole−dipole-induced spin relaxation. We investigated a total of 16 ligands, 10 of which have already been described in our previous work on CH−π interactions[9] (SI, Table 1).[9] In total, BRD4-BD1 contains 13 leucine and 7 isoleucine residues. Of these, 2 leucines at positions 92 and 94 as well as one isoleucine at position 146 are present in the binding pocket and consequently experience large shift changes upon binding (Figure 1).

Geometric Model of Methyl−π Interactions. To compare the measured chemical shift changes upon ligand binding to the ones expected, as calculated from available X-ray structures of the complexes, we apply the commonly used dipole model for chemical shifts induced by aromatic systems first introduced by Pople (Figure 2).[20]



$$\Delta\sigma = 10^{6} * \frac{ne^{2}a^{2}}{4\pi mc^{2}} * \frac{3cos^{2}\theta - 1}{r^{3}}$$

Figure 2. Top: Geometric model that unambiguously describes the position and orientation of a methyl vector relative to an aromatic ring system. Bottom: Pople point-dipole model was used to calculate the chemical shift change induced by a benzene ring.

Δσ refers to the change in the isotropic nuclear shielding constant in parts per million (ppm). n is the number of circulating electrons and is fixed at $n = 6$ in our case. e is the elementary charge in Franklin, m is the mass of an electron in gram, and a is the diameter of the aromatic ring in cm. The diameter was chosen based on the presence of a 5- or 6-ring and multiring systems were separated into their set of smallest rings (SI, Table 2).

r is the distance between the nucleus in question and the aromatic ring center in cm. θ is the angle between the vector connecting the nucleus under investigation with the ring center and the ring normal pointing in the same general direction. Both parameters were extracted from X-ray cocrystal structures. While this is straightforward for methyl carbons, methyl hydrogens are generally not present in X-ray structures. In addition, due to the fast rotation of the hydrogens around a common C−C symmetry axis, they are chemically equivalent, and their signals collapse to a single average over the three atoms. To address this, a more sophisticated geometric model must be used (Figures 2 and S1). In this model, we treat the

https://doi.org/10.1021/acs.jmedchem.4c01128
J. Med. Chem. 2024, 67, 13187−13196

Journal of Medicinal Chemistry pubs.acs.org/jmc Article

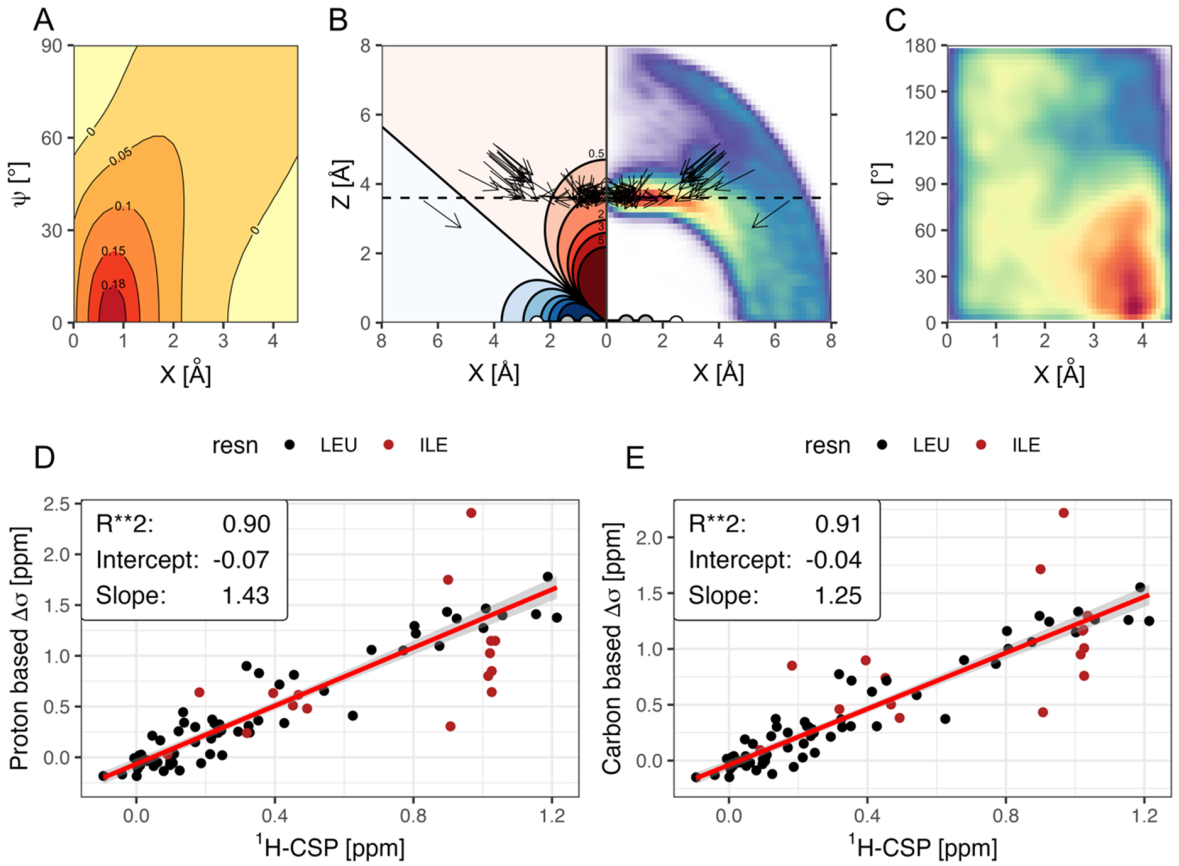


Figure 3. (A) Difference between the proton- and carbon-based models dependent on ψ angle and X-displacement along the ring plane at a constant normal distance (Z) for the methyl carbon of 3.8 Å and averaged over all φ angles. (B) Projection of the extracted methyl vectors (C−C symmetry axes) for all ligands under investigation above the aromatic ring plane. The colored areas indicate the expected Pople values (left) and the relative frequency of Leu methyl carbons in the PDB at the same position (right). (C) Frequency of φ values dependent on X-displacement along the ring plane. (D/E) Correlation between measured and calculated chemical shift changes for the leucine (black) and isoleucine (red) methyl groups. At this stage, the linear fit is based only on leucine shifts. Shifts are calculated either based on an average over ideal positions for the methyl protons (D) or the proxy position of the methyl carbons (E).

methyl carbon not as a point in space but as the origin of a unit vector along the C−C symmetry axis of the methyl hydrogens (methyl vector). We get two such vectors for leucine ($C\delta_1/C\delta_2$ to $C\gamma$) and two in isoleucine ($C\delta_1$ to $C\gamma_1$ and $C\gamma_2$ to $C\beta$). This leads to a new set of four parameters r, θ, φ, and ψ that uniquely describe the vector's position and orientation relative to the center and normal of an aromatic ring plane.

The parameters r and θ keep their previous meaning. In addition, we extracted parameters φ and ψ. In this case, φ is the angle formed between the respective projections of the methyl vector and the vector connecting the methyl carbon with the ring center onto the aromatic plane, and ψ is the angle between the methyl vector and the ring normal. By treating the aromatic ring as a simple plane, it holds no directional information, except for the ring normal. Therefore, the domains of both φ and ψ are limited to $[0, \pi]$. To reconstruct the expected position of the methyl hydrogens from this set of parameters, one can simply scale the methyl vector to the length of a CH−sp^3 bond (1.087 Å) and rotate it around the vector resulting from the cross product of the methyl vector and the ring normal, centered at the methyl-carbon position, by 109.5° toward the aromatic ring plane. This vector can then be rotated in discrete steps around the original methyl vector. For each step, we extract new r and θ for that respective hydrogen position. By averaging the chemical shifts calculated via the Pople model for each step, we get the approximate chemical shift expected for the methyl hydrogens when rotated around their symmetry axis (Figure S6). When comparing the calculated proton chemical shift changes based on this proton model with the values we get when simply using the carbon position as a proxy (carbon-based model), the biggest differences arise close to the ring center at low ψ angles (Figure 3A). While methyl-carbon positions seem to be a good proxy for the averaged methyl proton center, carbon shifts themselves are not discussed in this paper as they are heavily influenced by additional factors like side chain rotamers among others that dominate over the ring-current effect.[21]

The geometric parameters extracted from the protein−ligand X-ray structures are visualized in Figure 3B. The black arrows represent the position of the methyl vectors above the ligand rings, with the arrow tips centered on the methyl-carbon position. In this projection of the 3-dimensional geometry, the angle ψ is indicated by the pitch of the vector. Increasing yaw angles φ in turn are represented by shortened arrows as the methyl vector is rotated out of the projection plane. The methyl carbons are predominantly positioned above the ring plane at a normal distance of 3.6 Å (Figure 3B, dotted line) and a slight displacement from the ring center.[22] This is also the area in which one can expect a chemical shift change of around 1 ppm for the carbon atoms while the chemical shift

https://doi.org/10.1021/acs.jmedchem.4c01128
J. Med. Chem. 2024, 67, 13187−13196

Journal of Medicinal Chemistry pubs.acs.org/jmc Article

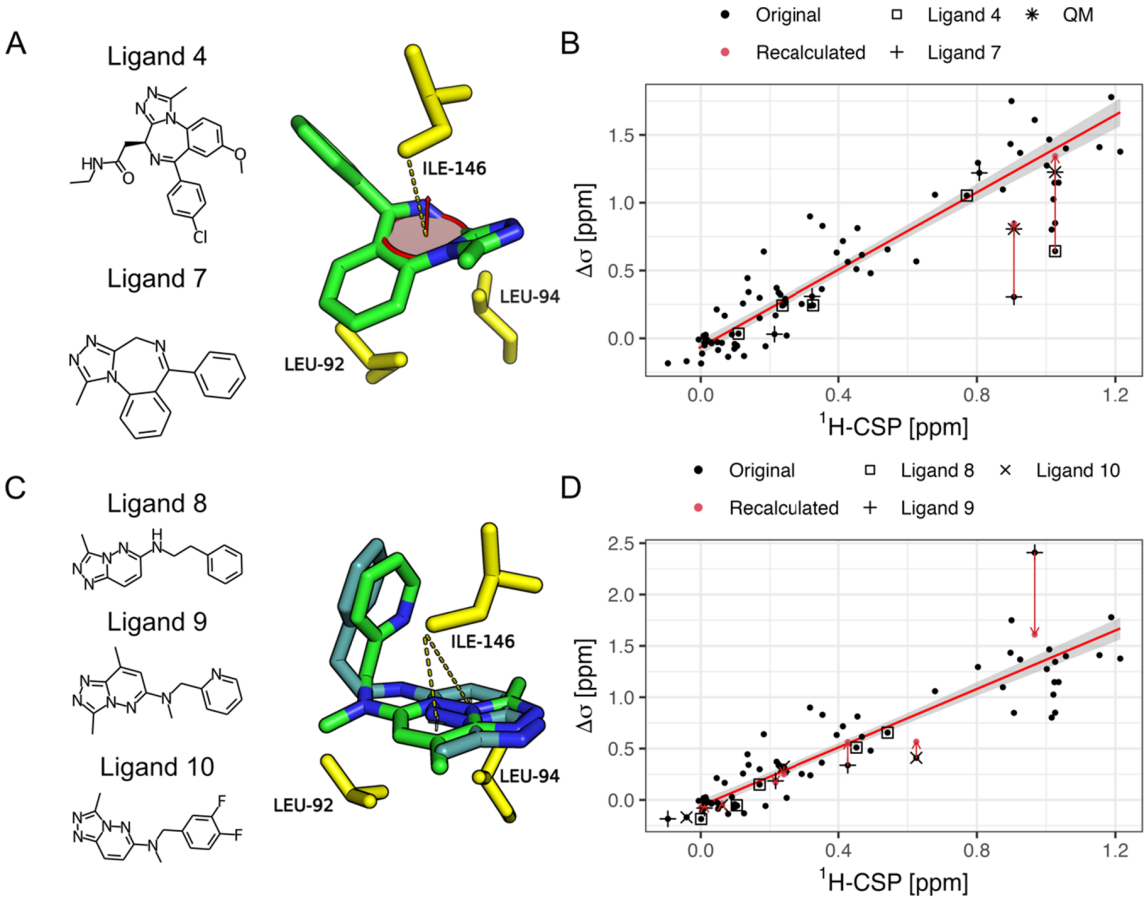


Figure 4. (A) X-ray structure of ligand 7. The central nonplanar diazepine ring found in ligands 4 and 7 can be converted to a consensus plane over all ring atoms in order extract the relevant parameters for the Pople model. 2D structures of ligands 4 and 7 are depicted in ascending order. (B) Correlation of calculated isoleucine shifts for ligands 4 and 7 is improved by the homoaromatic ring approximation (red arrows). QM-based shift calculations predict similar values after being scaled by the slope of the linear fit over all of the leucines. (C) Overlay of ligands 8 and 9. The aromatic polycycle shared among ligands 8 and 9 is rotated in ligand 8 around one of the ring centers which places the Ile-Cδ_1 directly above the other ring center. 2D structures of ligands 8, 9, and 10 are depicted in ascending order. (D) Calculated shifts for ligand 9 reflect measured data (red arrows) when they are calculated from an average over the ring position for ligands 8 and 9.

change for the methyl protons will be averaged by their rapid rotation around the methyl axis, and plotting their expected shift becomes dependent on both φ and ψ (Figure 3B left). The same preferred position for methyl-carbon atoms above aromatic rings can also be extracted from the Protein Data Bank[23] (PDB) as indicated by the heatmap in Figure 3B (right). It is apparent that while the pitch angle ψ takes up values around 60−70°, the yaw angle φ is either unconstrained above the center of the ring or very low beyond the ring diameter in order to point the methyl toward the ring center (Figures 3C and S4).

Measured vs Calculated CSPs. We measure the experimental chemical shift change for BRD4-BD1 induced by the addition of a ligand from a set of 16 known binders (Table S1) and compare them to the expected shift change calculated from protein−ligand cocrystal structures. Upon ligand binding, we can detect large shifts for the δ-methyl resonances of BRD4-BD1 Leu 92, 94, and Ile 146 (Figure 1). The correlation between measured chemical shift changes induced by ligand binding versus the expected values calculated using the Pople model is shown in Figure 3D,E. For Leu-δ methyl groups, both the carbon proxy model (Figure 3E) as well as the average over all possible proton rotamers (Figure 3D) show similar R^2-values of 0.90 and 0.91, respectively. The main difference between the two models is the deviation from the expected slope of 1, which is more pronounced for the proton model. The reason for this deviation is the lack of proper parametrization of the Pople model depending on the nature of the aromatic ring.[24,25] As the Pople model is originally parametrized for a benzene molecule, a ring-current intensity factor is commonly fitted as a scaling factor that relates the expected intensity in benzene to the one in the ring under investigation.[24] The carbon proxy model seems to be a good approximation for small changes in binding geometries that are averaged in solution which are not reflected in a static X-ray structure, compared to the proton model, as it is less sensitive to small changes in φ and ψ angles (Figure 3A) that are sampled in solution.

Outlier Correction. The biggest discrepancy between measured and calculated values is present for the signals measured for Ile146 in BRD4-BD1. For ligands 4 and 7, the Ile-δ_1 methyl group under investigation is positioned in the vicinity of a nonaromatic diazepine ring. In contrast, the leucine methyl groups are positioned above aromatic 5- and 6-rings and consequently fit the data. The 7-membered diazepine ring may experience through space conjugation of its p-orbitals leading to a weak homoaromatic ring-current effect, that in turn affects the methyl proton shifts of Ile146 upon ligand

https://doi.org/10.1021/acs.jmedchem.4c01128
J. Med. Chem. 2024, 67, 13187−13196

Journal of Medicinal Chemistry pubs.acs.org/jmc Article

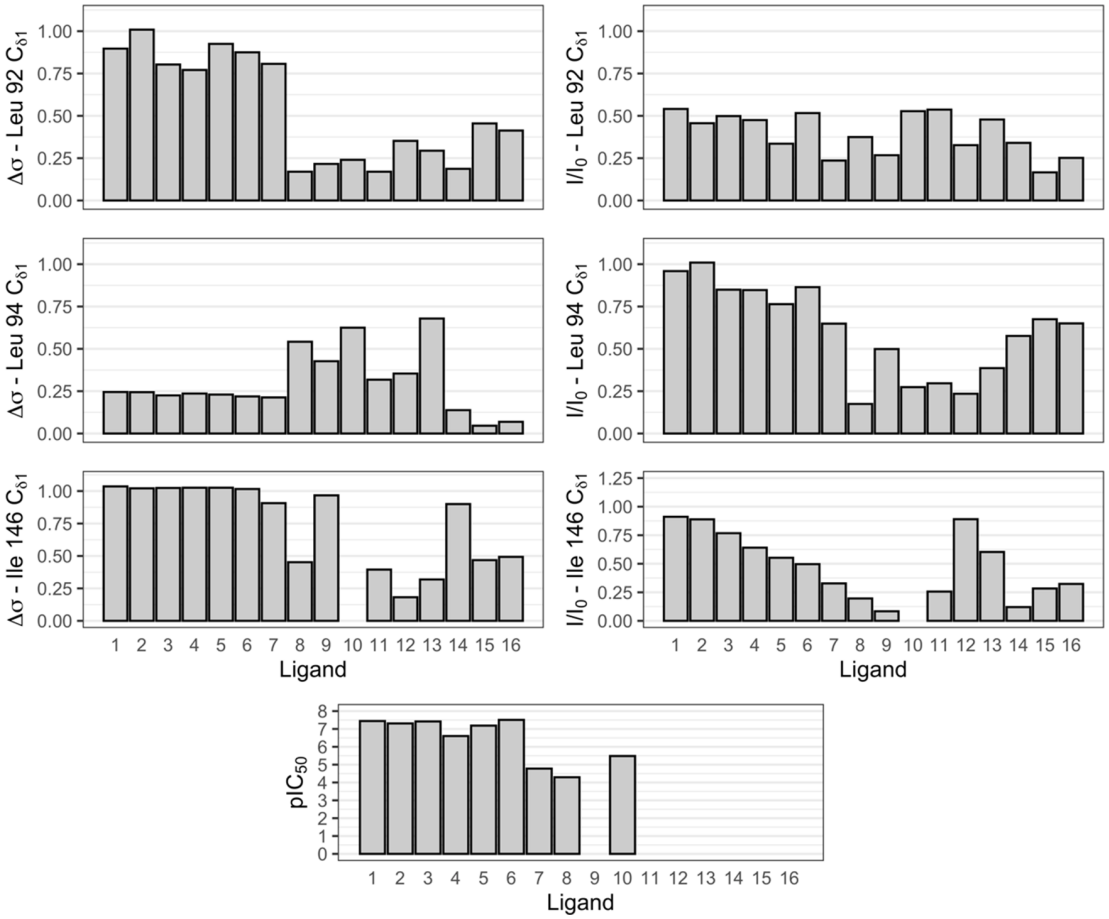


Figure 5. Chemical shift changes and intensity ratios between free and bound state for Cδ1 methyl groups as well as AlphaScreen[27] pIC50 values. For ligands 9 as well as 11−16, no affinity information is available as they were fragment hits that yielded X-ray co-structures but for which no K_D values were determined. Intensity ratios for leucine 92 Cδ1 are reduced due to partial overlap of the apo state (Figure S3).

binding.[25,26] The Pople model is not able to deal with systems like this and, therefore, underestimates the chemical shift change. As an approximation of the homoaromatic effect, a consensus plane was created based on the diazepine-ring atoms (Figure 4A). When the geometric parameters for the Pople model are extracted based on this plane, the agreement between experimental and calculated shifts for Ile146-δ1 improves drastically (Figure 4B). To test the validity of this approximation, we performed quantum mechanics (QM) based calculations of the bound and unbound shifts for ligands 4 and 7. The resulting theoretical CSPs can reproduce the experimental values and, after scaling by the slope of the linear fit, align well with our diazepine-ring pople approximation. (Figure 4B).

For ligand 9 in turn, the calculated shift for the Ile-δ_1 methyl group is much larger than the experimental value. Upon closer inspection, ligand 9 shares a central substructure with ligands 8 and 10. In the corresponding X-ray structures, it is apparent that the 5,6-ring substructure for ligand 8 is rotated around the center of the 5-ring compared to ligands 9 and 10. This leads to a major change of the methyl−π geometry between the Ile-δ_1 and the 6-ring center (Figure 4C). A plausible explanation for the overestimation of the calculated shift in this case is that the rings of ligands 9 and 10 also sample the corresponding positions occupied by ligand 8 and exchange between both states on a medium to fast time scale. Further support comes from the fact that these 3 ligands are also the ones showing the largest signal attenuation for the Ile-δ_1 signal upon ligand binding, in the case of ligand 10 even beyond detection (Figure 5). For all 3 of them, crystallographic B-factors for the Cδ_1 positions as well as the average over the respective ring atoms are increased compared to, e.g., ligand 1 (Table S2). This further suggests conformational exchange between both states. If the shift changes for ligand 9 are recalculated from an average of the ring position for ligands 8 and 9, then both the isoleucine and leucine shifts improve significantly (Figure 4D).

Shift Patterns and Relative Intensity. The intensity ratio of the respective side chains indicates additional flexibility of the methyl groups, alternative binding modes of the ligand, or exchange between free and bound states, all of which indicate an opportunity for Medicinal Chemists to increase binding affinity further, by pointing to the individual molecular interaction that is not optimal yet. Here, especially Ile146 is of interest as we observe a wide range of intensities from completely exchange broadened to no decrease compared with the apoprotein. Inspection of the intensity ratios in our set indicates that a stable binding mode was achieved over the course of compound optimization for high-affinity compounds.

In contrast to Ile146, for Leu94 the change in signal intensity in most cases seems to be dependent on the presence of an additional fused ring system in the ligand, which leads to less exchange broadening of Leu94. This suggests that the

https://doi.org/10.1021/acs.jmedchem.4c01128
J. Med. Chem. 2024, 67, 13187−13196

Journal of Medicinal Chemistry pubs.acs.org/jmc Article

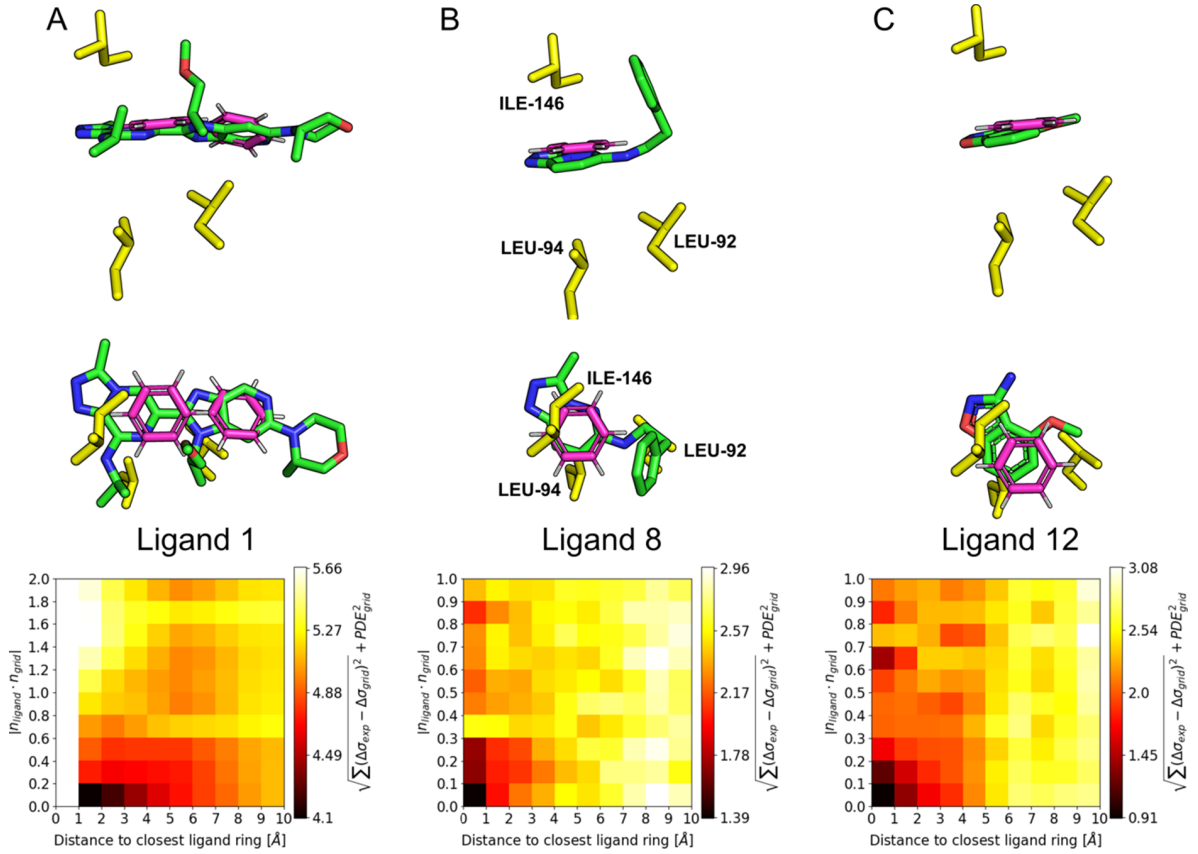


Figure 6. Examples for ring location and orientation fit based on methyl shifts and interaction geometry probabilities. Side and top views of the bound ligand structures overlaid with the best ring fits in magenta (top) and average fit value dependent on closest ligand ring position and orientation (bottom). Each cell of the heatmap represents the average fit value at the respective difference in position and orientation (absolute cosine value of the angle) of the closest aromatic ring of the ligand crystal structure. For all examples, we find the lowest fit values close to those of the true conformation of the aromatic rings. (A) Two aromatic rings must be fit simultaneously in order to cover both fused ring systems of ligand 1 (PDB 6XUZ). Due to the grid spacing, the combined distance to the closest ring is larger than 1 and the heatmap cells below 1 Å are therefore empty. (B, C) For ligands 8 (PDB entry 9FWX) and 12 (PDB entry 9FXP), a single aromatic ring is sufficient to find the correct position and orientation.

broadening experienced by Ile146 and Leu94 is not due to the global dynamics of the protein but rather the residual mobility of the ligands in the binding site. While the relative signal for Leu92 seems to be reduced for all ligands, the true effect is hard to gauge as there is peak overlap for the apo signal (Figure S3). While the magnitude of the chemical shift changes upon binding does not directly report on the affinity of the overall interaction, it does indicate which methyl groups are addressed by the aromatic ring systems. This is especially clear when comparing the ligands with either 1 (8−16) or 2 (1, 2, 3, 5, 6) fused ring systems in the set. In the former case, mostly Leu94 and Ile146 are addressed, while the addition of an additional ring system suggests a reduced interaction with Leu94 in favor of both Leu92 and Ile146.

Ring Fit Based on Shift and the PDB. In the early 2000s, McCoy et al. introduced the concept of *j*-surfaces in which they used backbone NH shifts induced by aromatic rings to locate the origin of the perturbation and therefore the most probable position of aromatic ligand rings.[28] As the methyl groups in our examples are directly interacting with the aromatic rings compared to the mostly indirect effects experienced by backbone NH atoms, as is apparent from the large magnitude of the induced shift, it should also be possible to fit the position and orientation of aromatic rings in our case. For this purpose, we examined the agreement between experimental and calculated shift perturbations induced by a single benzene ring within a grid of ring origin locations and ring normal orientations (Figure 6). In addition, to compensate for the rotational degeneracy of the simple Pople dipole model, we also incorporate the probability for the interaction geometry parameters (r, θ) as extracted from the PDB (Figures 3B and S7−S10).

For a single or fused aromatic ring system in the ligand, one benzene ring is sufficient to find the correct position and orientation of the ring responsible for the largest effect (Figure 6B,C). When a second fused ring system is present, a second benzene ring is necessary to find a good agreement between experimental and calculated values (Figure 6A). In practice, the necessity of a second ring can be inferred from both the structure of the ligand as well as the chemical shift perturbation pattern induced by ligand binding. In all examples, the smallest fit values are found for grid positions resembling the experimental structures. If the correct assumptions are made, we find ring conformations close to the experimental ones for our full set of ligands (Figure S11) As we have used a structure provided by AlphaFold[29,30] in all the fitting examples, this approach is available even in the absence of an experimentally determined apo-structure and is robust for minor rearrange-

https://doi.org/10.1021/acs.jmedchem.4c01128
J. Med. Chem. 2024, 67, 13187−13196

Journal of Medicinal Chemistry pubs.acs.org/jmc Article

ments of side-chain conformations upon binding. Apart from finding the most probable origin of the induced chemical shift as described above, one straightforward application is to evaluate and rank ligand docking poses based on the measured chemical shifts as well as the methyl-ring geometries.

CONCLUSIONS

FBDD relies on the initial discovery of weak binding ligands and subsequent improvement of noncovalent interactions between said ligands and the protein target. At both stages, NMR can serve as a valuable tool due to the sensitivity of the chemical shift, even to small variations of the chemical environment on an atomic level. One especially prominent form of interaction between proteins and their small molecule ligands are CH−hydrogen bonding interactions, involving aromatic ring systems and both sp^2- and sp^3-hybridized carbons, commonly termed CH−π interactions. Though they are weak individually, they serve to fine-tune the overall intermolecular interaction and consequently have a big impact on both affinity and selectivity. In this work, we show that one can directly detect and quantify methyl−π interactions even in the absence of a protein−ligand cocrystal structure. The binding site can be derived on a sequence level from the residue-specific experimental chemical shifts, while the contribution of methyl−π interactions can be gauged from the magnitude of the chemical shift change. The use of selectively labeled amino acid precursors leads to increased signal-to-noise levels and, if needed, a residue-specific simplification of the resulting methyl spectra.

While there are some discrepancies between the measured and expected chemical shift values, these can easily be attributed to increased flexibility of side chains and ligands as well as to the presence of homoaromatic groups that, while not being explicitly covered by the Pople model, can be approximated to a satisfactory degree as we demonstrated. The observed differences that can be attributed to the lack of ring parametrization follow a linear relationship and can therefore be estimated as a simple scaling factor from measuring multiple methyl positions and, if necessary, different but structurally related ligands.[31]

We further showed that there is a specific interaction geometry for methyl−π interactions both in our ligand set as well as in the PDB. The distinct orientational propensity of methyl groups relative to aromatic rings, in combination with the exquisite dependence of the proton chemical shift on these orientations, allows one to reliably determine or extract protein−ligand geometries.

The combination of the binding geometry probabilities extracted from the PDB and the simple chemical shift model to fit the position and orientation of aromatic rings from a small set of methyl shifts dramatically improves the fit. Positioning of the dummy aromatic centers should in many cases resolve the binding site and conformation of the ligand aromatic groups if they interact directly with a methyl group of the protein.

As the computation time to calculate the chemical shift is negligible compared to more sophisticated QM and density functional theory methods, this approach provides fast iterative feedback for medicinal chemists in the process of fragment refinement on the interaction geometry of crucial methyl−π interactions and allows for a more informed decision-making approach, especially as the relative intensity of the measured signal provides insights into residual mobility of the ligand in the binding pocket and therefore indirectly reports on binding energies as demonstrated within our set of BRD4-BD1 ligands.

Additional labeling strategies already presented or in development make additional aliphatic groups available for use with this approach.[32,33] In the future, we aim to be able to compile these into a novel and comprehensive refinement approach that can unanimously probe the geometries of CH−π interactions in protein−ligand binding sites and subsequently determine the complex structures at unprecedented speed. Due to the prevalence of aromatic centers in both small fragments and larger, highly optimized ligands, this will allow identification and evaluation of docked conformation at all stages of development.

EXPERIMENTAL SECTION

Ligand Synthesis. Ligands 1−10 were synthesized according to Platzer et al.[9] and are >95% pure by high-performance liquid chromatography (HPLC) analysis (Figures S12−S21). Compounds 11−16 were purchased from commercial vendors and tested without additional HPLC control.

Protein Expression and Purification. Human BRD4-BD1 (44−168) containing a TEV-cleavable His6-tag was overexpressed in *E. coli* BL21(DE3). 4L of bacteria were grown to an optical density of 0.6. After pelleting (20 °C) and resuspension in 1L of M9 minimal medium, expression was induced with a final IPTG concentration of 0.4 mM and continued overnight. After harvesting, the cells were resuspended in 30 mL of phosphate buffer (20 mM sodium phosphate, 500 mM NaCl, 20 mM Imidazole, pH 7.5), sonicated, and subsequently loaded onto a Ni^{2+} affinity column. After step gradient elution and buffer exchange of the protein-containing fractions under low imidazole conditions, the protein was subjected to TEV-cleavage overnight (4 °C). His6-TEV protease and cleaved his6-tag was removed by a second pass over a Ni^{2+} affinity columns. After subsequent buffer exchange (10 mM sodium phosphate, 100 mM NaCl, pH7.5) and concentration to 0.3 mM using a 3kD cutoff centrifugal filter device, the protein was stored at −20 °C.

All samples were uniformly ^{15}N-labeled. To achieve selective labeling of leucine and isoleucine δ positions, either 2-ketoisocaproate or 2-ketobutyrate was added at a final concentration of 100 mg/L to the minimal media prior to induction by IPTG. All precursors used in this work were provided by MAG-LAB.

NMR Measurements and Analysis. All protein NMR measurements were carried out at 298 K on a Bruker Avance HD3+ 600 MHz spectrometer. 2D $^1H-^{13}C$ HSQC spectra were acquired using the pulse sequence "hsqcetgpsi" of the Bruker library[34−36] using 64 (t_1) × 512 (t_2) complex points with acquisition times of 16.9 ms (t_1) and 53.2 ms (t_2). Assignment of Leu-δ methyl groups was achieved by mutational analysis. Four BRD4 mutants were produced selectively labeled (V87A, V90A, L92A, L94A) to assign resonances of those key residues in the binding site (Figure S2). Worth noticing here is that the mutation to an alanine leads to subtle rearrangements in the conformation of different neighboring residues, which in turn leads to chemical shift perturbations of other methyl resonances. Thus, in our case, unambiguous assignment of the four residues of interest (V87, V90, L92, L94) was only possible after the acquisition of spectra of all four mutants.

Stereospecific assignment of the two diastereotopic methyl groups of leucine and valine residues was achieved by fractional ^{13}C-labeling described previously.[37]

All samples contained a final protein concentration of 0.2 mM, and a final ligand concentration of 1 mM.

PDB Statistics. To generate statistics about distance and angle distributions between methyl groups of leucine residues and aromatic rings of ligand molecules, all X-ray structures in the PDB with a resolution smaller than 2.5 Å and a r-free smaller than 0.24 were used. All nonpolymer residues were treated as ligands and the parameters for all aromatic rings of this set were extracted as well as the parameters for all Leucines found in polymeric sequences. To reduce

https://doi.org/10.1021/acs.jmedchem.4c01128
J. Med. Chem. 2024, 67, 13187−13196

Journal of Medicinal Chemistry pubs.acs.org/jmc Article

the redundancy of the set, the protein sequences were clustered into groups with a sequence similarity ≥95%. Consequently, for each leucine, a new sequence number was assigned representing its position in a multiple sequence alignment of all cluster members. The set was then further filtered to contain only the one leucine with the smallest distance between methyl carbon and ring center for each combination of cluster ID, cluster sequence number, and ligand name.

The kernel density estimation (kde) visualizations in Figure 3A−C were produced in R using the "kde" package with a grid size of 50 points in both dimensions.

QM Chemical Shift Calculations. Calculations were performed for the structures of ligands 4 and 7 as well as the APO form. The systems consisted of Leu92, Leu94, and Ile146, along with the side chains of residues located within a 4 Å distance from the methyl groups of interest, namely Pro82, Val87, Cys136, Tyr139, and Asn140.[38] In the case of HOLO forms, the ligands were also included. For leucine and isoleucine residues, the N- and C-termini were capped as amides with the preceding carbonyl carbon and with the following amino group, respectively. The remaining residues were truncated at the Cβ atoms. The systems were completed with dummy hydrogen atoms.

Chemical shieldings were calculated using the gauge independent atomic orbital (GIAO) approach, at the wB97XD/def2-TZVP level of theory, as implemented in the Gaussian16 package.[39−41] To determine the chemical shift perturbations (CSPs), we subtracted the chemical shielding values of the APO form from those of the HOLO form.

Fit of Ring Positions. The structure of apo BRD4 was downloaded from AlphaFold[29,30] (Uniprot: Q5BJ26, version: 2022−11−01) and shortened to include residues 42−168. A grid representing potential ring centers was constructed around the geometric center of $C\delta_{1/2}$ atoms of leucine 92 and 94 as well as the $C\delta_1$ atom of isoleucine 146 with uniform dimensions of 10 Å and a grid spacing of 1 Å. This grid is further simplified by removing positions that are within 3 Å of any non-hydrogen atom in the structure. To sample the potential orientations of the ring normal, the φ and ψ spherical coordinates of the vertex positions of a once subdivided icosahedron are extracted. As parallel normals are equivalent due to the squared cosine term of the Pople equation, φ and ψ ranges of $[0, \pi]$ and $[0, \pi]$ are sufficient to effectively sample the complete icosphere. This leads to a final grid with dimensions 11 × 11 × 11 Å and 22 orientations. For each grid point, the distance r and angle θ are calculated for each above-mentioned methyl-carbon position, and the root of the squared difference between the experimentally measured CSP and the Pople model is calculated. If two rings are to be sampled simultaneously, the combination of each grid position with all others is evaluated, and combinations that are within 3.5 Å of one another are excluded.

To incorporate the probability of different methyl−π geometries, a two-dimensional Gaussian kernel density estimation (kde) was constructed based on the distance and θ angle between methyl carbons and aromatic rings in the PDB. The kde was constructed separately for $C\delta_{1/2}$ positions of leucine and $C\delta_1$ positions of isoleucine and weighted by $1/4 \times \pi \times r^2$. The kde probabilities are subsequently rescaled by subtracting them from the maximum value. Just as before, for each grid position, the kde is evaluated for each methyl-carbon position and the probabilities are summed up.

The delta shift root of squared differences and the squared kde probabilities are summed up to a single final fit value, and the grid position of the minimal value is reported as the expected ring position and orientation. Even small deviations from the true position and orientation of aromatic rings in the crystal structure lead to a major increase in the fit value both for the summed value and for its single components (Figures S7−S9B−D).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jmedchem.4c01128.

Compound structures, molecular formula strings, measured and calculated chemical shift values, B-factor comparison, Pople model coefficients, interaction geometry model, protein methyl assignment information, partial peak overlap, relative methyl ψ-X frequency in the PDB, top-down methyl vector projection, stepwise proton based chemical shift calculations, ring fit statistics, HPLC traces (PDF)

Molecular_formula_strings (CSV)

New PDB IDs: 9FWX and 9FXP. Authors will release the atomic coordinates upon article publication.

■ AUTHOR INFORMATION

Corresponding Author

Andreas Beier − *Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Department of Structural and Computational Biology, Max Perutz Laboratories, University of Vienna, 1030 Vienna, Austria; Vienna Doctoral School of Chemistry, University of Vienna, 1090 Vienna, Austria; Department of Structural and Computational Biology, Max Perutz Laboratories, 1030 Vienna, Austria;* orcid.org/0000-0002-1934-8926; Email: andreas.beier@univie.ac.at

Authors

Gerald Platzer − *MAG-LAB, 1030 Vienna, Austria*

Theresa Höfurthner − *Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Department of Structural and Computational Biology, Max Perutz Laboratories, University of Vienna, 1030 Vienna, Austria; Vienna Doctoral School of Chemistry, University of Vienna, 1090 Vienna, Austria; Department of Structural and Computational Biology, Max Perutz Laboratories, 1030 Vienna, Austria*

Aleksandra L. Ptaszek − *Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Department of Structural and Computational Biology, Max Perutz Laboratories, University of Vienna, 1030 Vienna, Austria; Vienna Doctoral School of Chemistry, University of Vienna, 1090 Vienna, Austria*

Roman J. Lichtenecker − *Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Institute of Organic Chemistry, University of Vienna, 1090 Vienna, Austria; MAG-LAB, 1030 Vienna, Austria*

Leonhard Geist − *Boehringer Ingelheim RCV GmbH & Co. KG, 1121 Vienna, Austria*

Julian E. Fuchs − *Boehringer Ingelheim RCV GmbH & Co. KG, 1121 Vienna, Austria*

Darryl B. McConnell − *Curie Bio, Boston, Massachusetts 02115-3153, United States;* orcid.org/0000-0002-2537-3458

Moriz Mayer − *Boehringer Ingelheim RCV GmbH & Co. KG, 1121 Vienna, Austria*

Robert Konrat − *Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Department of Structural and Computational Biology, Max Perutz Laboratories, University of Vienna, 1030 Vienna, Austria; Department of Structural and Computational Biology, Max Perutz Laboratories, 1030 Vienna, Austria*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jmedchem.4c01128

https://doi.org/10.1021/acs.jmedchem.4c01128
J. Med. Chem. 2024, 67, 13187−13196

Notes

The authors declare no competing financial interest.
The code used to extract interaction geometries from the PDB and fit the dummy ring positions is available upon reasonable request.

■ ACKNOWLEDGMENTS

Andreas Beier, Theresa Höfurthner, and Aleksandra L. Ptaszek were funded by the Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology. We want to thank Gerd Bader, Dirk Kessler, and Dirk Reinert for providing the crystallographic data used in this work.

■ REFERENCES

(1) Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H. Twenty years on: the impact of fragments on drug discovery. *Nat. Rev. Drug Discov* **2016**, *15*, 605−619.
(2) Shi, Y.; von Itzstein, M. How Size Matters: Diversity for Fragment Library Design. *Molecules* **2019**, *24*, 2838.
(3) Hann, M. M.; Leach, A. R.; Harper, G. Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856−864.
(4) Giordanetto, F.; Jin, C.; Willmore, L.; Feher, M.; Shaw, D. E. Fragment Hits: What do They Look Like and How do They Bind? *J. Med. Chem.* **2019**, *62*, 3381−3394.
(5) Nishio, M.; Umezawa, Y.; Fantini, J.; Weiss, M. S.; Chakrabarti, P. CH−π hydrogen bonds in biological macromolecules. *Phys. Chem. Chem. Phys.* **2014**, *16*, 12648−12683.
(6) Ferreira de Freitas, R.; Schapira, M. A systematic analysis of atomic protein−ligand interactions in the PDB. *Medchemcomm* **2017**, *8*, 1970−1981.
(7) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science* **1996**, *274* (5292), 1531−1534.
(8) Harner, M. J.; Frank, A. O.; Fesik, S. W. Fragment-based drug discovery using NMR spectroscopy. *J. Biomol NMR* **2013**, *56*, 65−75.
(9) Platzer, G.; et al. PI by NMR: Probing CH−π Interactions in Protein−Ligand Complexes by NMR Spectroscopy. *Angew. Chem.* **2020**, *132*, 14971−14978.
(10) Jiang, Y.; Kalodimos, C. G. NMR Studies of Large Proteins. *J. Mol. Biol.* **2017**, *429*, 2667−2676.
(11) Ardenkjaer-Larsen, J.-H.; et al. Facing and Overcoming Sensitivity Challenges in Biomolecular NMR Spectroscopy. *Angew. Chem., Int. Ed.* **2015**, *54*, 9162−9185.
(12) Tugarinov, V.; Hwang, P. M.; Ollerenshaw, J. E.; Kay, L. E. Cross-Correlated Relaxation Enhanced ^{1}H−^{13}C NMR Spectroscopy of Methyl Groups in Very High Molecular Weight Proteins and Protein Complexes. *J. Am. Chem. Soc.* **2003**, *125*, 10420−10428.
(13) Kainosho, M.; et al. Optimal isotope labelling for NMR protein structure determinations. *Nature* **2006**, *440*, 52−57.
(14) Dias, D. M.; Ciulli, A. NMR approaches in structure-based lead discovery: Recent developments and new frontiers for targeting multi-protein complexes. *Prog. Biophys. Mol. Biol.* **2014**, *116*, 101−112.
(15) Dhalluin, C.; et al. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **1999**, *399*, 491−496.
(16) Filippakopoulos, P.; et al. Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family. *Cell* **2012**, *149*, 214−231.
(17) Gacias, M.; et al. Selective Chemical Modulation of Gene Transcription Favors Oligodendrocyte Lineage Progression. *Chem. Biol.* **2014**, *21*, 841−854.
(18) Lichtenecker, R. J.; Coudevylle, N.; Konrat, R.; Schmid, W. Selective Isotope Labelling of Leucine Residues by Using α-Ketoacid Precursor Compounds. *ChemBioChem.* **2013**, *14*, 818−821.
(19) Gardner, K. H.; Kay, L. E. Production and Incorporation of ^{15}N, ^{13}C, ^{2}H (^{1}H-δ1 Methyl) Isoleucine into Proteins for Multidimensional NMR Studies. *J. Am. Chem. Soc.* **1997**, *119*, 7599−7600.
(20) Pople, J. A. Proton Magnetic Resonance of Hydrocarbons. *J. Chem. Phys.* **1956**, *24*, 1111−1111.
(21) Hansen, D. F.; Neudecker, P.; Kay, L. E. Determination of Isoleucine Side-Chain Conformations in Ground and Excited States of Proteins from Chemical Shifts. *J. Am. Chem. Soc.* **2010**, *132*, 7589−7591.
(22) Ran, J.; Wong, M. W. Saturated Hydrocarbon−Benzene Complexes: Theoretical Study of Cooperative CH/π Interactions. *J. Phys. Chem. A* **2006**, *110*, 9702−9709.
(23) Berman, H. M. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(24) Case, D. A. Calibration of ring-current effects in proteins and nucleic acids. *J. Biomol. NMR* **1995**, *6*, 341.
(25) Turchin, K. F. Paramagnetic ring current in the diazepine ring of the 1H-benzo-1,5-diazepinium monocation. *Chem. Heterocycl Compd (N Y)* **1974**, *10*, 720−723.
(26) Haketa, Y.; et al. Self-Associating Curved π-Electronic Systems with Electron-Donating and Hydrogen-Bonding Properties. *J. Am. Chem. Soc.* **2020**, *142*, 16420−16428.
(27) Eglen, R. M.; et al. The Use of AlphaScreen Technology in HTS: Current Status. *Curr. Chem. Genomics* **2008**, *1*, 2−10.
(28) McCoy, M. A.; Wyss, D. F. Spatial Localization of Ligand Binding Sites from Electron Current Density Surfaces Calculated from NMR Chemical Shift Perturbations. *J. Am. Chem. Soc.* **2002**, *124*, 11758−11763.
(29) Varadi, M.; et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **2022**, *50*, D439−D444.
(30) Jumper, J.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583−589.
(31) Agarwal, A.; Barnes, J. A.; Fletcher, J. L.; McGlinchey, M. J.; Sayer, B. G. Ring currents, local anisotropy, and the problem of the out-of-plane protons: a reinvestigation of the nuclear magnetic resonance spectrum of [10]-paracyclophane. *Can. J. Chem.* **1977**, *55*, 2575−2581.
(32) Toscano, G.; et al. ^{13}Cβ-Valine and ^{13}Cγ-Leucine Methine Labeling To Probe Protein Ligand Interaction. *ChemBioChem* **2024**, *25*, No. e202300762.
(33) Höfurthner, T.; et al. Synthesis of a 13C-methylene-labeled isoleucine precursor as a useful tool for studying protein side-chain interactions and dynamics. *J. Biomol NMR* **2024**, *78*, 1−8.
(34) Schleucher, J.; et al. A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J. Biomol. NMR* **1994**, *4*, 301.
(35) Kay, L.; Keifer, P.; Saarinen, T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **1992**, *114*, 10663−10665.
(36) Palmer, A. G.; Cavanagh, J.; Wright, P. E.; Rance, M. Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *Journal of Magnetic Resonance (1969)* **1991**, *93*, 151−170.
(37) Neri, D.; Szyperski, T.; Otting, G.; Senn, H.; Wuethrich, K. Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional carbon-13 labeling. *Biochemistry* **1989**, *28*, 7510−7516.
(38) Swails, J.; Zhu, T.; He, X.; Case, D. A. AFNMR: automated fragmentation quantum mechanical calculation of NMR chemical shifts for biomolecules. *J. Biomol NMR* **2015**, *63*, 125−139.
(39) Schreckenbach, G.; Ziegler, T. Calculation of NMR Shielding Tensors Using Gauge-Including Atomic Orbitals and Modern Density Functional Theory. *J. Phys. Chem.* **1995**, *99*, 606−611.
(40) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom−atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615.
(41) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn:

https://doi.org/10.1021/acs.jmedchem.4c01128
J. Med. Chem. 2024, 67, 13187−13196

Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.

https://doi.org/10.1021/acs.jmedchem.4c01128
J. Med. Chem. 2024, 67, 13187−13196

Supporting Information

Probing Protein-Ligand Methyl-π Interaction Geometries Through Chemical Shift Measurements of Selectively Labelled Methyl Groups

Andreas Beier[1,3,4*], Gerald Platzer[6], Theresa Höfurthner[1,3,4], Aleksandra L. Ptaszek[1,3], Roman J. Lichtenecker[2,6], Leonhard Geist[5], Julian E. Fuchs[5], Darryl B. McConnell[7] Moriz Mayer[5] and Robert Konrat[1,4]

[1] Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Department of Structural and Computational Biology, Max Perutz Laboratories, University of Vienna, Campus Vienna Biocenter 5, 1030 Vienna (Austria).

[2] Christian Doppler Laboratory for High-Content Structural Biology and Biotechnology, Institute of Organic Chemistry, University of Vienna, Währingerstraße 38, 1090 Vienna (Austria).

[3] Vienna Doctoral School of Chemistry, University of Vienna, Währingerstraße 38, 1090 Vienna (Austria)

[4] Max Perutz Laboratories, Department of Structural and Computational Biology, Campus Vienna Biocenter 5, 1030 Vienna.

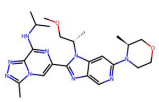
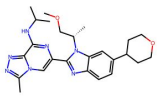
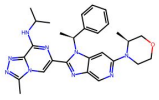
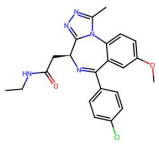
[5] Boehringer Ingelheim RCV GmbH & Co. KG, Dr. Boehringergasse 5-11, 1121 Vienna.

[6] MAG-LAB, Karl-Farkas-Gasse 22, 1030 Vienna.

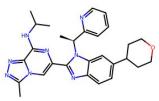
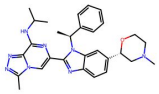
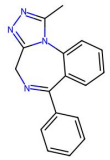
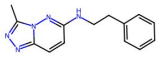
[7] Curie.Bio, 177 Huntington Ave Ste 1703, Boston, MA 02115-3153

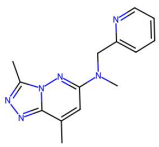
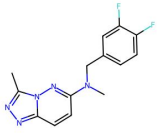
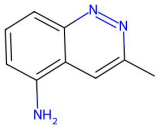
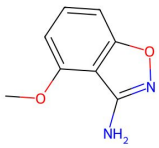
** corresponding author Andreas Beier, andreas.beier@univie.ac.at*

Supporting Figures and Tables

Index	IC50 (uM)	Smiles	Structure	ILE-146-CD1			LEU-92-CD1			LEU-92-CD2			LEU-94-CD1			LEU-94-CD2		
				exp	calc	I/I0	exp	calc	I/I0	exp	calc	I/I0	exp	calc	I/I0	exp	calc	I/I0
1	0.036	COC[C@H](C)n1c(-c2cn3c(C)nnc3c(NC(C)C)n2)nc2cnc(N3CCOC[C@@H]3C)cc21		1.04	1.15	0.91	0.90	1.43	0.54	1.06	1.40	0.63	0.25	0.26	0.96	0.03	-0.04	1.01
2	0.049	COC[C@H](C)n1c(-c2cn3c(C)nnc3c(NC(C)C)n2)nc2ccc(C3CCOCC3)cc21		1.02	1.03	0.89	1.01	1.47	0.46	1.21	1.38	0.64	0.24	0.29	1.01	0.03	-0.03	0.97
3	0.038	Cc1nnc2c(NC(C)C)nc(-c3nc4cnc(N5CCOC[C@@H]5C)cc4n3[C@@H](C)c3ccccc3)cn12		1.02	1.15	0.77	0.80	1.29	0.50	1.00	1.27	0.58	0.23	0.34	0.85	0.02	-0.01	0.93
4	0.249	CCNC(=O)C[C@@H]1N=C(c2ccc(Cl)cc2)c2cc(OC)ccc2-n2c(C)nnc21		1.03	0.64	0.64	0.77	1.05	0.47	0.33	0.24	0.72	0.24	0.24	0.85	0.11	0.03	0.82

S2

Index	IC50 (uM)	Smiles	Structure	ILE-146-CD1			LEU-92-CD1			LEU-92-CD2			LEU-94-CD1			LEU-94-CD2		
				exp	calc	I/I0	exp	calc	I/I0	exp	calc	I/I0	exp	calc	I/I0	exp	calc	I/I0
5	0.065	Cc1nnc2c(NC(C)C)nc(-c3nc4ccc(C5CCOCC5)cc4n3[C@@H](C)c3ccccn3)cn12		1.03	0.85	0.55	0.93	1.37	0.34	1.15	1.41	0.35	0.23	0.32	0.76	0.01	-0.01	0.51
6	0.031	Cc1nnc2c(NC(C)C)nc(-c3nc4ccc([C@H]5CN(C)CCO5)cc4n3[C@@H](C)c3ccccc3)cn12		1.02	0.80	0.50	0.88	1.10	0.52	1.19	1.78	0.36	0.22	0.37	0.86	0.01	0.01	0.94
7	16.596	Cc1nnc2n1-c1ccccc1C(c1ccccc1)=NC2		0.91	0.31	0.33	0.81	1.22	0.24	0.32	0.31	0.71	0.21	0.03	0.65	broadened		
8	50.926	Cc1nnc2ccc(NCCc3ccccc3)nn12		0.45	0.51	0.20	0.17	0.15	0.37	0.00	-0.19	0.59	0.54	0.66	0.17	0.10	-0.05	0.55

Index	IC50 (uM)	Smiles	Structure	ILE-146-CD1			LEU-92-CD1			LEU-92-CD2			LEU-94-CD1			LEU-94-CD2		
				exp	calc	I/I0	exp	calc	I/I0	exp	calc	I/I0	exp	calc	I/I0	exp	calc	I/I0
9		Cc1cc(N(C)Cc2ccccn2)nn2c(C)nnc12		0.97	2.41	0.08	0.22	0.18	0.27	-0.10	-0.18	0.83	0.43	0.34	0.50	0.01	-0.08	0.79
10	3.283	Cc1nnc2ccc(N(C)Cc3ccc(F)c(F)c3)nn12		broadened			0.24	0.32	0.53	-0.04	-0.17	1.49	0.63	0.41	0.27	0.06	-0.05	0.62
11		Cc1cc2c(N)cccc2nn1		0.40	0.63	0.26	0.17	0.30	0.54	0.08	-0.14	0.88	0.32	0.90	0.30	0.05	-0.03	1.07
12		COc1cccc2onc(N)c12		0.18	0.64	0.89	0.35	0.36	0.33	0.10	-0.08	0.79	0.35	0.83	0.23	0.01	-0.03	0.15

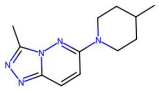
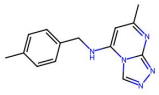
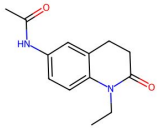
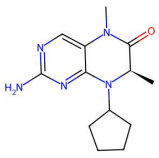
Index	IC50 (uM)	Smiles	Structure	ILE-146-CD1			LEU-92-CD1			LEU-92-CD2			LEU-94-CD1			LEU-94-CD2		
				exp	calc	I/I0	exp	calc	I/I0	exp	calc	I/I0	exp	calc	I/I0	exp	calc	I/I0
13		Cc1nnc2ccc(N3CCC(C)CC3)nn12		0.32	0.24	0.60	0.29	0.25	0.48	0.05	-0.09	0.75	0.68	1.06	0.39	0.11	-0.06	0.62
14		Cc1ccc(CNc2cc(C)nc3nncn23)cc1		0.90	1.75	0.12	0.19	-0.06	0.34	0.13	-0.13	0.58	0.14	0.34	0.58	0.00	-0.11	0.15
15		CCN1C(=O)CCc2cc(NC(C)=O)ccc21		0.47	0.61	0.28	0.46	0.81	0.17	0.25	0.02	0.54	0.05	0.21	0.67	0.01	0.03	1.05
16		C[C@@H]1C(=O)N(C)c2cnc(N)nc2N1C1CCCC1		0.49	0.48	0.32	0.41	0.72	0.25	0.10	-0.04	0.84	0.07	0.17	0.65	0.01	0.02	0.93

Table S1: List of ligands with their AlphaScreen IC_{50} values, as well as the measured and calculated shifts based on the Pople model

	B factor			
Ligand	Leu 92 CD1	Leu 94 CD1	Ile 146 CD1	Ring average
Ligand 1	11.35	13.55	11.29	7.79
Ligand 8	14.22	16.08	11.42	11.81
Ligand 9	18.09	19.26	20.64	19.38
Ligand 10	13.58	11.99	20.24	15.11

Table S2: B-factors for selected atoms and structures. Ring average is the average over all ring atoms of the double ring directly beneath Ile-146-CD1

n	number of ring electrons	6
e	elemetary charge	4.8032×10^{-10} Fr
m	electron mass	9.1094×10^{-28} g
c	speed of light	2.998×10^{8} cm/s
a	ring diameter (5 ring)	1.18×10^{-8} cm
	ring diameter (6 ring)	1.39×10^{-8} cm

Table S3: Pople model parameters used in cgs units.

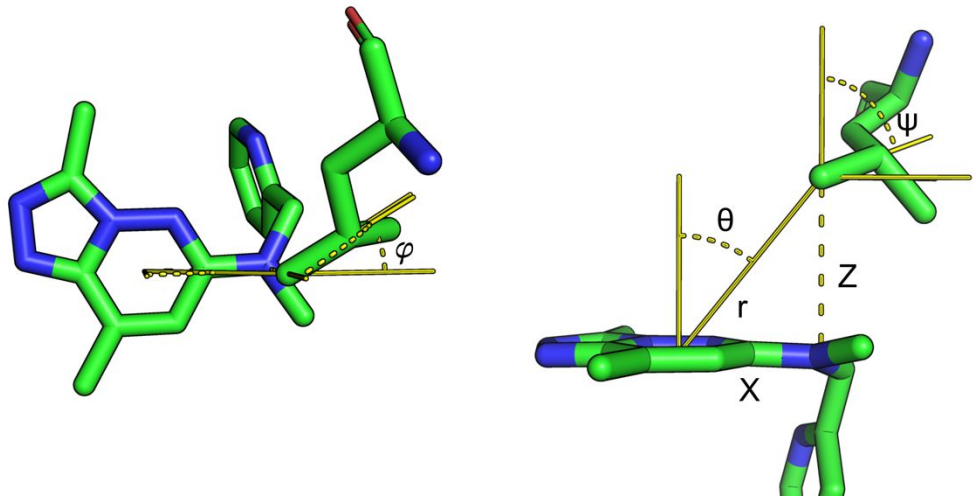


Figure S1: Interaction geometry parameters visualized on Ligand 8. For the pople model only θ and r are used as well as the ring size to select the correct diameter.

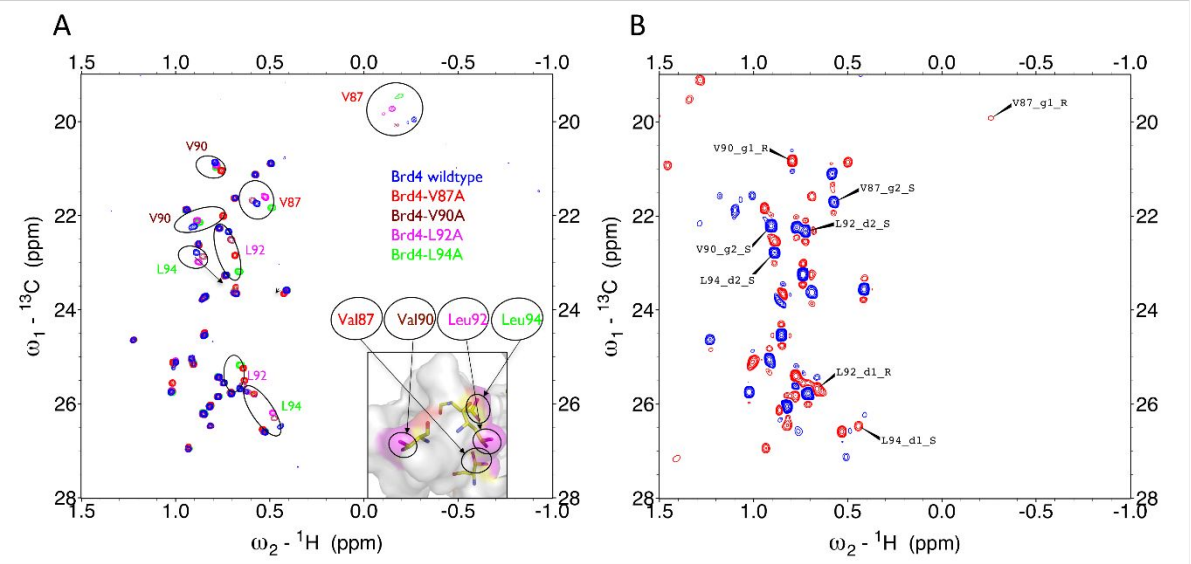


Figure S2: Assignment of selected Leucine and Valine resonances: A: Overlay of the spectrum of wildtype Brd4-BD1 in blue with the spectra of the four Brd4-BD1 mutants (V87A, V90A, L92A, L94A). The structure in (A) shows the 4 residues that are closest to the ligand binding site with V87, L92, L94 facing the ligand. B: Constant time ^{1}H-^{13}C HSQC of fractionally labeled Brd4-BD1 for stereospecific assignment of Leucine Cδ-resonances and Valine Cγ-resonances. Positive resonances (blue) are pro-S methyl groups, negative resonances (red) are pro-R methyl groups of Valine and Leucine residues.

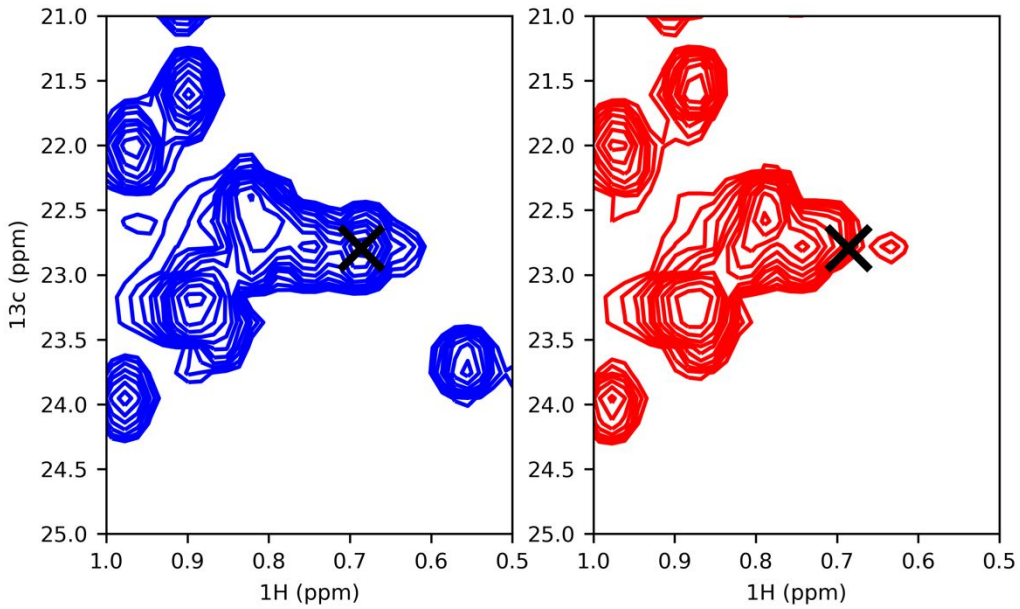


Figure S3: Overlap for position of apo Leu92 $C\delta_1$-$H\delta_1$ resonances (blue) in comparison with ligand 4 bound state (red)

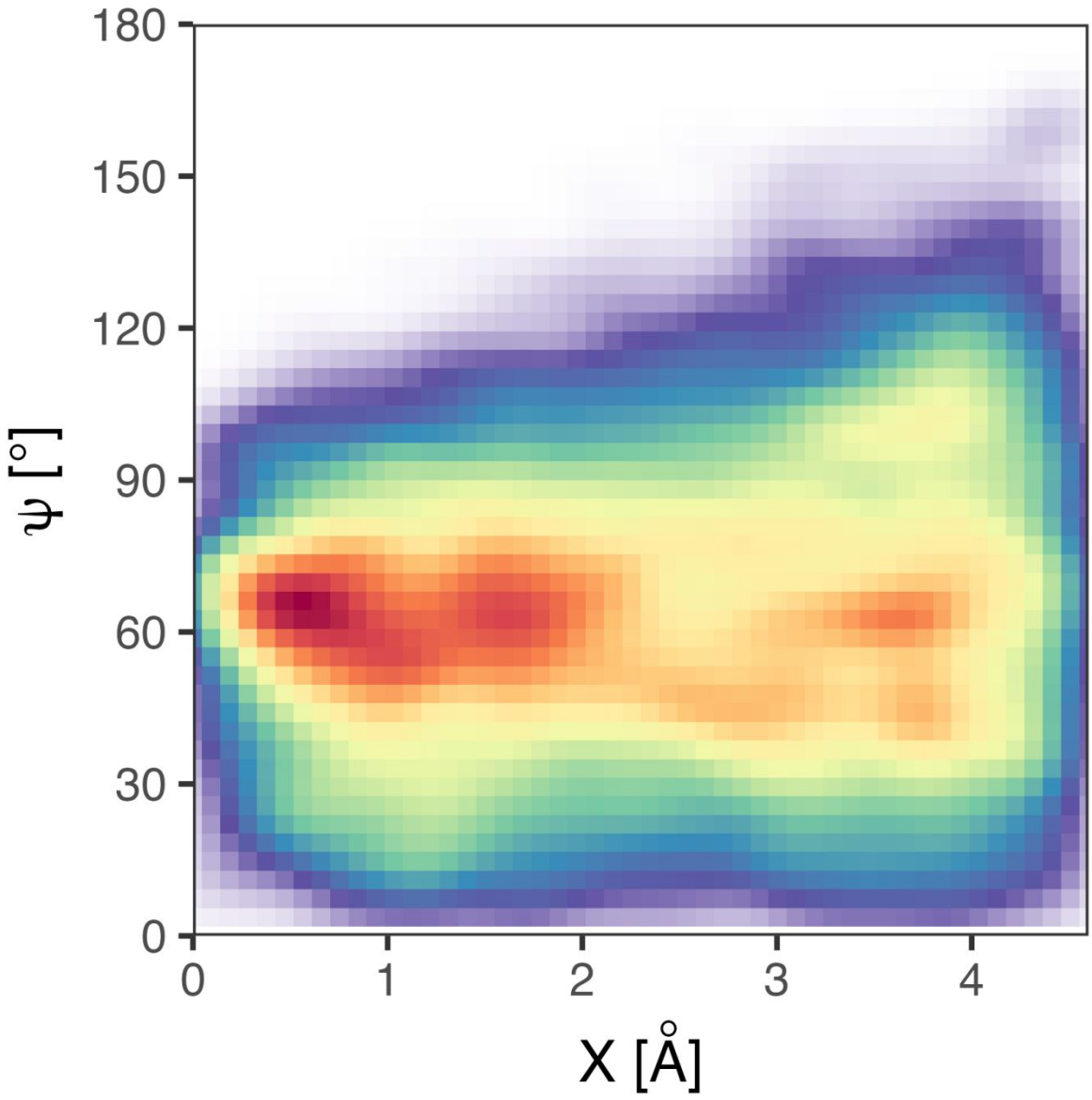


Figure S4: Relative frequency of methyl-vector ψ-values dependent in displacement along the ring plane (X)

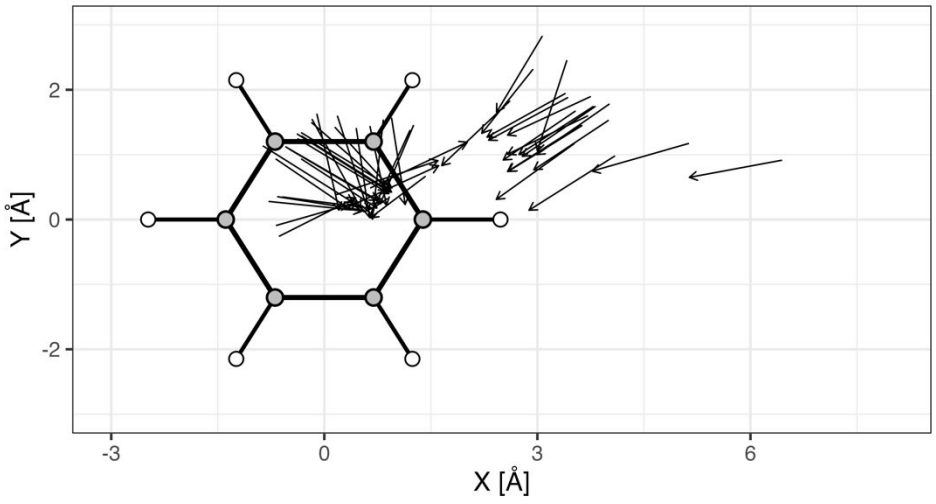


Figure S5: Top-down projection of methyl vectors for ligands 1-10

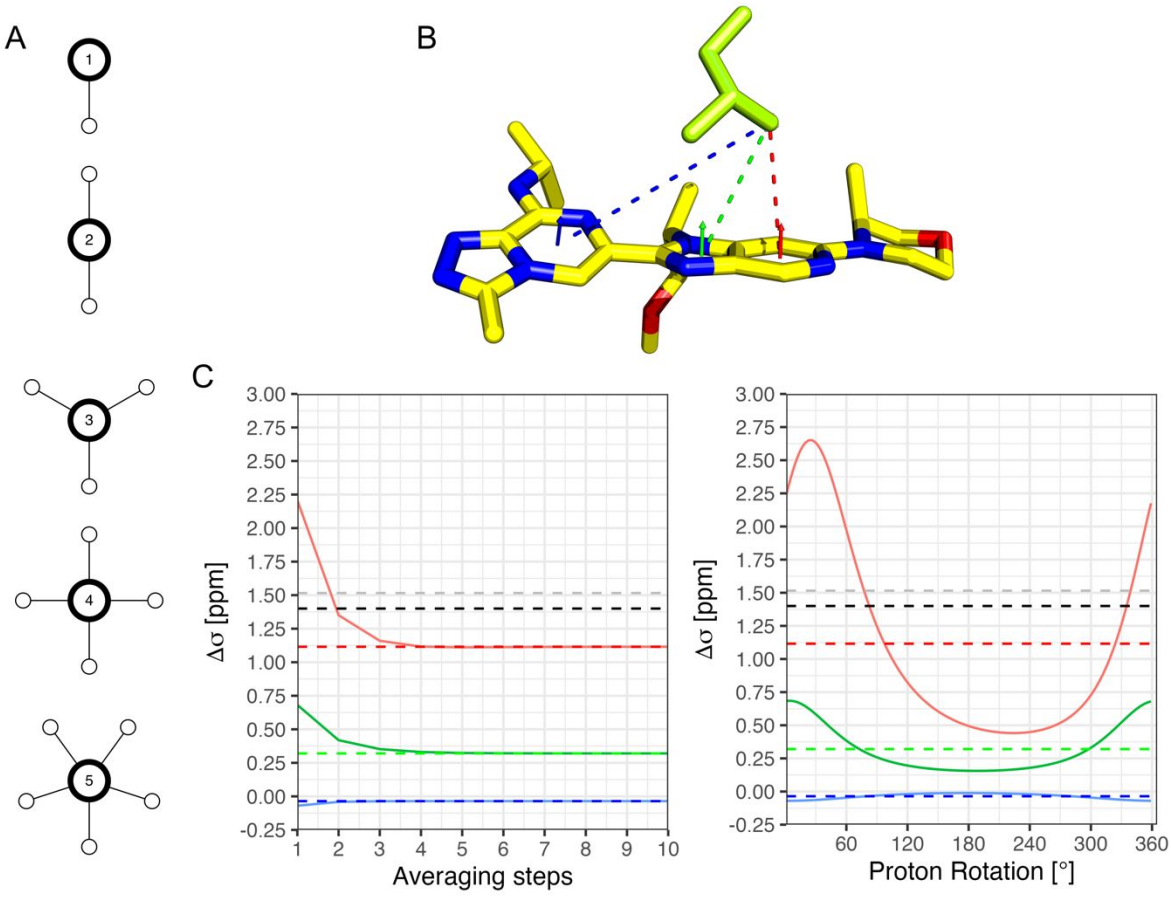


Figure S6: Comparison of calculated shift values based on proton rotation around the C-C symmetry axis. A: For stepwise averaging, the hydrogen is positioned along the normal direction of the ring and consequently rotated by a specified number of equal sized steps around the central C-C symmetry axis. B: For the calculation of the chemical shift change of Leu92-Cδ_2, 3 aromatic rings of ligand 1 within 8 Angstrom are taken into consideration (PDB: 6XUZ). Ring normals and center distances are indicated by arrows and lines. C: Changes in calculated shifts dependent on the number of hydrogen positions averaged over (left, see A) or as a function of the rotation angle for a single hydrogen atom (right). Colored lines represent ring positions as indicated in B. Dashed lines represent the average over 100 uniformly spaced rotation steps. Black and gray lines represent the sum of the averages for all 3 rings and the scaled experimental value respectively.

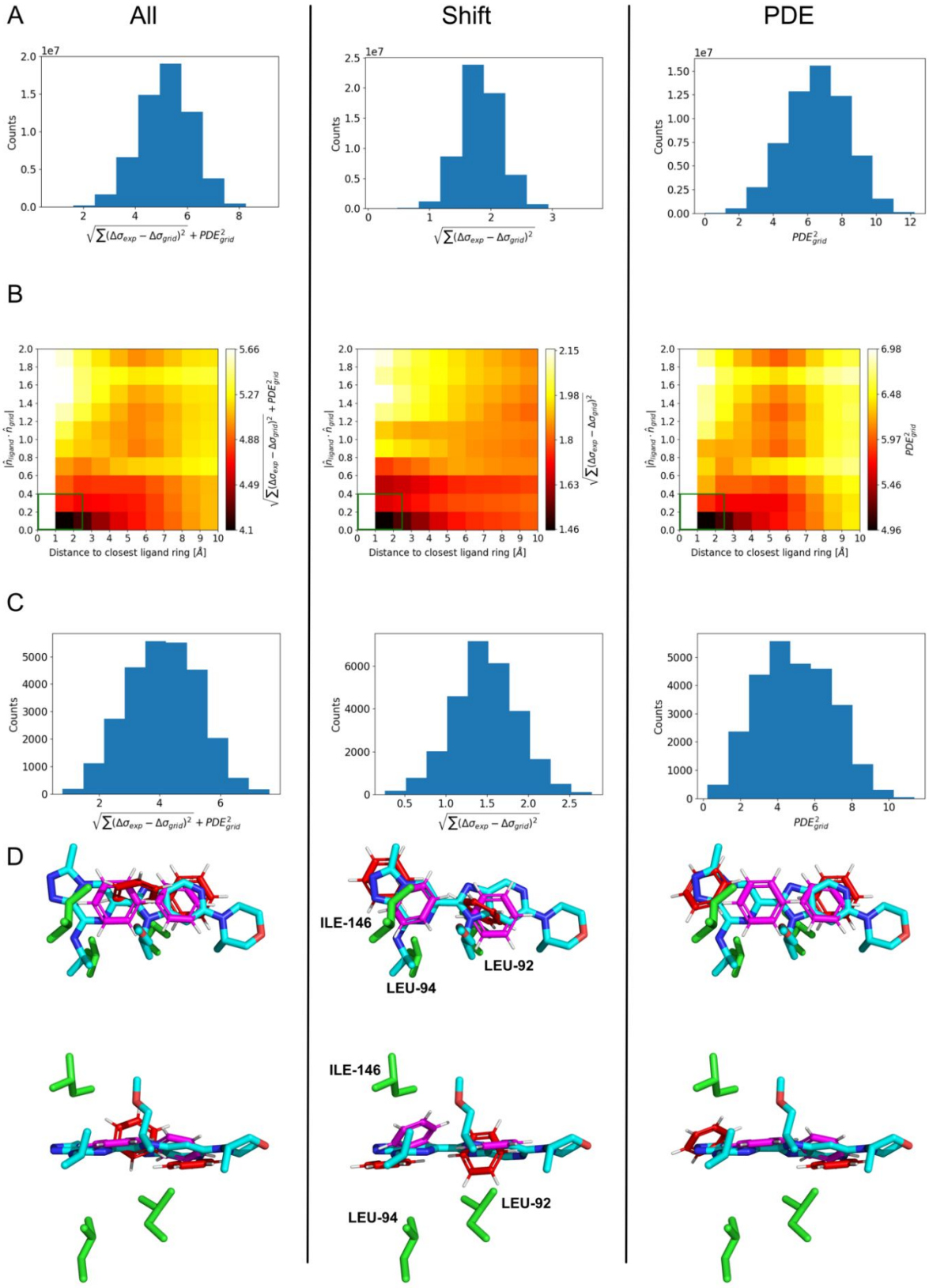


Figure S7: Ring fit statistics for Ligand 1 (PDB: 6XUZ) and the respective shift and PDB probability density estimation components. A: Histogram of the fit score over all grid positions. B: Average fit score binned by the summed distance to the closest ring in the respective crystal structure (x-axis) and the sum of the absolute dot products between the respective ring normals in the fit and the experimental structure (y-axis). C: Histogram of the region indicated by a green box in C. D: Best (magenta) and worst (red) ring position based on fit values in the region indicated by a green box in C. Fitting 2 rings simultaneously and therefore evaluating expanding the single-ring grid by itself leads to a large increase in histogram counts compared to ligands 8 (Figure S8) and 12 (Figure S9)

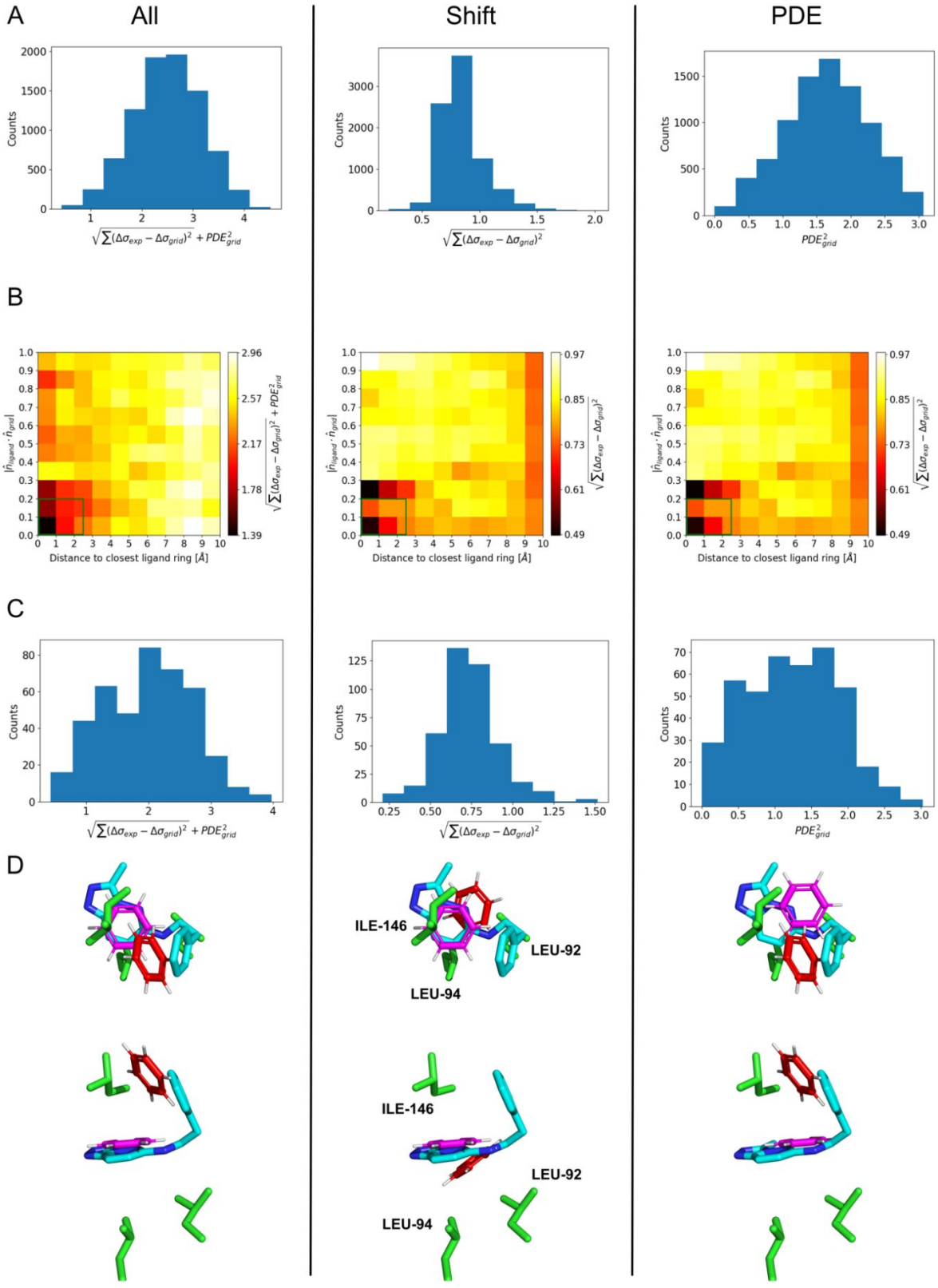


Figure S8: Ring fit statistics for Ligand 8 (PDB: 9FWX) and the respective shift and PDB probability density estimation components. A: Histogram of the fit score over all grid positions. B: Average fit score binned by the distance to the closest ring in the respective crystal structure (x-axis) and the absolute dot products between the respective ring normals in the fit and the experimental structure (y-axis). C: Histogram of the region indicated by a green box in C. D: Best (magenta) and worst (red) ring position based on fit values in the region indicated by a green box in C.

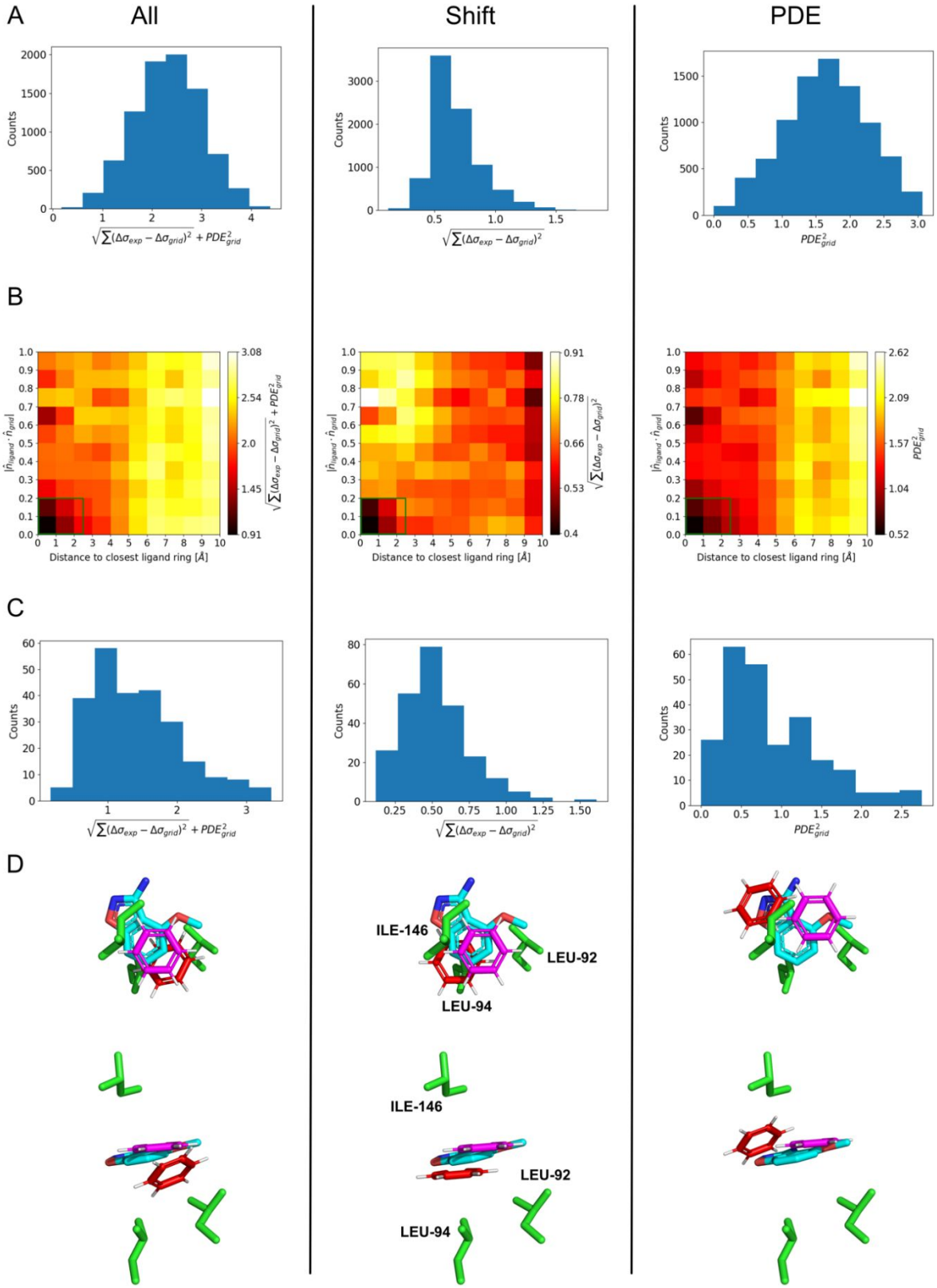


Figure S9: Ring fit statistics for Ligand 12 (PDB: 9FXP) and the respective shift and PDB probability density estimation components. A: Histogram of the fit score over all grid positions. B: Average fit score binned by the distance to the closest ring in the respective crystal structure (x-axis) and the absolute dot products between the respective ring normals in the fit and the experimental structure (y_axis). C: Histogram of the region indicated by a green box in C. D: Best (magenta) and worst (red) ring position based on fit values in the region indicated by a green box in C.

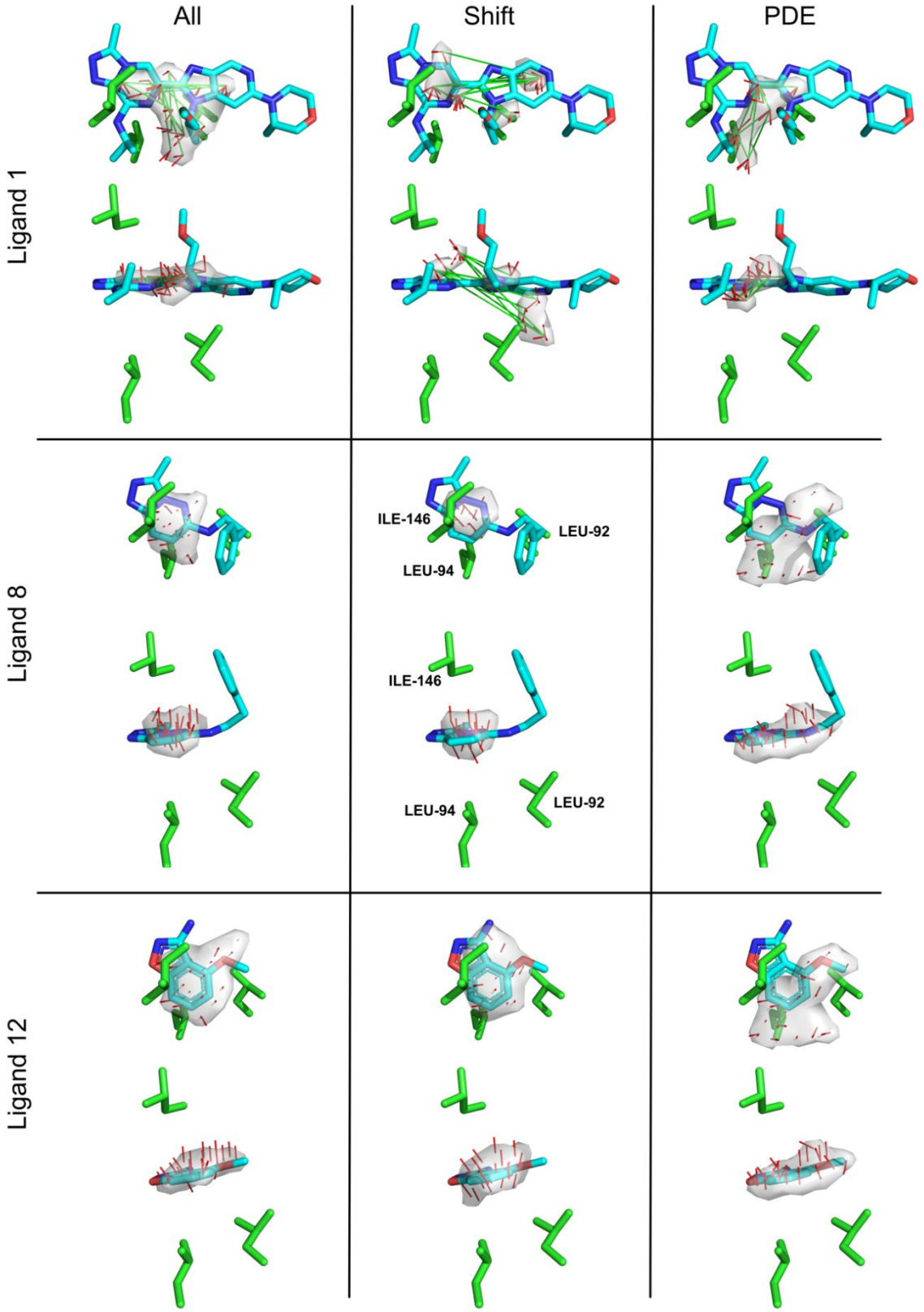


Figure S10: Top and side view of the top 20 result of the ring fit based on overall fit value and its respective shift and PDB probability density estimation components for ligand 1, 8 and 12. Surfaces indicate variability in the ring position while red lines represent the fitted ring normal and therefore ring orientations. For ligand 1, two rings are fit simultaneously, and the green lines indicate connecting vectors between ring pairs.

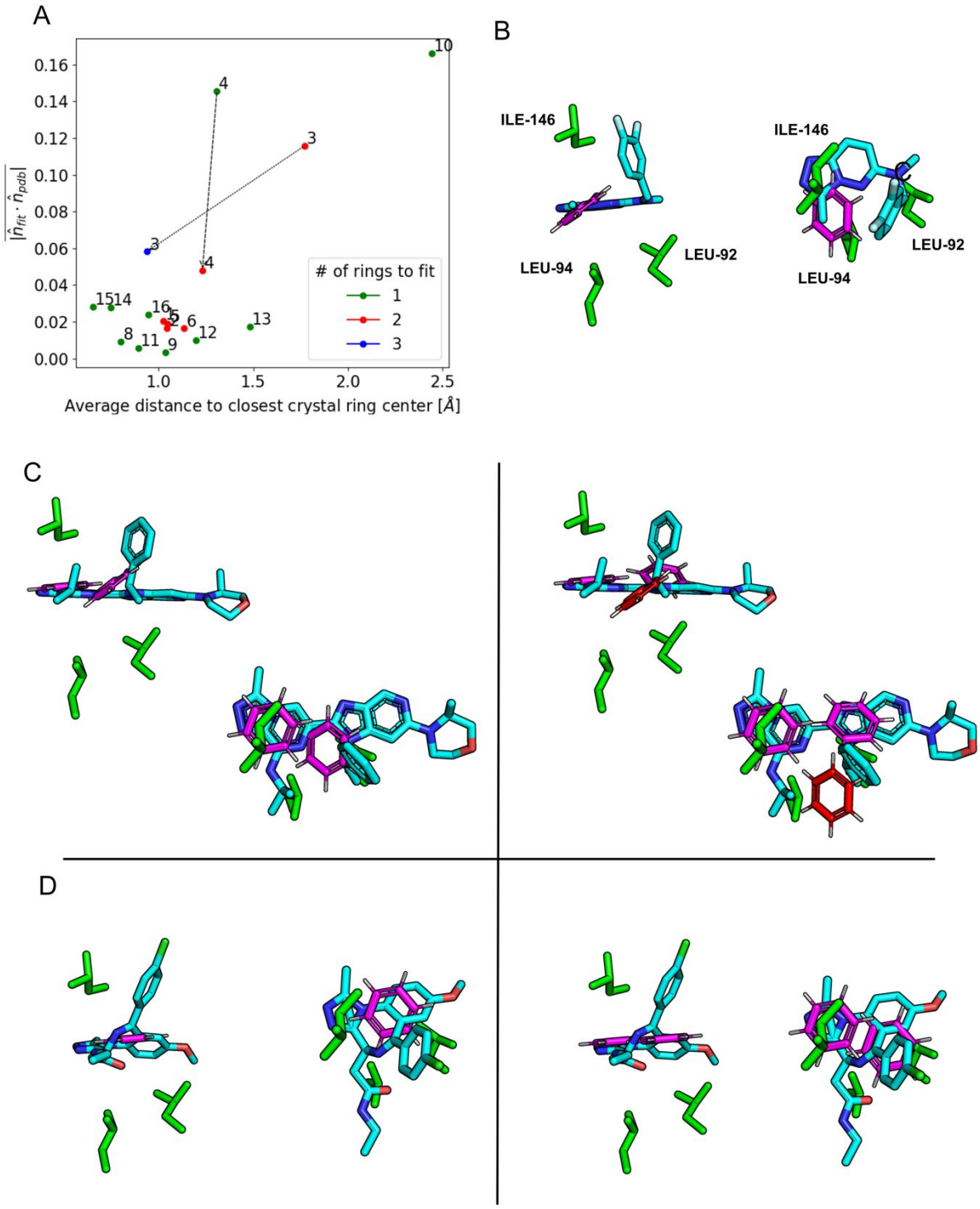


Figure S11: A: Comparison of best ring fit results for the set of ligands. A: Difference in orientation (y-axis) vs distance of ring centers (x-axis) between the best fit and the closest ring in the respective crystal structure. B: For ligand 10 we are missing shift data for Ile-Cδ which leads to a large offset for the best fit. C: For ligand 3, there is an additional ring that has to be accounted for and leads to an intermediate position when fitting only 2 rings (left). When fitting 3 rings (right), the position of the 2 main fit rings improves accordingly, while the 3rd ring covers a potential ring position of the 3rd ligand ring after change of one dihedral angle. D: For ligand 4, the presence of the non-planar 7-membered ring hinders good orientational agreement (left). This is improved by including a second ring in the fit (right).

HPLC chromatograms for compounds tested:

All samples were analyzed on an Agilent 1200 series LC system coupled with an Agilent 6140 mass spectrometer. Purity was determined via UV detection with a bandwidth of 170nm in the range from 230-400nm. LC parameters were as follows: Waters Xbridge C18 column, 2.5µm particle size, 2.1 x 20mm. Run time 2.1 minutes, flow 1ml/min, column temperature 60°C and 5µl injections. Solvent A (20mM NH4HCO3/ NH3 pH 9), solvent B (MS grade acetonitrile). Start 10% B, gradient 10% - 95% B from 0.0 - 1.5min, 95% B from 1.5 - 2.0min, gradient 95% - 10% B from 2.0 – 2.1min.

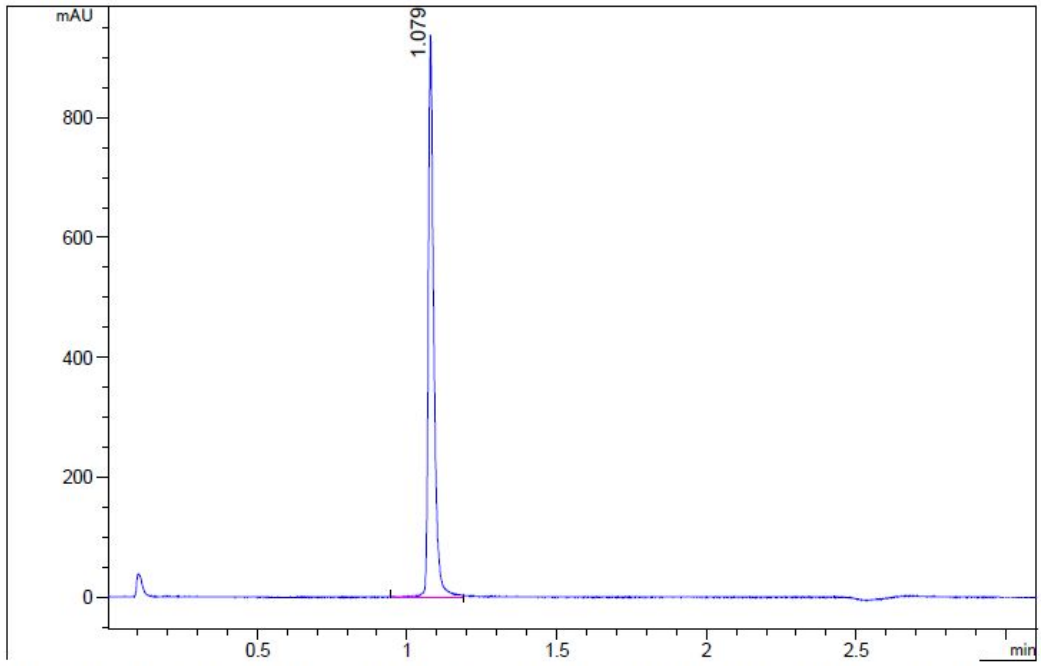


Figure S12: HPLC trace of Ligand 1

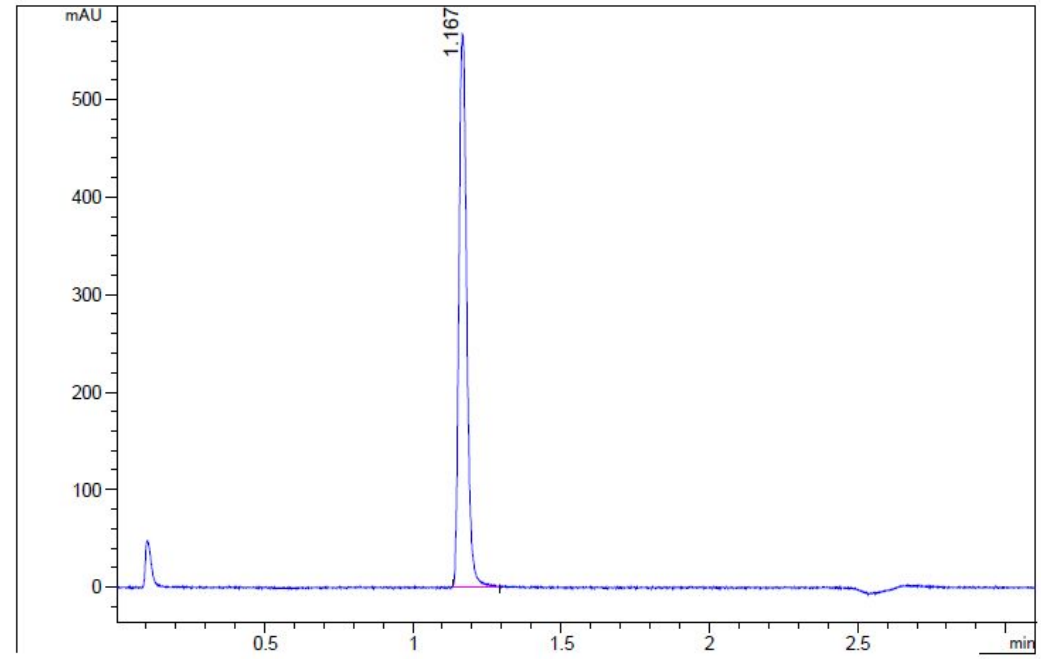


Figure S13: HPLC trace of Ligand 2

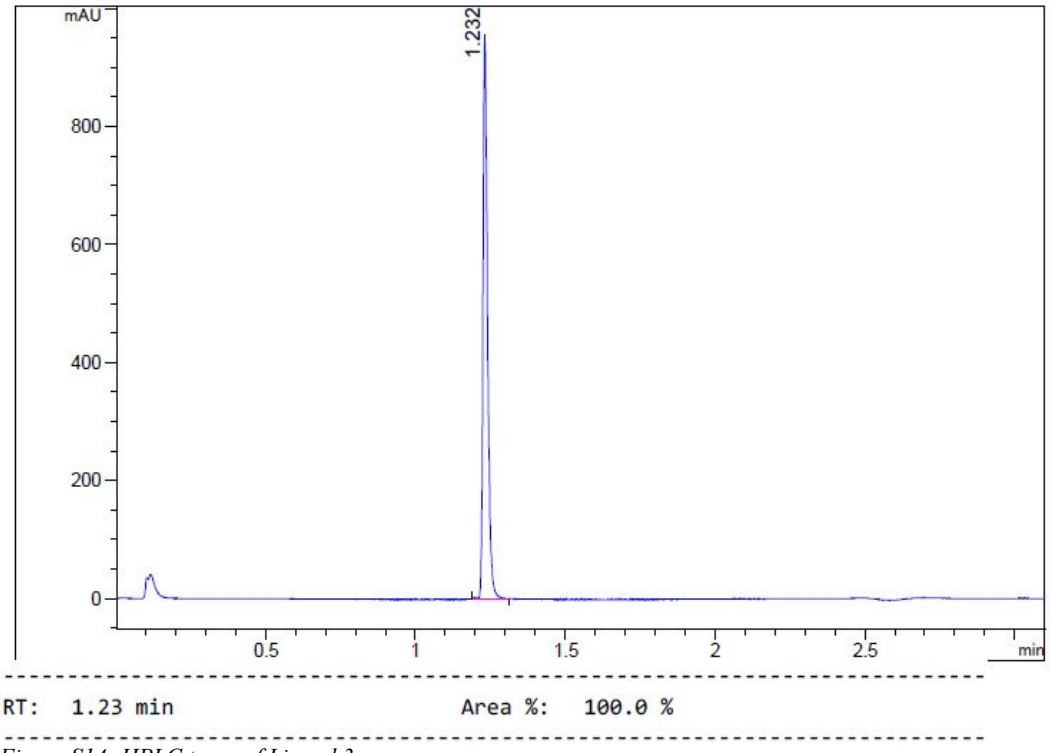
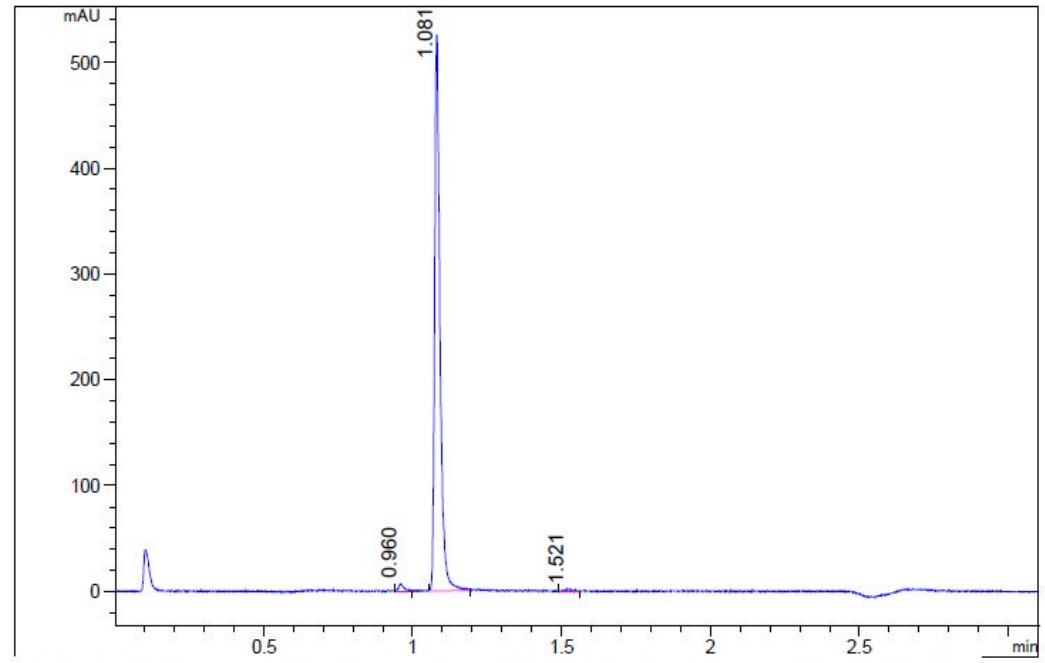
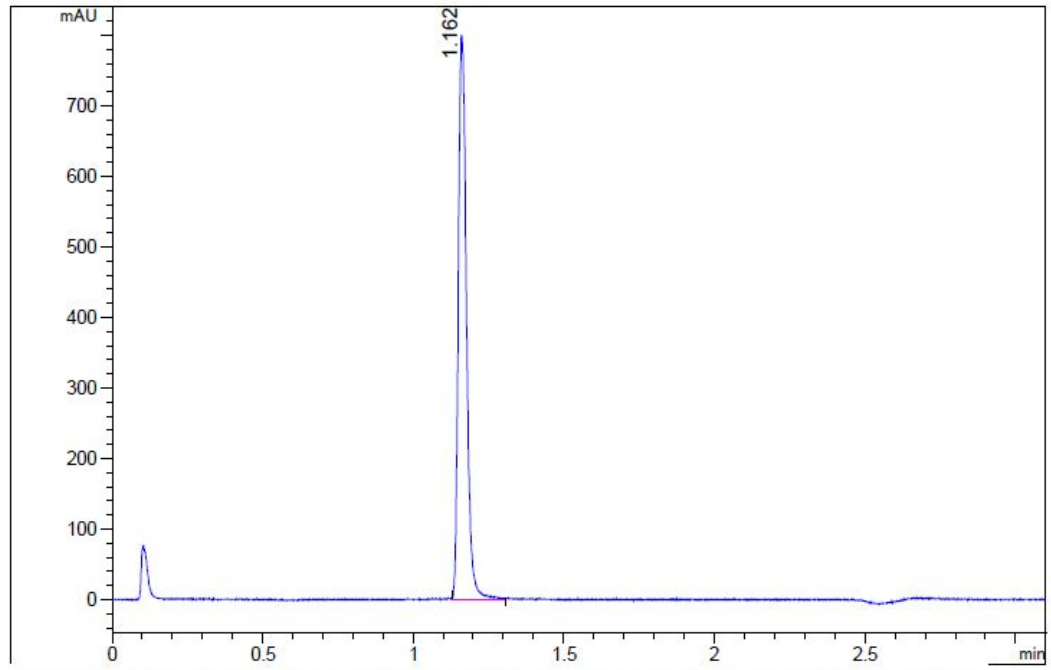


Figure S14: HPLC trace of Ligand 3



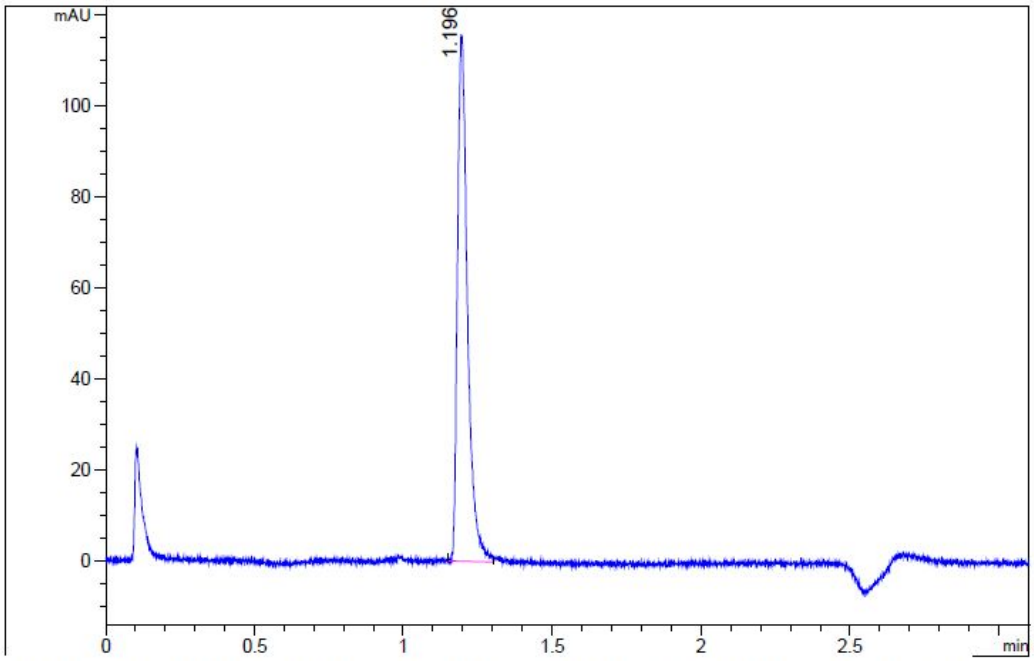
RT:	0.96 min	Area %:	1.2 %
RT:	1.08 min	Area %:	98.4 %
RT:	1.52 min	Area %:	0.4 %

Figure S15: HPLC trace of Ligand 4



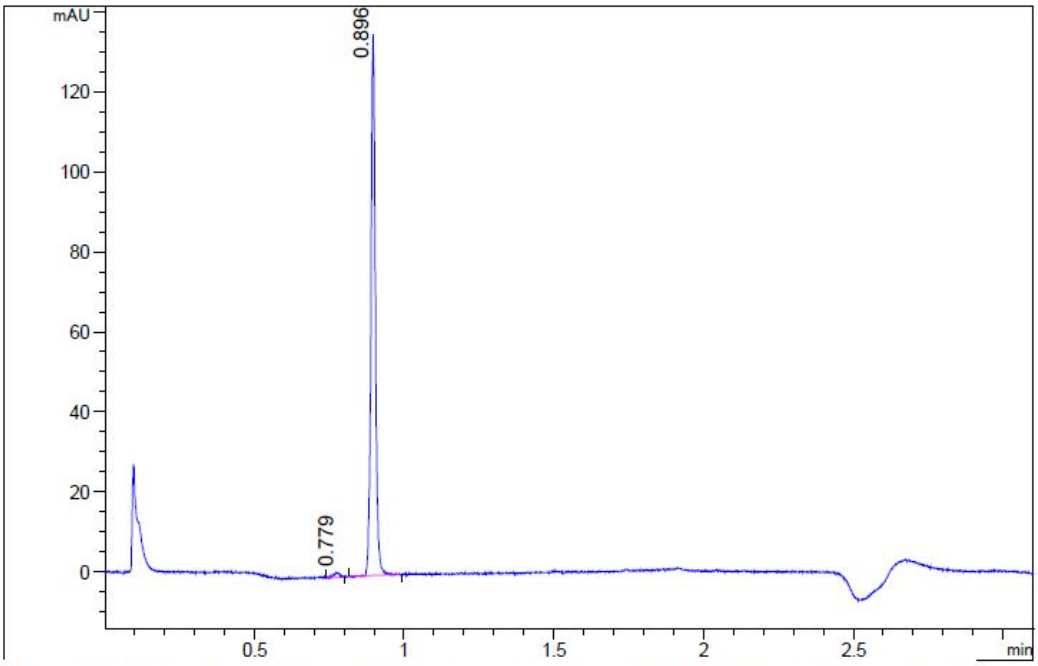
RT:	1.16 min	Area %:	100.0 %

Figure S16: HPLC trace of Ligand 5



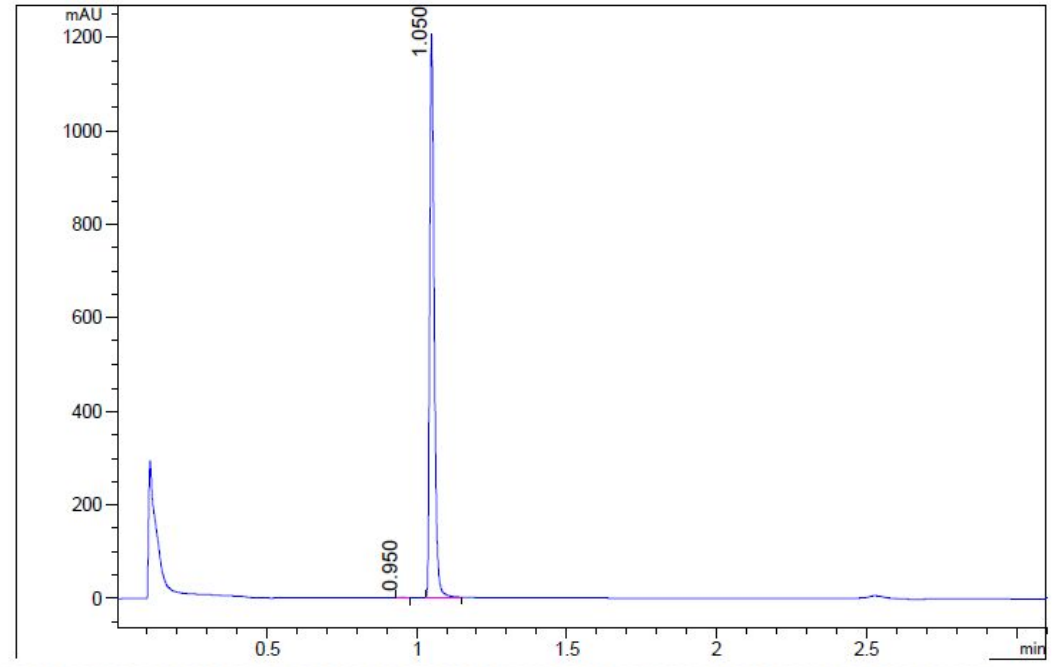
RT: 1.20 min Area %: 100.0 %

Figure S17: HPLC trace of Ligand 6



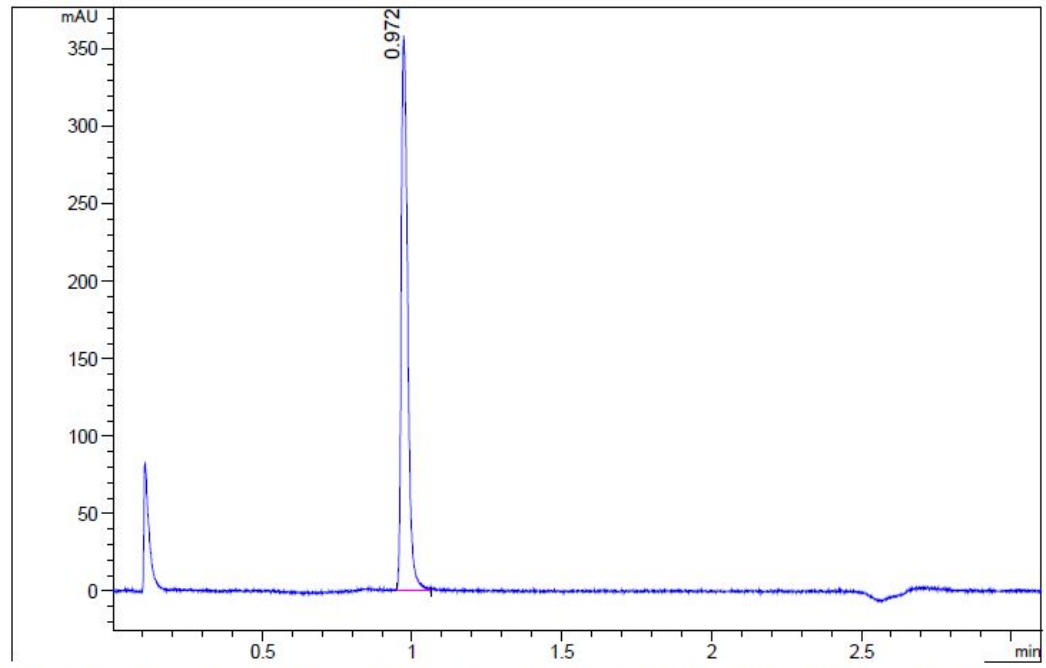
RT: 0.78 min Area %: 1.5 %
RT: 0.90 min Area %: 98.5 %

Figure S18: HPLC trace of ligand 7



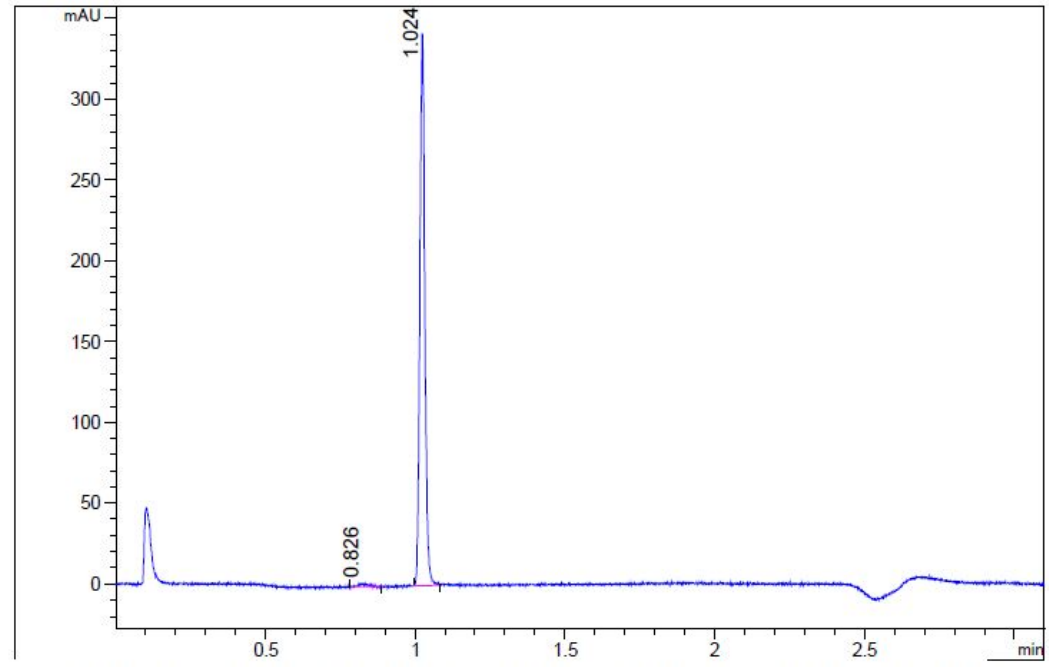
RT: 0.95 min Area %: 0.1 %
RT: 1.05 min Area %: 99.9 %

Figure S19: HPLC trace of ligand 8



RT: 0.97 min Area %: 100.0 %

Figure S20: HPLC trace of ligand 9



RT:	0.83 min	Area %:	0.9 %
RT:	1.02 min	Area %:	99.1 %

Figure S21: HPLC trace of ligand 10

Part III

DISCUSSION AND OUTLOOK

7

DISCUSSION AND OUTLOOK

Despite recent advances in deep learning models that put into question many traditional structure determination methods, NMR still remains strong as a technique that is able to answer questions that arise when one needs to look beyond fixed coordinates [80]–[82]. At the same time, in many other areas like fragment screening and drug discovery and optimization, NMR has confidently secured its role as a reliable method to provide information in a fast and iterative manner [83]–[85]. In this thesis we cover both of these topics.

7.1 PRE AND PRI BASED DETERMINATION OF A COMPACT α-SYNUCLEIN STRUCTURE

The first paper presented in chapter 3 aims to use an integrative approach combining paramagnetic NMR with cross-linking mass spectrometry to investigate the membrane bound compact conformations of α-sunyclein. It is well known, that α-synuclein adopts a compact α-helical conformation when in contact with biological membranes as well as common membrane mimics [86], [87]. We were able to determine the backbone amide proton PRE rates of the protein when bound to SDS micelles for multiple different spin tag positions. We also determined the PRI rates for all double tagged combinations of these sites. We developed a novel PRI based energy potential for the simulated annealing software Xplor-NIH and used it in combination with already available PRE potentials to calculate the most probable monomeric ensemble of membrane bound α-synuclein [42].

It is also commonly stated, that α-synuclein molecules form dimers and multimers on the surface of biological membranes [88]. To investigate this, we developed a novel isotope-edited corss-linking mass spectrometry approach. With this, we were able to determine the most probable intra- and intermolecular contact sites in multimeric α-synuclein. Using this data, we were able to refine monomeric structures into dimers and multimers.

Our proposed dimeric structures and existing experimental data both suggest the existence of 2 prototypical oligomeric topologies, namely a circular ring-like arrangement, as well as an extended twisted oligomer which exist in a dynamic equilibrium [49]. Due to the different amphiphilic character of both multimer species, they can either lead to bilayer disruption or facilitate aggregate formation [89], [90].

Common disease associated mutations are located at the multimer interfaces and lead to clear changes in the intermolecular cross-linking patterns. Thus, we propose, that these mutations change the balance between multimeric α-synuclein species and the disruption of proper regulation ultimately leads to common synucleopathies [91].

Both NMR as well as cross-linking data provide evidence that the novel drug candidate UCB0599 specifically targets oligomeric α-synuclein species and disrupts both aggregate formation as well as membrane interaction and ultimately promotes their dissociation into monomers and release into the cytosol.

7.2 CHARACTERIZATION OF METHYL-π INTERACTIONS THROUGH CHEMICAL SHIFTS

The second paper in chapter 3 is interested in the possible applications of NMR to identify and characterize methyl-π interactions during any stage of a drug discovery campaign and specifically in the process of fragment screening and hit validation. While NMR is already established as a standard technique during fragment screening to determine fragment binding through CSP measurements, the elucidation of the bound protein-ligand conformation is still traditionally carried out via crystallization studies [85]. In the highlighted article, we show that CH-π interactions are ubiquitous in protein-ligand systems and can be reliably characterized via CSP measurements. They show clear preferences in interaction geometries and the expected CSPs, as back-calculated from crystal data, match the experimental values. We then use the refined Pople ring-current model as well as the interaction geometries extracted from the PDB to determine the most probable position and orientation of the ligand aromatic centers in the complex. The fact that this is achieved using a pure in-silico structure highlights the robustness of the approach even for small changes in backbone and side chain dihedrals.

In the process we were able to identify and account for discrepancies between the calculated shifts and the measured values. We were able to extend the Pople ring-current model to include uncommon homoaromatic systems and we were able to show that residual ligand and side chain dynamics in the bound state have to be taken into account [92]. This is especially true for residues like isoleucine that posses highly mobile side chains [93]. We hope to be able to improve the approach even further in the future by incorporating, for instance representative frames of molecular dynamics simulations in the determination and refinement of the bound conformation.

As stated in the introduction, for all ring-current models, a ring specific intensity factor should be determined independently for any aromatic system under investigation. In practice however, this factor

is often neglected. In the future, we hope to be able to determine these intensity factors by QM and DFT methods for a representative set of aromatic systems which can then serve as a standard library to mix and match. On top of that, there have been multiple suggestions for additional factors that effect the induced chemical shift in CH-π interactions like local anisotropies arising from the separate ring atoms [67], [68].

In addition to methyl groups, we have also recently investigated the effect of nearby aromatic rings on aromatic protons in the protein as well as methylene and methine groups [94], [95]. We hope to be able to combine these in the future into a single concise standard approach applicable to a broad range of protein-ligand systems. With a good coverage of the standard amino acids, it will hopefully be straightforward to quickly and unanimously determine the bound ligand conformation in the context of the protein without the need of additional crystallization studies.

BIBLIOGRAPHY

[1] E. Fischer, "Einfluss der configuration auf die wirkung der enzyme," *Berichte der deutschen chemischen Gesellschaft*, vol. 27, no. 3, pp. 2985–2993, 1894.

[2] A. K. Dunker, J. D. Lawson, C. J. Brown, *et al.*, "Intrinsically disordered protein," *Journal of molecular graphics and modelling*, vol. 19, no. 1, pp. 26–59, 2001.

[3] F. H. Crick, "On protein synthesis," in *Symp Soc Exp Biol*, vol. 12, 1958, p. 8.

[4] J. S. Fruton, *Contrasts in scientific style: research groups in the chemical and biochemical sciences*. American Philosophical Society, 1990, vol. 191.

[5] E. Callaway, "The revolution will not be crystallized," *Nature*, vol. 525, no. 7568, p. 172, 2015.

[6] K. Takeuchi, K. Baskaran, and H. Arthanari, "Structure determination using solution nmr: Is it worth the effort?" *Journal of Magnetic Resonance*, vol. 306, pp. 195–201, 2019.

[7] H. M. Berman, J. Westbrook, Z. Feng, *et al.*, "The protein data bank," *Nucleic acids research*, vol. 28, no. 1, pp. 235–242, 2000.

[8] P. R. Markwick, T. Malliavin, and M. Nilges, "Structural biology by nmr: Structure, dynamics, and interactions," *PLoS computational biology*, vol. 4, no. 9, e1000168, 2008.

[9] R. M. LeBlanc and M. F. Mesleh, "A drug discovery toolbox for nuclear magnetic resonance (nmr) characterization of ligands and their targets," *Drug Discovery Today: Technologies*, vol. 37, pp. 51–60, 2020.

[10] G. Ortega, M. Pons, and O. Millet, "Protein functional dynamics in multiple timescales as studied by nmr spectroscopy," *Advances in protein chemistry and structural biology*, vol. 92, pp. 219–251, 2013.

[11] D. Khago, I. J. Fucci, and R. A. Byrd, "The role of conformational dynamics in the recognition and regulation of ubiquitination," *Molecules*, vol. 25, no. 24, p. 5933, 2020.

[12] A. S. Holehouse and B. B. Kragelund, "The molecular basis for cellular function of intrinsically disordered protein regions," *Nature Reviews Molecular Cell Biology*, vol. 25, no. 3, pp. 187–211, 2024.

[13] B. Brutscher, I. C. Felli, S. Gil-Caballero, *et al.*, "Nmr methods for the study of instrinsically disordered proteins structure, dynamics, and interactions: General overview and practical guidelines," *Intrinsically disordered proteins studied by NMR spectroscopy*, pp. 49–122, 2015.

[14] R. Konrat, "Nmr contributions to structural dynamics studies of intrinsically disordered proteins," *Journal of Magnetic Resonance*, vol. 241, pp. 74–85, 2014.

[15] D. V. Laurents, "Alphafold 2 and nmr spectroscopy: Partners to understand protein structure, dynamics and function," *Frontiers in molecular biosciences*, vol. 9, p. 906 437, 2022.

[16] G. M. Clore and J. Iwahara, "Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes," *Chemical reviews*, vol. 109, no. 9, pp. 4108–4139, 2009.

[17] G. M. Clore, "Practical aspects of paramagnetic relaxation enhancement in biological macromolecules," *Methods in enzymology*, vol. 564, pp. 485–497, 2015.

[18] C. Hartlmüller, E. Spreitzer, C. Göbl, F. Falsone, and T. Madl, "Nmr characterization of solvent accessibility and transient structure in intrinsically disordered proteins," *Journal of biomolecular NMR*, vol. 73, no. 6, pp. 305–317, 2019.

[19] C. N. Johnson and D. S. Libich, "Paramagnetic relaxation enhancement for detecting and characterizing self-associations of intrinsically disordered proteins," *Journal of visualized experiments: JoVE*, no. 175, 2021.

[20] B. Mateos, R. Konrat, R. Pierattelli, and I. C. Felli, "Nmr characterization of long-range contacts in intrinsically disordered proteins from paramagnetic relaxation enhancement in 13c direct-detection experiments," *ChemBioChem*, vol. 20, no. 3, pp. 335–339, 2019.

[21] Y. Xue, I. S. Podkorytov, D. K. Rao, *et al.*, "Paramagnetic relaxation enhancements in unfolded proteins: Theory and application to drkn sh3 domain," *Protein Science*, vol. 18, no. 7, pp. 1401–1424, 2009.

[22] X. Fan, T. Myers, B. Sukra, and G Gabrielse, "Measurement of the electron magnetic moment," *Physical review letters*, vol. 130, no. 7, p. 071 801, 2023.

[23] I. Solomon, "Relaxation processes in a system of two spins," *Physical Review*, vol. 99, no. 2, p. 559, 1955.

[24] N Bloembergen and L. Morgan, "Proton relaxation times in paramagnetic solutions. effects of electron spin relaxation," *The Journal of Chemical Physics*, vol. 34, no. 3, pp. 842–850, 1961.

[25] Q. Miao, C. Nitsche, H. Orton, *et al.*, "Paramagnetic chemical probes for studying biological macromolecules," *Chemical Reviews*, vol. 122, no. 10, pp. 9571–9642, 2022.

[26] O. H. Griffith and A. Waggoner, "Nitroxide free radicals: Spin labels for probing biomolecular structure," *Accounts of Chemical Research*, vol. 2, no. 1, pp. 17–24, 1969.

[27] F. Rodriguez-Castañeda, P. Haberz, A. Leonov, and C. Griesinger, "Paramagnetic tagging of diamagnetic proteins for solution nmr," *Magnetic Resonance in Chemistry*, vol. 44, no. S1, S10–S16, 2006.

[28] I. Bertini, C. Luchinat, G. Parigi, and E. Ravera, *NMR of paramagnetic molecules: applications to metallobiomolecules and models*. Elsevier, 2016.

[29] L. Lee and B. D. Sykes, "Use of lanthanide-induced nuclear magnetic resonance shifts for determination of protein structure in solution: Ef calcium binding site of carp parvalbumin," *Biochemistry*, vol. 22, no. 19, pp. 4366–4373, 1983.

[30] Y. W. Ebright, Y. Chen, P. S. Pendergrast, and R. H. Ebright, "Incorporation of an edta-metal complex at a rationally selected site within a protein: Application to edta-iron dna affinity cleaving with catabolite gene activator protein (cap) and cro," *Biochemistry*, vol. 31, no. 44, pp. 10 664–10 670, 1992.

[31] P. H. Keizers, A. Saragliadis, Y. Hiruma, M. Overhand, and M. Ubbink, "Design, synthesis, and evaluation of a lanthanide chelating protein probe: Clanp-5 yields predictable paramagnetic effects independent of environment," *Journal of the American Chemical Society*, vol. 130, no. 44, pp. 14 802–14 812, 2008.

[32] L. Banci, I. Bertini, K. Bren, *et al.*, "The use of pseudocontact shifts to refine solution structures of paramagnetic metalloproteins: Met80ala cyano-cytochrome c as an example," *JBIC Journal of Biological Inorganic Chemistry*, vol. 1, pp. 117–126, 1996.

[33] I. Bertini, M. B. Janik, Y.-M. Lee, C. Luchinat, and A. Rosato, "Magnetic susceptibility tensor anisotropies for a lanthanide ion series in a fixed protein matrix," *Journal of the American Chemical Society*, vol. 123, no. 18, pp. 4181–4188, 2001.

[34] R. Barbieri, I. Bertini, G. Cavallaro, *et al.*, "Paramagnetically induced residual dipolar couplings for solution structure determination of lanthanide binding proteins," *Journal of the American Chemical Society*, vol. 124, no. 19, pp. 5581–5587, 2002.

[35] T. Mu¨ntener, D. Joss, D. Ha¨ussinger, and S. Hiller, "Pseudocontact shifts in biomolecular nmr spectroscopy," *Chemical Reviews*, vol. 122, no. 10, pp. 9422–9467, 2022.

[36] J. Boisbouvier, P. Gans, M. Blackledge, B. Brutscher, and D. Marion, "Long-range structural information in nmr studies of paramagnetic molecules from electron spin- nuclear spin cross-correlated relaxation," *Journal of the American Chemical Society*, vol. 121, no. 33, pp. 7700–7701, 1999.

[37] J. Iwahara, C. Tang, and G. M. Clore, "Practical aspects of 1h transverse paramagnetic relaxation enhancement measurements on macromolecules," *Journal of Magnetic Resonance*, vol. 184, no. 2, pp. 185–195, 2007.

[38] G. M. Clore, "Exploring sparsely populated states of macromolecules by diamagnetic and paramagnetic nmr relaxation," *Protein Science*, vol. 20, no. 2, pp. 229–246, 2011.

[39] M. A. Lietzow, M. Jamin, H. J. Dyson, and P. E. Wright, "Mapping long-range contacts in a highly unfolded protein," *Journal of molecular biology*, vol. 322, no. 4, pp. 655–662, 2002.

[40] A. Kulikov and G. Likhtenstein, "The use of spin relaxation phenomena in the investigation of the structure of model and biological systems by the method of spin labels," *Advances in molecular relaxation and interaction processes*, vol. 10, no. 1, pp. 47–79, 1977.

[41] J. R. Biller, H. Elajaili, V. Meyer, *et al.*, "Electron spin–lattice relaxation mechanisms of rapidly-tumbling nitroxide radicals," *Journal of Magnetic Resonance*, vol. 236, pp. 47–56, 2013.

[42] J. Iwahara, C. D. Schwieters, and G. M. Clore, "Ensemble approach for nmr structure refinement against 1h paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule," *Journal of the American Chemical Society*, vol. 126, no. 18, pp. 5879–5896, 2004.

[43] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, and G. M. Clore, "The xplor-nih nmr molecular structure determination package," *Journal of magnetic resonance*, vol. 160, no. 1, pp. 65–73, 2003.

[44] C. D. Schwieters, J. J. Kuszewski, and G. M. Clore, "Using xplor–nih for nmr molecular structure determination," *Progress in nuclear magnetic resonance spectroscopy*, vol. 48, no. 1, pp. 47–62, 2006.

[45] D. Kurzbach, A. Vanas, A. G. Flamm, *et al.*, "Detection of correlated conformational fluctuations in intrinsically disordered proteins through paramagnetic relaxation interference," *Physical Chemistry Chemical Physics*, vol. 18, no. 8, pp. 5753–5758, 2016.

[46] D. Kurzbach, A. Beier, A. Vanas, *et al.*, "Nmr probing and visualization of correlated structural fluctuations in intrinsically disordered proteins," *Physical Chemistry Chemical Physics*, vol. 19, no. 16, pp. 10 651–10 656, 2017.

[47] G. Platzer, A. Schedlbauer, A. Chemelli, *et al.*, "The metastasis-associated extracellular matrix protein osteopontin forms transient structure in ligand interaction sites," *Biochemistry*, vol. 50, no. 27, pp. 6113–6124, 2011.

[48] D. Kurzbach, G. Platzer, T. C. Schwarz, *et al.*, "Cooperative unfolding of compact conformations of the intrinsically disordered protein osteopontin," *Biochemistry*, vol. 52, no. 31, pp. 5167–5175, 2013.

[49] T. C. Schwarz, A. Beier, K. Ledolter, *et al.*, "High-resolution structural information of membrane-bound α-synuclein provides insight into the moa of the anti-parkinson drug ucb0599," *Proceedings of the National Academy of Sciences*, vol. 120, no. 15, e2201910120, 2023.

[50] M. A. Hass and F. A. Mulder, "Contemporary nmr studies of protein electrostatics," *Annual Review of Biophysics*, vol. 44, no. 1, pp. 53–75, 2015.

[51] D. S. Wishart, B. D. Sykes, and F. M. Richards, "Relationship between nuclear magnetic resonance chemical shift and protein secondary structure," *Journal of molecular biology*, vol. 222, no. 2, pp. 311–333, 1991.

[52] A. Pardi, G. Wagner, and K. Wüthrich, "Protein conformation and proton nuclear-magnetic-resonance chemical shifts," *European journal of biochemistry*, vol. 137, no. 3, pp. 445–454, 1983.

[53] M. P. Williamson, "Secondary-structure dependent chemical shifts in proteins," *Biopolymers: Original Research on Biomolecules*, vol. 29, no. 10-11, pp. 1423–1431, 1990.

[54] K. Tamiola and F. A. Mulder, "Using nmr chemical shifts to calculate the propensity for structural order and disorder in proteins," *Biochemical Society Transactions*, vol. 40, no. 5, pp. 1014–1020, 2012.

[55] J. Pople, W. Schneider, H. Bernstein, and D. Ingram, "High-resolution nuclear magnetic resonance," *Physics Today*, vol. 13, no. 5, pp. 46–48, 1960.

[56] H. Saitô, I. Ando, and A. Ramamoorthy, "Chemical shift tensor–the heart of nmr: Insights into biological aspects of proteins," *Progress in nuclear magnetic resonance spectroscopy*, vol. 57, no. 2, pp. 181–228, 2010.

[57] F. A. Anet and D. J. O'Leary, "The shielding tensor. part i: Understanding its symmetry properties," *Concepts in Magnetic Resonance*, vol. 3, no. 4, pp. 193–214, 1991.

[58] J. C. Facelli, "Chemical shift tensors: Theory and application to molecular structural problems," *Progress in nuclear magnetic resonance spectroscopy*, vol. 58, no. 3-4, p. 176, 2010.

[59] L. Pauling, "The diamagnetic anisotropy of aromatic molecules," *The Journal of chemical physics*, vol. 4, no. 10, pp. 673–677, 1936.

[60] D. A. Case, "Calibration of ring-current effects in proteins and nucleic acids," *Journal of biomolecular NMR*, vol. 6, pp. 341–346, 1995.

[61] C. Haigh and R. Mallion, "Ring current theories in nuclear magnetic resonance," *Progress in nuclear magnetic resonance spectroscopy*, vol. 13, no. 4, pp. 303–344, 1979.

[62] J. Pople, "Proton magnetic resonance of hydrocarbons," *The Journal of Chemical Physics*, vol. 24, no. 5, pp. 1111–1111, 1956.

[63] H. Bernstein, W. G. Schneider, and J. A. Pople, "The proton magnetic resonance spectra of conjugated aromatic hydrocarbons," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 236, no. 1207, pp. 515–528, 1956.

[64] H. M. McConnell, "Theory of nuclear magnetic shielding in molecules. i. long-range dipolar shielding of protons," *The Journal of Chemical Physics*, vol. 27, no. 1, pp. 226–229, 1957.

[65] J. Waugh and R. Fessenden, "Nuclear resonance spectra of hydrocarbons: The free electron model," *Journal of the American Chemical Society*, vol. 79, no. 4, pp. 846–849, 1957.

[66] C. Johnson Jr and F. Bovey, "Calculation of nuclear magnetic resonance spectra of aromatic hydrocarbons," *The Journal of Chemical Physics*, vol. 29, no. 5, pp. 1012–1014, 1958.

[67] M. Barfield, D. M. Grant, and D. Ikenberry, "Importance of local anisotropic effects and ring currents on proton shieldings in aromatic hydrocarbons," *Journal of the American Chemical Society*, vol. 97, no. 24, pp. 6956–6961, 1975.

[68] A. Agarwal, J. A. Barnes, J. L. Fletcher, M. J. McGlinchey, and B. G. Sayer, "Ring currents, local anisotropy, and the problem of the out-of-plane protons: A reinvestigation of the nuclear magnetic resonance spectrum of [10]-paracyclophane," *Canadian Journal of Chemistry*, vol. 55, no. 13, pp. 2575–2581, 1977.

[69] F. London, "The general theory of molecular forces," *Transactions of the Faraday Society*, vol. 33, 8b–26, 1937.

[70] E. Hückel, "Quantentheoretische beiträge zum problem der aromatischen und ungesättigten verbindungen. iii," *Zeitschrift für Physik*, vol. 76, pp. 628–648, 1932.

[71] C. Haigh and R. Mallion, "New tables of 'ring current'shielding in proton magnetic resonance," *Organic Magnetic Resonance*, vol. 4, no. 2, pp. 203–228, 1972.

[72] C. N. Pace, "Conformational stability of globular proteins," *Trends in biochemical sciences*, vol. 15, no. 1, pp. 14–17, 1990.

[73] P. Andrews, D. Craik, and J. Martin, "Functional group contributions to drug-receptor interactions," *Journal of medicinal chemistry*, vol. 27, no. 12, pp. 1648–1657, 1984.

[74] F. Carver, C. Hunter, and E. Seward, "Structure–activity relationship for quantifying aromatic interactions," *Chemical Communications*, no. 7, pp. 775–776, 1998.

[75] R. F. de Freitas and M. Schapira, "A systematic analysis of atomic protein–ligand interactions in the pdb," *Medchemcomm*, vol. 8, no. 10, pp. 1970–1981, 2017.

[76] T. J. Ritchie and S. J. Macdonald, "The impact of aromatic ring count on compound developability–are too many aromatic rings a liability in drug design?" *Drug discovery today*, vol. 14, no. 21-22, pp. 1011–1020, 2009.

[77] S. Soga, H. Shirai, M. Kobori, and N. Hirayama, "Use of amino acid composition to predict ligand-binding sites," *Journal of chemical information and modeling*, vol. 47, no. 2, pp. 400–406, 2007.

[78] L. M. Salonen, M. Ellermann, and F. Diederich, "Aromatic rings in chemical and biological recognition: Energetics and structures," *Angewandte Chemie International Edition*, vol. 50, no. 21, pp. 4808–4842, 2011.

[79] A. Beier, G. Platzer, T. Ho¨furthner, *et al.*, "Probing protein–ligand methyl- π interaction geometries through chemical shift measurements of selectively labeled methyl groups," *Journal of Medicinal Chemistry*, vol. 67, no. 15, pp. 13 187–13 196, 2024.

[80] R. Krishna, J. Wang, W. Ahern, *et al.*, "Generalized biomolecular modeling and design with rosettafold all-atom," *Science*, vol. 384, no. 6693, eadl2528, 2024.

[81] J. Jumper, R. Evans, A. Pritzel, *et al.*, "Highly accurate protein structure prediction with alphafold," *nature*, vol. 596, no. 7873, pp. 583–589, 2021.

[82] M. Varadi, S. Anyango, M. Deshpande, *et al.*, "Alphafold protein structure database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models," *Nucleic acids research*, vol. 50, no. D1, pp. D439–D444, 2022.

[83] M. J. Harner, A. O. Frank, and S. W. Fesik, "Fragment-based drug discovery using nmr spectroscopy," *Journal of biomolecular NMR*, vol. 56, pp. 65–75, 2013.

[84] C. Dalvit, "Nmr methods in fragment screening: Theory and a comparison with other biophysical techniques," *Drug Discovery Today*, vol. 14, no. 21-22, pp. 1051–1057, 2009.

[85] A.-H. Emwas, K. Szczepski, B. G. Poulson, *et al.*, "Nmr as a "gold standard" method in drug design and discovery," *Molecules*, vol. 25, no. 20, p. 4597, 2020.

[86] T. Bartels, L. S. Ahlstrom, A. Leftin, *et al.*, "The n-terminus of the intrinsically disordered protein α-synuclein triggers membrane binding and helix folding," *Biophysical journal*, vol. 99, no. 7, pp. 2116–2124, 2010.

[87] T. S. Ulmer, A. Bax, N. B. Cole, and R. L. Nussbaum, "Structure and dynamics of micelle-bound human α-synuclein," *Journal of Biological Chemistry*, vol. 280, no. 10, pp. 9595–9603, 2005.

[88] J. Burré, M. Sharma, and T. C. Südhof, "α-synuclein assembles into higher-order multimers upon membrane binding to promote snare complex formation," *Proceedings of the National Academy of Sciences*, vol. 111, no. 40, E4274–E4283, 2014.

[89] J. Pan, A. Dalzini, N. K. Khadka, C. M. Aryal, and L. Song, "Lipid extraction by α-synuclein generates semi-transmembrane defects and lipoprotein nanoparticles," *ACS omega*, vol. 3, no. 8, pp. 9586–9597, 2018.

[90] G. Fusco, S. W. Chen, P. T. Williamson, *et al.*, "Structural basis of membrane disruption and cellular toxicity by α-synuclein oligomers," *Science*, vol. 358, no. 6369, pp. 1440–1443, 2017.

[91] M. H. Polymeropoulos, C. Lavedan, E. Leroy, *et al.*, "Mutation in the α-synuclein gene identified in families with parkinson's disease," *science*, vol. 276, no. 5321, pp. 2045–2047, 1997.

[92] D. Geuenich, K. Hess, F. Köhler, and R. Herges, "Anisotropy of the induced current density (acid), a general method to quantify and visualize electronic delocalization," *Chemical reviews*, vol. 105, no. 10, pp. 3758–3772, 2005.

[93] M. W. MacArthur and J. M. Thornton, "Protein side-chain conformation: A systematic variation of $\chi 1$ mean values with resolution–a consequence of multiple rotameric states?" *Acta Crystallographica Section D: Biological Crystallography*, vol. 55, no. 5, pp. 994–1004, 1999.

[94] T. Höfurthner, G. Toscano, G. Kontaxis, *et al.*, "Synthesis of a 13c-methylene-labeled isoleucine precursor as a useful tool for studying protein side-chain interactions and dynamics," *Journal of Biomolecular NMR*, vol. 78, no. 1, pp. 1–8, 2024.

[95] G. Toscano, T. Höfurthner, B. Nagl, *et al.*, "13cβ-valine and 13cγ-leucine methine labeling to probe protein ligand interaction," *ChemBioChem*, vol. 25, no. 6, e202300762, 2024.