UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Structural Studies on Flexible Small Molecules Based on NMR in Oriented Media. Methodology and Application to Natural Products

by

Pablo Trigo-Mouriño

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy

in the

Facultade de Química

Departamento de Química Orgánica

April 2013

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Facultade de Química

Departamento de Química Orgánica

Declaration of Supervision

We, Dr. Víctor Sánchez Pedregal, Profesor Titular at Universidade de Santiago de Compostela, and Dr. Armando Navarro Vázquez, Ramón y Cajal fellow at Universidade de Vigo,

INFORM

That the present thesis dissertation titled "Structural Studies on Flexible Small Molecules Based on NMR in Oriented Media. Methodology and Application to Natural Products", presented by Mr. PABLO TRIGO MOURIÑO, in partial fulfillment for the obtainment of the international Ph.D. degree, has been done under our direction in the Departamento de Química Orgánica of the Universidade de Santiago de Compostela and meets the requirements ordered by applicable regulations to be examined by the corresponding tribunal in public defense.

This report is signed for legal statement for the appropriate purposes,

Signed:

Dr. Víctor Sánchez Pedregal

_____

Signed:

Dr. Armando Navarro Vázquez

_____

Date: _____

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Facultade de Química

Departamento de Química Orgánica

Declaración de Consentimiento

El Dr. Víctor Sánchez Pedregal, profesor titular universitario en la Universidade de Santiago de Compostela, y el Dr. Armando Navarro Vázquez, investigador Ramón y Cajal en la Universidade de Vigo,

INFORMAN

Que la presente memoria titulada "Structural Studies on Flexible Small Molecules Based on NMR in Oriented Media. Methodology and Application to Natural Products", presentada por D. PABLO TRIGO MOURIÑO, para optar al grado internacional de doctor por la Universidade de Santiago de Compostela, ha sido realizada bajo su dirección en el Departamento de Química Orgánica de la Universidade de Santiago de Compostela y reúne los requisitos exigidos por la normativa aplicable para ser valorada en pública defensa por el tribunal correspondiente.

Y para que así conste a los efectos oportunos, se firma el presente informe,

Firmado:

Dr. Víctor Sánchez Pedregal

_____

Firmado:

Dr. Armando Navarro Vázquez

_____

Fecha: _____

Declaration of Authorship

I, Pablo Trigo-Mouriño, declare that this thesis titled, "Structural Studies on Flexible Small Molecules Based on NMR in Oriented Media. Methodology and Application to Natural Products" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: _____

Date: _____

Declaración de Autoría

Yo, Pablo Trigo Mouriño, declaro que esta memoria de tesis titulada "Structural Studies on Flexible Small Molecules Based on NMR in Oriented Media. Methodology and Application to Natural Products" así como el trabajo presentado en la misma es de mi autoría. Yo confirmo que:

- Este trabajo ha sido realizado mayoritariamente durante mi estancia en esta universidad con el objetivo de obtener el título de doctor.
- Si alguna parte de esta tesis ha sido previamente depositada para otra titulación diferente en esta Universidad o cualquier otra, esto se ha indicado claramente.
- El trabajo publicado por otros, se ha reconocido siempre con claridad.
- Siempre se ha citado la fuente original del trabajo realizado por otros. Con la excepción de esas citas, esta tesis es completamente mi propio trabajo.
- Toda ayuda recibida se menciona y agradece en las secciones correspondientes de la tesis.
- Cuando partes de esta tesis se han basado en trabajo realizado en colaboración con otros, he establecido claramente cuál ha sido la contribución de cada colaborador así como la mía propia.

Firmado: _____

Fecha: _____

"Do. Or do not. There's no try."

Yoda

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Abstract

Facultade de Química
Departamento de Química Orgánica

Doctor of Philosophy

by Pablo Trigo-Mouriño

This thesis describes the development and application of structural elucidation methodologies based on NMR in aligned media. Nuclear magnetic resonance is arguably the most important technique for the structural analysis of organic molecules in solution. In the last decade, Residual Dipolar Coupling (RDC) analysis emerged as a powerful tool for the determination of the three-dimensional structure of organic molecules in solution, complementing and even outperforming the approach based on the classical NMR observables such as NOE or ^3J couplings. While application of RDCs to the structural analysis of proteins developed rapidly, their use with "small" molecules (typically organic compounds and natural products with MW $<$ 1000 Da) is still scarce. From the spectroscopic point of view, two features of small molecules pose the main obstacles to the application of RDC to their analysis: the scarcity of observable couplings and the complexity stemming from conformational flexibility in solution. Besides, sample preparation with the optimal degree of alignment is still an issue for most classes of compounds.

In this thesis, all these topics are addressed and new experimental and computational advancements are presented.

i) Sample preparation. Weak alignment in water and aligning properties of polyacrylamide gels.

ii) New observables. Long-range proton–carbon RDCs.

iii) Analysis of flexible organic molecules.

Regarding the preparation of weakly oriented samples in solution, technical advancements are presented, involving the development, orientational properties investigation, and quality assessment of new alignment media both in organic solvents and in water, comprising lyotropic liquid crystals and polymeric gels.

Conformational flexibility is addressed using the single-tensor approach, that makes possible to fit experimental RDC to mixtures of conformers in fast-exchange equilibrium, facilitating the quantification of the conformational exchange. Bioactive compounds, such as salsolidine (metabolite) and lorcaserin (anti-obesity drug), are analyzed in this way.

Long-range RDC are introduced to alleviate the lack of enough one-bond C–H couplings in molecules with few protons, as is usually the case with many small molecules. A new experimental approach (SJS-HQSC) is presented for the measurement of long-range RDC, that is based on the insertion of the well-known proton-flip element in an HSQC-type experiment. The SJS-HSQC experiment provides very accurate long-range couplings in an experimentally simple and fast way.

These methodological advances have been applied to model molecules such as salsolidine, 10-epi, or the anti-obesity drug lorcaserin as well as to the structural determination of a new natural vinca alkaloid.

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Resumen

Facultade de Química
Departamento de Química Orgánica

Doctor of Philosophy

por Pablo Trigo-Mouriño

Esta tesis describe el desarrollo metodológico y su aplicación en la elucidación estructural mediante RMN en medios alineados. La resonancia magnética nuclear (RMN) es probablemente la técnica más importante para el análisis estructural de moléculas orgánicas en disolución. En la última década, el análisis de acoplamientos dipolares residuales (RDC) se ha convertido en una poderosa herramienta para la determinación de la estructura tridimensional de moléculas orgánicas en disolución, complementando los observables clásicos de RMN como NOE o acoplamientos ^3J. Mientras que la aplicación de los RDCs para el análisis estructural de proteínas se desarrolló rápidamente, su uso con moléculas "pequeñas" (normalmente compuestos orgánicos y productos naturales con MM $<$ 1000 Da) es todavía escaso. Desde el punto de vista espectroscópico, dos son los principales obstáculos para la aplicación de RDC para el análisis de moléculas pequeñas: el escaso número de acoplamientos observables y la complejidad derivada de la flexibilidad conformacional en disolución. Además, la preparación de muestras con el grado óptimo de alineamiento sigue siendo un problema para la mayoría de las clases de compuestos.

En esta tesis se tratan estos temas y se presentan nuevos avances experimentales y computacionales.

i) Preparación de la muestra. Alineamiento débil en medio acuoso y estudio de las propiedades de alineamiento de geles de poliacrilamida.

ii) Nuevos observables. RDCs de largo alcance protón–carbono.

iii) Análisis de moléculas orgánicas flexibles.

En cuanto a la preparación de muestras débilmente orientadas en disolución, los avances técnicos presentados abarcan el desarrollo, la investigación de las propiedades de orientación, y la evaluación de la calidad de los nuevos medios de alineamiento, tanto en disolventes orgánicos como en agua, incluyendo cristales líquidos liotrópicos y geles poliméricos.

La flexibilidad conformacional se aborda utilizando el enfoque de tensor único, que hace posible ajustar RDC experimentales para mezclas de confórmeros en equilibro de intercambio rápido, facilitando la cuantificación del equilibrio conformacional. De este modo se analizan compuestos bioactivos, como salsolidina (metabolito) o lorcaserina (medicamento contra la obesidad).

Los RDC de largo alcance se introducen para paliar la falta de un número suficiente de acoplamientos a un enlace C–H en moléculas con pocos protones, como suele ser el caso de muchas moléculas pequeñas. Se presenta un nuevo enfoque experimental (SJS-HQSC) para la medida de RDC de largo alcance, que se basa en la inserción del conocido elemento *proton-flip* en un experimento de tipo HSQC. El SJS-HSQC proporciona acoplamientos de largo alcance con mucha precisión mediante un procedimiento experimental sencillo y rápido.

Estos avances metodológicos se han aplicado a moléculas modelo, como salsolidina o 10-epi, al fármaco anti-obesidad lorcaserina, así como a la determinación estructural de un nuevo alcaloide vinca de origen natural.

Acknowledgements

This work would not have been possible without the support of many people, to whom I want to express my deepest gratitude:

My advisors, Dr. Víctor Sánchez Pedregal and Dr. Armando Navarro Vázquez, for giving me the opportunity to work on these very interesting projects and for their continuous encouragement, advice and guidance.

Dr. Roberto R. Gil for teaching me so much about the handling of the spectrometers and the structural analysis of natural products, his contagious enthusiasm and his invaluable support.

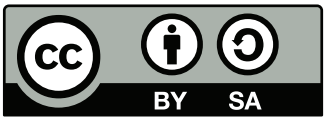
Dr. Manuel Marcos and Dr. Sonia Escudero for their invaluable help with the NMR spectrometers hosted by Universidade de Vigo.

Our collaborators Prof. Dr. Christian Griesinger, Dr. Ad Bax, Dr. Roberto R. Gil, Dr. Helena Maruenda, Dr. Maria del Carmen de la Fuente, for their inspiration and fruitful cooperations.

My colleagues at the departments of Chemistry at Universidade de Santiago de Compostela and Universidade de Vigo, NMR users at the Mellon Institute, my colleagues at the Department of NMR-based Structural Biology at MPI-BPC, specially the guys in the *cubicle*, and all other colleagues I found in this way so far, for their help, advice, friendship and fun.

Finally, I thank my parents, brother and close friends for their interest in my work and their moral support.

Made in LaTeX2ε



Contents

List of Figures

List of Tables

Abbreviations

AIBN	2,2′-azobis(2-methylpropionitrile)
AMPS	2-(acrylamido)-2-methyl-propanesulfonic acid
BIS	*N,N′*-methylenebis(acrylamide)
DBP	Dibenzoylperoxide
DFT	Density functional theory
DMAA	*N,N*-dimethylacrylamide
DVB	Divinylbenzene
EGDMA	Ethylene glycol dimethacrylate
GIAO	Gauge-independent atomic orbital
MMA	Methylmethacrylate
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
PAA	Polyacrylamide
PAN	Poly(acrylonitryle)
PDMS	Poly(dimethylsiloxane)
PMMA	Poly(methylmethacrylate)
PS	Polystyrene
PVac	Poly(vinyl acetate)
RDC	Residual dipolar coupling
SAG	strained aligning gel
SCF	Self-consistent field
SVD	Singular value decomposition
TD-DFT	Time-dependent DFT
TEMED	*N,N,N′,N′*-tetramethylethylenediamine
V70	2,2′-azobis(2,4-dimethyl-4-methoxyvaleronitrile)

Physical Constants

Gyromagnetic ratio of ^{1}H	γ_H	=	$26.752\,2208 \times 10^7$ rad s^{-1} T^{-1} (defined)
Gyromagnetic ratio of ^{13}C	γ_C	=	$6.728\,2860 \times 10^7$ rad s^{-1} T^{-1} (defined)
Permeability of vacuum	μ_0	=	$4\pi \times 10^7$ H m^{-1} (defined)
Dirac's constant	$\hbar$	=	$1.054\,457\,1726 \times 10^{34}$ J s (exact)
Boltzmann constant	k	=	$1.380\,6488 \times 10^{23}$ J K^{-1} (exact)

"We have thought of various investigations in which this effect can be used fruitfully"[1]

Chapter 1

Introduction

This thesis describes the development and application of structural elucidation methodologies for small molecules based on NMR in weakly aligning media. From the initial discovery of chemical shift by Knight,[3] Proctor and Yu,[4] and Dickinson,[5] followed by the ^{1}H chemical shift identification and the first high-resolution spectra[2] (Figure 1.1), liquid-state NMR has become an extraordinarily powerful tool for the structural elucidation of organic compounds, either of natural or synthetic origin. Furthermore, from the pioneering work by Jeener and Ernst,[6] a plethora of 2D correlated experiments have made possible the determination of the structure of very complex organic compounds and biomolecules.[7,8]

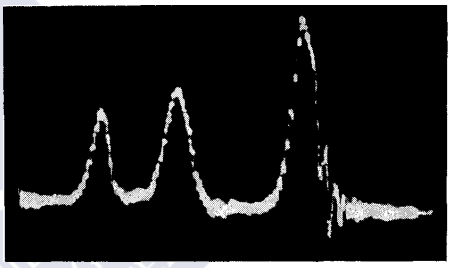


FIGURE 1.1: First, high-resolution NMR spectra of ethyl alcohol. Taken from[2].

NMR makes possible to determine not only the chemical constitution but also the relative configuration and conformation through the use of scalar couplings[9–12] and NOE experiments.[13–15] Most of applications until recent years involved the use of NOE correlations and $^{3}J_{\mathrm{HH}}$ vicinal couplings analysis. Even when this "classic NMR approach" has demonstrated useful for a wide variety of molecules, it relies on local (NOE) and sequential ($^{3}J_{\mathrm{HH}}$) information. This local character of the information furnished by NOE and $^{3}J_{\mathrm{HH}}$ couplings might hamper the structural analysis in cases in which the molecule lacks protons or has magnetically inactive linkers.

In the case of a molecule with scarce number of protons, the main drawback is the loss of the sequential $^{3}J_{\mathrm{HH}}$ information, that needs the presence of at least one proton on each atom of the chain under study. Similarly, NOE structural analysis fails if there are

not enough distance restraints available (i.e. not enough protons close to each other). Magnetically inactive linkers deprive of information about the relative orientation of the linked subunits as they interrupts $^3J_{HH}$ sequential information.

Moreover, molecular flexibility causes different averaging on NOE and $^3J_{HH}$ couplings, making difficult the combined analysis. This may hamper stereospecific assignment, which is sometimes impossible without assuming one preferential conformational state.

In the general case, and more importantly in the examples presented in this thesis dissertation, Residual Dipolar Couplings (RDCs) provide more information for solving the structural problem under study. A decisive feature is the non-local character of RDC-based restraints —the 3D angle of a vector interconnecting two interacting atoms— contrary to "classical NMR observables". Accordingly, they have been employed for solving configurational, conformational and even constitutional problems in a wide number of molecules.

1.1 Nuclear Magnetic Resonance

Historically, NMR spectroscopy has been classified in two extreme regimes: liquid-state NMR and solid-state NMR. The fundamental difference between the two regimes resides in whether the anisotropic part of the nuclear spin Hamiltonians terms, such as dipolar coupling, J coupling, chemical shift, and quadrupolar coupling is visible in the spectra. Rapid isotropic motions of molecules contribute to the averaging of the anisotropic parts of the nuclear spin Hamiltonians in liquid-state. This results in a high-resolution NMR spectrum that consists of isotropic chemical shifts and J couplings, which are usually straightforward to interpret, whereas dipolar couplings are completely averaged out. On the contrary, solid-state NMR also reveals these anisotropic parts of the nuclear spin Hamiltonians. Usually, the solid-state spectrum is dominated by these anisotropic parts and contains information-rich but hard-to-interpret broad lines. Commonly, Magic Angle spinning (MAS) and pulse techniques are employed to tailor the spectral information and make its interpretation easier.

Oriented media, such as liquid crystals or strained polymer gels, give the advantages of both regimes. Rapid motions of molecules average out *most* of the anisotropic interactions and result in spectral lines nearly as narrow as conventional high-resolution liquid-state NMR. In addition, since the molecular motions in these systems are not completely isotropic, the resulting spectra still contain information about the

anisotropic parts of the nuclear spin Hamiltonians. For example, the dipolar couplings, which are used extensively in this thesis dissertation, scale down —but do not vanish— in weak alignment media. The information about molecular geometry can still be obtained, often much more easily, as the Residual part of Dipolar Couplings (the so called RDCs) can be measured as an addition on top of *J*-coupling splitting in the spectra.

1.2 NMR Hamiltonians

A short theoretical introduction of the NMR parameters used along this thesis dissertation is presented in the following. For the derivation of the Hamiltonians and other parameters, the following books were used and the reader can refer to them for a deeper and extensive review on the theoretical (and practical) aspects of NMR spectroscopy: Ernst,[16] Cavanagh,[17], Levitt[18] Keeler,[19] and Abragam.[20] Additionally, a remarkable series of publications following the derivation of NMR Hamiltonians have been published by Smith *et al.*,[21–24] and a very illustrative review comprehending second-rank tensor rotations in NMR was recently published by Mueller.[25] In the particular case of RDCs, the comprehensive derivation of the key equations has been published by Glaser and co-workers.[26] Additionally, a good analytical description of RDCs and frame transformations can be found in the review by Blackledge.[27]

In NMR spectroscopy, we need to consider two different components of the spin Hamiltonian: the external part of interaction $\mathscr{H}_{ex}$, which describes the interaction between the spin(s) and the external static magnetic field, and the internal part $\mathscr{H}_{int}$, which describes intrinsic interactions between the spins

$$\mathscr{H} = \mathscr{H}_{ex} + \mathscr{H}_{int} \ . \tag{1.1}$$

The external Hamiltonian arises from the interaction of the spin angular momentum with the static field $\mathbf{B}_0$. In classical physics, when a magnetic dipole is placed inside a magnetic field, its potential energy is given by

$$U = -\vec{\mu} \cdot \mathbf{B}, \tag{1.2}$$

where $\vec{\mu}$ denotes a classical magnetic dipole moment. Likewise, when a nuclear spin is placed inside a magnetic field, its nuclear spin Hamiltonian depends on both its own nuclear magnetic dipole moment and the external magnetic field it experiences. The Hamiltonian operator ($\mathscr{H}$) corresponds to the total energy in classical mechanics, determines the time evolution of a system and describes the corresponding interactions

in a quantified manner. The nuclear dipole moment of a given nucleus is given by

$$\vec{\boldsymbol{\mu}} = \gamma\hbar\mathbf{I} \ , \tag{1.3}$$

in which γ is the gyromagnetic ratio of the nucleus and $\mathbf{I}$ is a dimensionless spin angular momentum operator. This finally gives the nuclear spin Hamiltonian in angular frequency (SI) units, which corresponds to the Zeeman Hamiltonian ($\mathscr{H}_{Zeeman}$) and is essential for the observation of NMR signals, as

$$\mathscr{H}_{Zeeman} = -\gamma\mathbf{B}_0\mathbf{I}_z \ , \tag{1.4}$$

where $\mathbf{I}_z$ is the spin angular momentum operator along the z axis.

Additionally, the external Hamiltonian is perturbed in the presence of oscillating radio frequency fields $\mathbf{B}_{RF}$ that are applied to the sample in pulsed NMR spectroscopy

$$\begin{aligned}\mathscr{H}_{RF} &= \omega_1 \left[\mathbf{I}_x \cos\left(\omega_{RF}t + \phi\right) + \mathbf{I}_y \sin\left(\omega_{RF}t + \phi\right)\right] \\ &= -\gamma\mathbf{B}_{RF} \left[\mathbf{I}_x \cos\left(\omega_{RF}t + \phi\right) + \mathbf{I}_y \sin\left(\omega_{RF}t + \phi\right)\right] \ ,\end{aligned} \tag{1.5}$$

where $\mathbf{I}_x$, $\mathbf{I}_y$, and $\mathbf{I}_z$ are the spin angular momentum operators along the x, y, and z axes, respectively, ω_1 is the Larmor frequency of the involved nucleus, ω_{RF} is the applied external frequency, and ϕ is the phase of the RF pulse.

The internal spin Hamiltonian accounts for the internal time-dependent elements, including contributions to relaxation. The internal interactions comprise the chemical shielding ($\mathscr{H}_{CS}$), the J-coupling ($\mathscr{H}_J$), the dipolar coupling ($\mathscr{H}_D$), and the quadrupolar coupling ($\mathscr{H}_Q$) [1]

$$\mathscr{H}_{int} = \mathscr{H}_{CS} + \mathscr{H}_J + \mathscr{H}_D + \mathscr{H}_Q \ . \tag{1.6}$$

1.2.1 Chemical Shielding

The chemical shift is the most evident and important parameter in NMR spectroscopy for the characterization of atoms in a molecule. Chemical shift differences between atoms of the same element with different electronic environments arise from the modification of the Zeeman interactions (the interactions with the static field) by the chemical shielding, thus changing the resonance frequency of the nucleus.

[1] The quadrupolar interaction term ($\mathscr{H}_Q$) only exists in the case of spins $I > \frac{1}{2}$.

The Hamiltonian of the chemical shielding can be expressed in terms of a second rank tensor $\hat{\sigma}$ [28,29]

$$\mathscr{H}_{CS} = \hbar\gamma \sum^{spins} \mathbf{I}_I \cdot \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \cdot \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix} \quad . \tag{1.7}$$

Considering only the time-independent part of the $\mathscr{H}_{CS}$, i.e. neglecting the CSA (anisotropic component) relaxation in liquid state samples, we retain only the isotropic scalar component (CSI), which is rotationally invariant and can not produce relaxation

$$\begin{aligned} \mathscr{H}_{CS}^{(\mathrm{LAB})} &= \hbar B_0 \gamma \sum^{spins} \frac{1}{3} \mathrm{Tr}\{\hat{\sigma}\} \mathbf{I}_z \\ &= \hbar B_0 \gamma \sum^{spins} \sigma_0 \mathbf{I}_z \ , \end{aligned} \tag{1.8}$$

where $\mathbf{I}_z$ is the nuclear spin angular momentum operator along the z axis. As this is derived in the LAB frame, the z axis is collinear with the $\mathbf{B}_0$ static field. This is the Hamiltonian that applies in isotropic liquids.

In general, the chemical shielding tensor can be decomposed into independent symmetric and antisymmetric parts, of which only the symmetric one contributes to the chemical shift. In its principal axis system (PAS), only the diagonal elements have non-zero values and these three elements (σ_{xx}, σ_{yy}, σ_{zz}) are the parameters defining the isotropic shielding constant σ_0, the chemical shift anisotropy $\Delta\sigma$, and the chemical shift biaxiality η [18]

$$\sigma_0 = \frac{\mathrm{Tr}\{\sigma\}}{3} = \frac{\sigma_{xx} + \sigma_{yy} + \sigma_{zz}}{3} \ , \tag{1.9}$$

$$\Delta\sigma = \sigma_{zz} - \sigma_0 \ , \tag{1.10}$$

$$\eta = \frac{\sigma_{yy} - \sigma_{xx}}{\sigma_{zz} - \sigma_0} \ . \tag{1.11}$$

The CSA tensor is represented by an ellipsoid, which adopts different dimensions in the three possible directions depending on the value of these parameters

$\Delta\sigma = 0$; isotropic, spherical tensor,

$\Delta\sigma \neq$; $\eta = 0$; uniaxial cigar-shaped tensor,

$\Delta\sigma \neq$; $\eta \neq 0$; biaxial flattened ellipsoid-shaped tensor.

The trace of a tensor remains invariant when the tensor is rotated to different axis frames. Because of this property, the isotropic shielding σ_0 is constant. In contrast, the anisotropic part of the chemical shift depends on the orientation of the molecule in the

magnetic field, which can be expressed in terms of the polar angles θ and ϕ of $\mathbf{B}_0$ in the Principal Axis system (PAS) and hence decomposed into an axial part $\left(\cos^2\theta - 1\right)$ and a rhombic part $\left(\sin^2\theta\ \cos 2\phi\right)$.

In weakly aligned media, statistically some of the molecule orientations are slightly preferred, thus causing a small portion of the chemical shift anisotropy from this orientation to contribute to the observed chemical shift. The relationship between this so-called *residual chemical shift anisotropy* (δ_{CSA}) and the orientational properties of molecules can be expressed as

$$\delta_{CSA} = \sum_{k=\{x,y,z\}} A_{kl}\delta_{kl} \ , \tag{1.12}$$

in which $\hat{\mathbf{A}}$ is the *order matrix* or *alignment tensor*. This order matrix is identical to the one that is used in the Residual Dipolar Coupling (RDC) analysis and the reader is referred to that section (Section 1.4).

1.2.2 Scalar coupling *J*

The *J*-coupling arises from an indirect interaction mediated by the electronic cloud of the bonds between two nuclei. It provides local structural information of key importance such as dihedral angles.[11] The Hamiltonian of the *J* coupling between spins *I* and *S* is expressed as[16]

$$\mathscr{H}_J = 2\pi \sum_{I<S} \mathbf{I}_I \mathbf{J}_{IS} \mathbf{I}_S \ , \tag{1.13}$$

where $\mathbf{I}_I$ and $\mathbf{I}_S$ are the spin angular momentum operators of the *I* and *S* spins, and $\mathbf{J}_{IS}$ is the indirect spin-spin scalar coupling tensor, which can be separated into isotropic and anisotropic parts. In secular form

$$\mathscr{H}_J^{iso} = 2\pi \sum_{I<S} J_{IS} \mathbf{I}_I \mathbf{I}_S \ , \text{with} \tag{1.14}$$

$$J_{IS} = \frac{J_{xx} + J_{yy} + J_{zz}}{3} \ ;$$

$$\mathscr{H}_J^{aniso} = 2\pi \sum_{I<S} \mathbf{I}_I \mathbf{J}_{IS}^{aniso} \mathbf{I}_S \ . \tag{1.15}$$

The isotropic part gives rise to the scalar coupling between the involved spins and can be directly measured from the NMR spectrum, while the anisotropic part of the *J* coupling $\mathscr{H}_J^{aniso}$ can be neglected in high-field liquid-state NMR, given that the anisotropy associated to ^{13}C–^{1}H or ^{1}H–^{1}H *J* couplings is usually low and its impact on weakly oriented nuclei can be safely ignored.[30] This turns into an advantage for the measurement of the residual dipolar couplings in aligned media, as there is no need to

disentangle the anisotropic part of the scalar *J*-coupling from the measured dipolar coupling.[30]

Regarding the sign of the *J* coupling, when the bonded spins have the same sign of gyromagnetic ratio, their *J* coupling is positive, while with opposite gyromagnetic ratio sign, *J* is negative. In the case of long-range *J* couplings, the sign can be either positive or negative and depends on the molecular geometry and other factors.

Fortunately, for most coupled hetereonuclear and many homonuclear spin-spin pairs, the resonance (Larmor) frequency difference $|\omega_I - \omega_S|$ fulfills the following criterion[19]

$$\sin 2\theta = \frac{J_{IS}}{\sqrt{(\omega_I - \omega_S)^2 + J_{IS}^2}} \approx 0 \ , \tag{1.16}$$

meaning that the spins are in the so-called *weak coupling regime*. The angle θ is defined as the *strong coupling parameter*. Importantly, under weak coupling conditions, i.e. $2\pi|J_{IS}| \ll |\omega_I - \omega_S|$, the non-secular components vanish and the scalar coupling Hamiltonian simplifies to

$$\mathscr{H}_J = 2\pi \sum_{I<S} J_{IS} \mathbf{I}_{Iz} \mathbf{I}_{Sz} \ , \tag{1.17}$$

in which $\mathbf{I}_{Iz}$ and $\mathbf{I}_{Sz}$ are the spin angular momentum operators of the spins I and S along the z axis, and J_{IS} is the scalar coupling constant.

However, many coupled spin systems do not fulfill this condition. In such *strongly-coupled* systems, the intensities and positions of the signals in spectra are affected, leading to multiplet distortions. In such situations, numerical simulations of the spin system furnish the actual *J* couplings and frequencies.[31]

Three-bond homo- and hetero-nuclear couplings are especially sensitive to the dihedral angles formed by the bonds connecting the coupled spins. The relationship between dihedral angle and ^{1}H–^{1}H $^3J_{\mathrm{HH}}$ coupling can be approximated by a three-term Fourier series

$$^3J_{\mathrm{HH}} = A\cos 2\theta + B\cos\theta + C \ . \tag{1.18}$$

The terms in the series were first parametrized by Karplus to fit experimental couplings.[9–11] However, attachment of electronegative atoms to the carbons in the coupling pathway also influences the value of the *J*-coupling.[12] Several equations were proposed in the past to take into account this effect, the most popular being the one proposed by Altona and co-workers.[32]

1.2.3 Dipolar Coupling

Dipolar coupling arises from the direct through-space interaction between two nuclear spin moments, $\mathbf{I}_I$ and $\mathbf{I}_S$, and is the most dominant interaction in solid-state NMR of $I = \frac{1}{2}$ nuclei, with a magnitude ranging in the kHz. On the contrary, dipolar couplings completely average out in liquids. An intermediate situation occurs in oriented liquids (e.g. liquid crystals), where the size of the dipolar coupling can be scaled down to the order of J couplings or less. This so-called Residual Dipolar Coupling retains the structural information relative to the orientation of the internuclear vector $\mathbf{r}_{IS}$ with respect to a molecule-fixed reference frame (Molecular Axis Frame, MAF). Dipolar couplings have proven to be a very useful restraint for the study of structure and dynamics of biological and organic molecules.[27,33,34] Residual Dipolar Couplings are the key source of structural information for the determination of configuration and conformation of organic molecules along this PhD thesis dissertation. The theoretical aspects of dipolar couplings will be discussed in the following.

The Hamiltonian of the dipolar coupling has the form[16]

$$\begin{aligned}\mathscr{H}_D &= \sum_{I<S} d_{IS}\mathbf{I}_I\mathbf{D}_{IS}\mathbf{I}_S, \text{ explicitly} \\ &= \sum_{I<S} d_{IS}\begin{pmatrix} I_{Ix} & I_{Iy} & I_{Iz}\end{pmatrix}\cdot\begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz}\end{pmatrix}\cdot\begin{pmatrix} I_{Sx} \\ I_{Sy} \\ I_{Sz}\end{pmatrix} \\ &= \sum_{I<S} d_{IS}\left\{\mathbf{I}_I\mathbf{I}_S - 3\frac{1}{r_{IS}^2}\left(\mathbf{I}_I\mathbf{r}_{IS}\right)\left(\mathbf{I}_S\mathbf{r}_{IS}\right)\right\}\ ,\end{aligned} \tag{1.19}$$

with $d_{IS} = \mu_0\gamma_I\gamma_S\hbar/\left(4\pi r_{IS}^3\right)$; where γ_I and γ_S are the gyromagnetic ratios of the interacting nuclei I and S, $\mathbf{r}_{IS}$ is the internuclear vector, and μ_0 is the permeability of vacuum. The internuclear unit vector $\mathbf{r}_{IS}/|\mathbf{r}_{IS}|$ can be expressed in polar coordinates (θ_{IS} , ϕ_{IS}, defined in Figure 1.2), leading to a equation in terms of irreducible spherical tensor operators

$$\mathscr{H}_D = \sum_{I<S}\sum_{q=-2}^{2} F_{IS}^{(q)} A_{IS}^{(q)}\ , \tag{1.20}$$

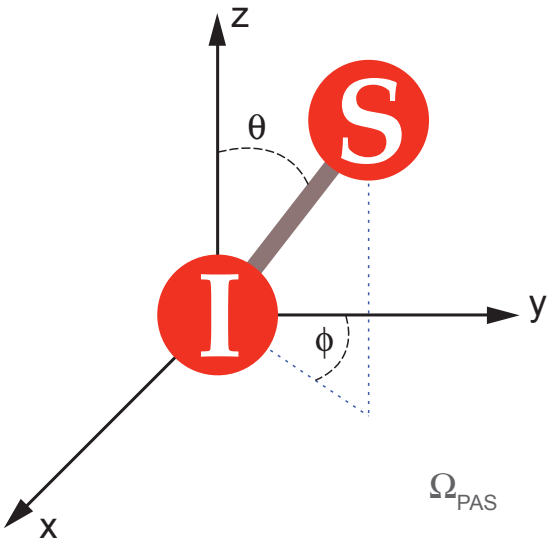


FIGURE 1.2: Definition of the $(\theta_{IS}\ ,\phi_{IS})$ polar angles used for dipolar coupling definition (see text). The angles are expressed in the PRINCIPAL AXIS SYSTEM (Ω_{PAS}).

where $F_{IS}^{(q)}$ describe the orientation of the internuclear vector and $A_{IS}^{(q)}$ contain the $\mathbf{I}_I$, $\mathbf{I}_S$ spin operators, which are defined as follows

$$
\begin{aligned}
A_{IS}^{(0)} &= d_{IS}\left\{I_{Iz}I_{Sz}-\frac{1}{4}\left(I_I^+I_S^-+I_I^-I_S^+\right)\right\}\ , & F_{IS}^{(0)} &= 1-3\cos^2\theta_{IS}\ ,\\
A_{IS}^{(1)} &= -\frac{3}{2}\,d_{IS}\left(I_{Iz}I_S^+ + I_I^+I_{Sz}\right)\ , & F_{IS}^{(1)} &= \sin\theta_{IS}\cos\theta_{IS}\exp\left\{-i\phi_{IS}\right\}\ ,\\
A_{IS}^{(-1)} &= -\frac{3}{2}\,d_{IS}\left(I_{Iz}I_S^- + I_I^-I_{Sz}\right)\ , & F_{IS}^{(-1)} &= \sin\theta_{IS}\cos\theta_{IS}\exp\left\{+i\phi_{IS}\right\}\ ,\\
A_{IS}^{(2)} &= -\frac{3}{4}\,d_{IS}I_I^+I_S^+\ , & F_{IS}^{(2)} &= \sin^2\theta_{IS}\exp\left\{-2i\phi_{IS}\right\}\ ,\\
A_{IS}^{(-2)} &= -\frac{3}{4}\,d_{IS}I_I^-I_S^-\ , & F_{IS}^{(-2)} &= \sin^2\theta_{IS}\exp\left\{+2i\phi_{IS}\right\}\ ,
\end{aligned}
$$

where θ_{IS} is the angle between the static magnetic field $\mathbf{B}_0$ and the $\mathbf{r}_{IS}$ internuclear vector, and ϕ_{IS} is the azimuthal angle with respect to the x-axis (see Figure 1.2). Using

the concept of "dipolar alphabet",[20] [2] which depends on $F_{IS}^{(q)}$ and $A_{IS}^{(q)}$

$$\begin{aligned} A + B &= \frac{1}{d_{IS}} F_{IS}^{(0)} A_{IS}^{(0)} \ , \\ C &= \frac{1}{d_{IS}} F_{IS}^{(1)} A_{IS}^{(1)} \ , \\ D &= \frac{1}{d_{IS}} F_{IS}^{(-1)} A_{IS}^{(-1)} \ , \\ E &= \frac{1}{d_{IS}} F_{IS}^{(2)} A_{IS}^{(2)} \ , \\ F &= \frac{1}{d_{IS}} F_{IS}^{(-2)} A_{IS}^{(-2)} \ , \end{aligned}$$

the dipolar coupling Hamiltonian can be written as

$$\mathscr{H}_D = d_{IS} \left(A + B + C + D + E + F \right) \ . \tag{1.21}$$

Under the high-field approximation, only the A and B ($q = 0$) terms are retained, as they commute with the Zeeman interaction ($\mathbf{I}_z$) and they are time-independent, while C, D, E and F —non-secular terms— are time-dependent and average out with $\mathbf{I}_z$. After truncation of the non-secular terms, the dipolar Hamiltonian as expressed in Equation (1.19) simplifies to the following expression

$$\mathscr{H}_D = \frac{1}{2} d_{IS} \left(1 - 3\cos^2 \theta_{IS} \right) \left(2\mathbf{I}_{Iz}\mathbf{I}_{Sz} - \frac{1}{2} \left(I_I^+ I_S^- + I_I^- I_S^+ \right) \right) \ . \tag{1.22}$$

Invoking the following relationship between the Cartesian angular momentum operators and the raising-and-lowering operators

$$I_I^+ I_S^- = \frac{1}{2} (2I_{Ix}I_{Sx} + 2I_{Iy}I_{Sy} - i2I_{Ix}I_{Sy} + i2I_{Iy}I_{Sx}) \ , \tag{1.23}$$

$$I_I^- I_S^+ = \frac{1}{2} (2I_{Ix}I_{Sx} + 2I_{Iy}I_{Sy} - i2I_{Sx}I_{Iy} + i2I_{Sy}I_{Ix}) \ , \tag{1.24}$$

the dipolar Hamiltonian can be further simplified to

$$\mathscr{H}_D = \frac{1}{2} d_{IS} (1 - 3\cos^2 \theta_{IS}) (2\mathbf{I}_{Iz}\mathbf{I}_{Sz} - I_{Ix}I_{Sx} - I_{Iy}I_{Sy}) \ . \tag{1.25}$$

In heteronuclear spin systems, and in homonuclear systems in the weak coupling regime (see Equation (1.16)), the terms involving transverse spin operators can be

[2] the reader can find the direct relation between the dipolar alphabet and normalized spherical harmonics, as well as with dipolar spin tensors in reference [22]

dropped out

$$\begin{aligned}\mathscr{H}_D &= d_{IS}\left(1-3\cos^2\theta_{IS}\right)\mathbf{I}_{Iz}\mathbf{I}_{Sz} \\ &= 2\pi D_{IS}\mathbf{I}_{Iz}\mathbf{I}_{Sz}\ ,\end{aligned} \tag{1.26}$$

where the Dipolar Coupling D_{IS} is defined as

$$\langle\overline{D}_{IS}\rangle = -\frac{\mu_0\gamma_I\gamma_S\hbar}{4\pi^2 r_{IS}^3}\frac{\left(3\cos^2\theta_{IS}-1\right)}{2}\ . \tag{1.27}$$

From this equation, it easily follows that the dipolar coupling reaches its maximum value, D_{IS}^{max}, when the internuclear vector ($\mathbf{r}_{IS}$) is collinear the static magnetic $\mathbf{B}_0$ field (i.e. $\theta_{IS} = 0°$or $180°$).

Taking as an example an archetypical C(sp^3)–H aliphatic bond with length (r_{IS}) of 1.10 Å, by using Equation (1.27) we find a maximum dipolar coupling for $\theta_{IS} = 0$ or π, of

$$D_{IS} = -\frac{\mu_0\gamma_I\gamma_S\hbar}{4\pi^2 r_{IS}^3} \approx -45 \text{ kHz}. \tag{1.28}$$

Dipole-dipole coupling, with a magnitude of tens of kHz, dominate the solid-state NMR spectra impairing the acquisition of high-resolution NMR spectra. Furthermore, it is no possible to extract the angular information contained in the dipolar coupling from a NMR spectrum dominated by such interaction.

1.2.4 Quadrupolar Coupling

Nuclei with quantum number $I > \frac{1}{2}$ have a non-spherical distribution of the electric charge, resulting in a nuclear electric quadrupole moment Q. The interaction of the quadrupole moment with the Electric Field Gradient (EFG) generated by their surroundings is called quadrupolar coupling. The quadrupolar coupling Hamiltonian can be expressed as

$$\mathscr{H}_Q = \sum_{I=1}^{N}\mathbf{I}_I\mathbf{Q}_I\mathbf{I}_I\ , \tag{1.29}$$

where $\mathbf{Q}_I$ is the quadrupole coupling tensor, which can be expressed in terms of the EFG tensor $\mathbf{V}_I$ at the site of nucleus I

$$\mathbf{Q}_I = \frac{eQ_I}{2\mathbf{I}_I\left(2\mathbf{I}_I-1\right)\hbar}\mathbf{V}_I\ . \tag{1.30}$$

Given that, generally, the quadrupolar interaction is not much less than the Zeeman interaction, the secular approximation can not be used to drop out some of the terms. This leads to a distinction between two interaction orders:

i) the *first order* quadrupolar coupling, equivalent to the secular approximation and sufficient when the quadrupolar coupling is small; it averages out in isotropic solution and produces the splitting of the signal in $2I - 1$ multiplets in anisotropic media, and

ii) the *second order* quadrupolar coupling; it only needs to be considered in the case of strong quadrupoles, is not averaged to zero in isotropic solution, and is field-dependent.

The splitting observed in the deuterium signal of aligned samples ($|\Delta\nu_Q|$) is the consequence of the first order quadrupolar interaction, and is a useful indicator that reflects the degree of alignment in a particular sample.

1.3 Reintroduction of anisotropic interactions in high-resolution liquid-state NMR spectroscopy. Residual Dipolar Couplings

We have seen in the previous section that NMR anisotropic interactions (e.g. chemical shielding and dipolar coupling) are rich in structural information, though difficult to interpret, as evidenced by the difference in resolution between liquid-state (high-resolution) and solid-state NMR spectra. Fortunately, the large dipole-dipole interactions can be scaled down to a manageable magnitude, while keeping the benefits of high-resolution NMR spectra (e.g. narrow spectral lines), by introducing some degree of order in the solution. This can be achieved by any means that perturbs the free reorientation of solute molecules in solution, thus increasing the probability of some orientations over the others. The most general approach is the use of the so-called *alignment media*, such as liquid crystals or anisotropic gels (see Section 1.9). Besides, some molecules may orient on their own due to the anisotropy of their magnetic susceptibility.

So, samples in aligned media can be considered as an intermediate state between the solid-state and the liquid-state, in which all anisotropic Hamiltonians average out due to the fast isotropic tumbling of the molecules.

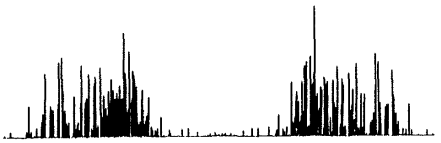


FIGURE 1.3: ^{1}H NMR spectra of 4,4′-dichlorobiphenyl (DCB) in I52 nematic phase. Adapted from[35]. The spectral complexity due to the large size of dipolar couplings is easily appreciated.

The dipolar couplings of a guest molecule dissolved in an oriented medium was first described in the 1960s By Saupe,[36] and Emsley and Lindon,[37] although with very small molecules such as benzene. In this early work, nematic phases were used as alignment media, which induced a too strong degree of orientation, resulting in dipolar couplings in the kHz range.[38,39] Such strong alignment hampered the application of dipolar couplings to structural determination due to the extreme complexity of the strongly-coupled spectra (Figure 1.3).

The field changed dramatically with the introduction of weak alignment media that induce partial alignment of biological macromolecules by Tjandra and Bax.[34] With these new alignment media, the dipolar couplings where scaled down from several kHz to a few Hz. These are referred to as residual dipolar couplings (RDCs). The persistence of only a residual part of anisotropic interactions allows easy retrieval of the information from high-resolution NMR spectra.

In isotropic liquids, molecules tumble and adopt all orientations with equal probability, hence dipolar couplings average to zero, as can be seen by evaluating this integral

$$\int_0^{\pi} d\theta_{IS} \sin\theta_{IS} (3\cos^2\theta_{IS} - 1) = 0 \ . \tag{1.31}$$

In anisotropic liquids, some orientations are more populated than others and dipolar couplings average as follows

$$\int_0^{\pi} d\theta_{IS} p(\theta_{IS}) (3\cos^2\theta_{IS} - 1) = -\frac{\mu_0 \gamma_I \gamma_S \hbar}{4\pi^2 r_{IS}^3} \left\langle \frac{(3\cos^2\theta_{IS} - 1)}{2} \right\rangle \neq 0 \ . \tag{1.32}$$

where $p(\theta_{IS})$ is the probability distribution of the orientation of vector IS in such aligned medium. This small —residual— part of dipole-dipole coupling is called Residual Dipolar Coupling (RDC). This topic will be expanded in section Section 1.9.

Provided that the size of the dipolar coupling is scaled down in a sufficient degree, the weak coupling condition is fulfilled and Equation (1.26) holds, giving rise to first-order spectra. As the scalar coupling Hamiltonian also operates, the total spin-spin

coupling term is

$$\begin{aligned}\mathscr{H}_{spin-spin} &= \mathscr{H}_J + \mathscr{H}_D \\ &= 2\pi\langle\overline{D}_{IS}\rangle\mathbf{I}_{Iz}\mathbf{I}_{Sz} + 2\pi J_{IS}\mathbf{I}_{Iz}\mathbf{I}_{Sz} \ .\end{aligned} \tag{1.33}$$

Note that the dipolar coupling Hamiltonian (Equation (1.26)) has the same form as the scalar coupling Hamiltonian. Hence, the total coupling (Hamiltonian) is

$$\mathscr{H}_{spin-spin} = 2\pi T_{IS}\mathbf{I}_{Iz}\mathbf{I}_{Sz} \ , \tag{1.34}$$

where ${}^nT_{\mathrm{CH}}$ is the total coupling

$${}^nT_{\mathrm{CH}} = {}^n J_{\mathrm{CH}} + {}^n D_{\mathrm{CH}} \ . \tag{1.35}$$

This means that the residual dipolar coupling appears as an addition to the splitting caused by the J coupling and can be directly extracted as ${}^nD_{\mathrm{CH}} = {}^nT_{\mathrm{CH}} - {}^nJ_{\mathrm{CH}}$. In the literature, the reader can find a different definition of the total splitting as follows

$${}^nT_{\mathrm{CH}} = {}^n J_{\mathrm{CH}} + 2{}^n D_{\mathrm{CH}} \ , \tag{1.36}$$

which stems from an alternative definition of the dipolar coupling Hamiltonian (compare with Equation (1.26))

$$\mathscr{H}_D = 2\pi D_{IS} 2\mathbf{I}_{Iz}\mathbf{I}_{Sz} \ . \tag{1.37}$$

Interestingly, RDCs retain all the structural information intrinsic to dipolar couplings, i.e. spin-spin distances and angles. In the following, we will finish the derivation of the treatment of residual dipolar coupling. Additionally, a brief introduction to the history of RDC couplings, and the derivation of the key equations that permit the use of RDC for structural analysis will be presented.

1.3.1 Treatment of molecular alignment in partially oriented media. The alignment tensor

By dissolving the molecule in weakly aligned media, certain orientations respect to the static field will be slightly preferred, causing a residual part of the dipole-dipole coupling to show up. In the LABoratory frame (LAB), the static field $\mathbf{B_0}$ points along the z axis and remains constant ($\mathbf{B_0} = B\vec{b}$). As the molecule tumbles in solution, the internuclear vector $\vec{R}_{IS} = R_{IS}\vec{r}_{IS}$ is time-dependent. $\vec{r}_{IS}$ is the $\vec{R}_{IS}$ unit vector and its time dependence is contained in the angle θ_{IS} that forms with $\mathbf{B_0}$. If we assume complete rigidity of the molecule, so that the time dependence of $\vec{R}$ is merely due

to the molecular tumbling motion, the term $\cos^2\theta_{IS}$ (where θ_{IS} is the angle formed by $\vec{r}_{IS}$ and $\mathbf{B_0}$) becomes time-dependent, as well as the dipolar coupling D_{IS}. The (time-averaged) residual dipolar coupling depends on the average alignment of the molecule, is expressed in polar coordinates as follows

$$\overline{D}_{IS} = -\frac{\mu_0\gamma_I\gamma_S\hbar}{4\pi^2 r_{IS}^3}\frac{\left(\overline{3\cos^2\theta_{IS}-1}\right)}{2} \ . \tag{1.38}$$

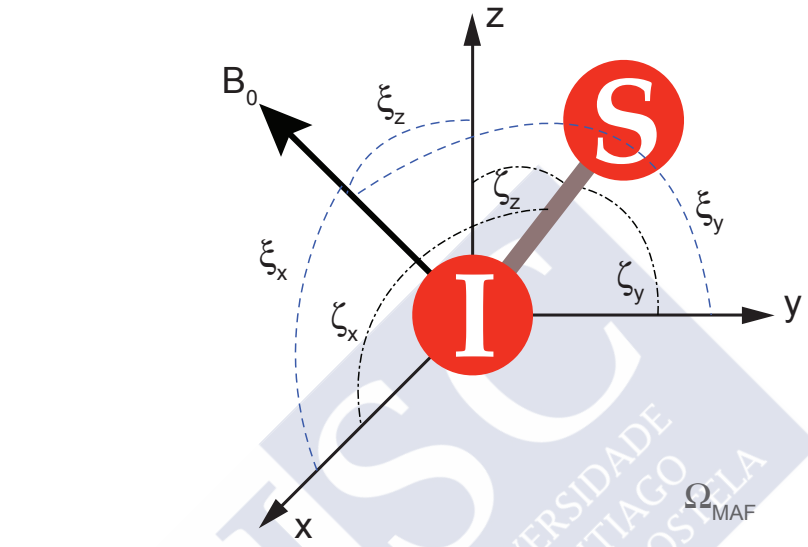


FIGURE 1.4: Definition of the (ξ_{IS} and ζ_{IS}) angles used for the connection of the MOLECULE-FIXED AXIS FRAME with the PAS, see text. The angles are expressed in the MOLECULE-FIXED AXIS FRAME.

Two coordinate frames will be considered in the next: the LAB frame and the internal molecular frame (MAF). The angle θ_{IS} formed by the internuclear vector with the static magnetic field (Figure 1.2), can be expressed in terms of the orientation of the molecule (i.e. of its MAF) respect to the magnetic field $(\xi_x,\ \xi_y,\ \xi_z)$, and the orientation of the internuclear vector with respect to the molecular frame (MAF) axes $(\zeta_x,\ \zeta_y,\ \zeta_z)$.[27] The definition of these angles is given in Figure 1.4. Note that, in this derivation, the MAF is considered as the fixed frame and the orientation of the LAB frame (hence the static $\vec{b}$) is considered to vary respect to this frame upon molecular tumbling (i.e. angles ζ_k remain constant while all molecular motion is accounted for by angles ξ_k).

$$\begin{aligned}\cos\theta_{IS} &= \begin{pmatrix}\cos\zeta_x & \cos\zeta_y & \cos\zeta_z\end{pmatrix}\cdot\begin{pmatrix}\cos\xi_x\\ \cos\xi_y\\ \cos\xi_z\end{pmatrix}\\ &= \cos\zeta_x\cos\xi_x + \cos\zeta_y\cos\xi_y + \cos\zeta_z\cos\xi_z \ .\end{aligned} \tag{1.39}$$

With the previous relationship, the time average given by the second element in Equation (1.38) can be expressed in terms of the director cosines

$$\frac{\left(\overline{3\cos^2\theta_{IS}-1}\right)}{2} = \frac{3}{2}\left(\overline{\cos\zeta_x\cos\xi_x+\cos\zeta_y\cos\xi_y+\cos\zeta_z\cos\xi_z}\right)^2-\frac{1}{2}\ , \tag{1.40}$$

which can be reordered, using $C_i=\cos\zeta_i$ and $c_i=\cos\xi_i$, for simplicity

$$\begin{aligned}\frac{\left(\overline{3\cos^2\theta_{IS}-1}\right)}{2} =& \frac{3}{2}\left[\overline{c_x^2}C_x^2+\overline{c_y^2}C_y^2+\overline{c_z^2}C_z^2\right]\\ &+2\overline{c_xc_y}C_xC_y+2\overline{c_xc_z}C_xC_z\\ &+2\overline{c_yc_z}C_yC_z-\frac{1}{2}\ .\end{aligned} \tag{1.41}$$

The average orientation of the molecule with respect to the field $\left(\overline{c_x^2},\ \overline{c_y^2},\ \overline{c_z^2},\ \overline{c_xc_y},\ \overline{c_xc_z},\text{ and }\overline{c_yc_z}\right)$ can be described by an order tensor $\hat{\mathbf{S}}$ (Saupe matrix) that was introduced by Saupe.[40] In an arbitrarily chosen Molecular Alignment Frame (MAF), the relationship between the RDC and the $\hat{\mathbf{S}}$ tensor can be expressed as follows

$$\overline{D}_{IS} = -\frac{\mu_0\gamma_I\gamma_S\hbar}{4\pi^2r_{IS}^3}\sum_{k,l=\{x,y,z\}}S_{kl}\overline{\cos\zeta_k\cos\zeta_l}\ ,\text{ with} \tag{1.42}$$

$$S_{kl} = \frac{\overline{3\cos\xi_k\cos\xi_l-\delta_{kl}}}{2}\ , \tag{1.43}$$

in which δ_{kl} is the Kronecker delta with values $\delta_{kl}=\begin{cases}1 & \text{if } k=l\\ 0 & \text{if } k\neq l\end{cases}$.

Alternatively, we can express $\overline{D}_{IS}$ as

$$\overline{D}_{IS} = -\frac{3\mu_0\gamma_I\gamma_S\hbar}{8\pi^2r_{IS}^3}\left(\overline{\cos^2\theta_{IS}}-\frac{1}{3}\right)\ . \tag{1.44}$$

In this case, the term $\left(\overline{\cos^2\theta_{IS}}-\frac{1}{3}\right)$ will be equal to

$$\overline{\cos^2\theta_{IS}}-\frac{1}{3} = \sum_{k,l=\{x,y,z\}}A_{kl}\overline{\cos\zeta_k\cos\zeta_l} \tag{1.45}$$

$$= \begin{pmatrix}\cos\zeta_x & \cos\zeta_y & \cos\zeta_z\end{pmatrix}\begin{pmatrix}A_{xx} & A_{xy} & A_{xz}\\ A_{xy} & A_{yy} & A_{yz}\\ A_{xz} & A_{yz} & A_{zz}\end{pmatrix}\begin{pmatrix}\cos\zeta_x\\ \cos\zeta_y\\ \cos\zeta_z\end{pmatrix} \tag{1.46}$$

$$= r^T\hat{\mathbf{A}}r\ . \tag{1.47}$$

The alignment tensor $\hat{\mathbf{A}}$ is a symmetric and traceless matrix ($A_{kl} = A_{lk}$; $A_{xx} + A_{yy} + A_{zz} = 0$), hence it has only five different elements. Tensors $\hat{\mathbf{A}}$ and $\hat{\mathbf{S}}$ differ only in a scaling factor

$$\hat{\mathbf{A}} = \frac{2}{3}\hat{\mathbf{S}} \ , \tag{1.48}$$

Therefore, the alignment tensor of a given molecule can be calculated, in principle, as soon as five *independent* experimental RDCs are available (although more experimental values are often desirable).

It is important to stress the linear independence requirement of, at least, five RDCs. For instance, in the model pentose shown in Figure 1.5, up to six one-bond C–H RDCs can be obtained from NMR experiments, but only three of them are in fact independent (non-parallel). Note that all the axial C–H vectors (in red) point in the same direction, and two of the equatorial ones (in green) are also non-independent. This kind of low-informative systems are quite common in natural products and sugar moieties.

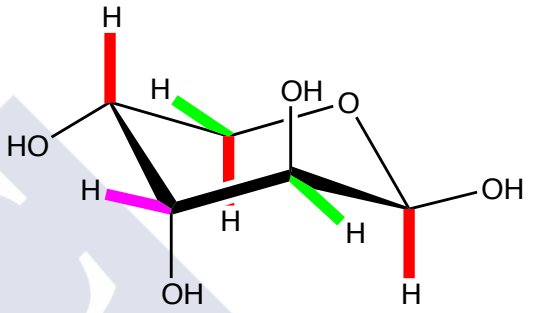


FIGURE 1.5: Model pentose showing linearly dependent vectors shaded in the same color.

The alignment tensor can be rotated from the molecule-fixed frame (MAF) to its *principal axis system* (PAS), in which the off-diagonal elements are zero. In this frame, and by following the convention $|A'_x| \leq |A'_y| \leq |A'_z|$, the axial ($A_{ax}$) and rhombic ($A_{rh}$) asymmetries of $\hat{\mathbf{A}}$ are defined as follows

$$A_{ax} = \frac{3}{2}A'_z = S'_z \ , \tag{1.49}$$

$$A_{rh} = A'_x - A'_y = \frac{2}{3}\left(S'_x - S'_y\right) \ . \tag{1.50}$$

The rhombic component measures the deviation of the tensor from axial symmetry.

The relationship between the residual dipolar coupling and A_{ax}, A_{rh} can be expressed in terms of the polar angles (θ, ϕ; defined in Figure 1.2) of the inter-spin vector in the PAS frame

$$D_{IS} = -\frac{\mu_0 \gamma_I \gamma_S \hbar}{4\pi^2 r_{IS}^3}\left[A_{ax}(3\cos^2\theta_{IS} - 1) + \frac{3}{2}A_{rh}\sin^2\theta_{IS}\cos 2\phi_{IS}\right] \ . \tag{1.51}$$

Another useful parameter that describes the deviation of the alignment tensor from axial symmetry is the asymmetry parameter (η)

$$\eta = \frac{A'_{xx} - A'_{yy}}{A'_{zz}} \ . \tag{1.52}$$

The alignment tensor encodes the orientational probability of the molecule in the aligned conditions. With the equations derived so far, it is possible to predict the expected RDCs of a molecule if its alignment tensor is known. Therefore, the problem of interpretation of experimental RDCs can be reduced to determining the alignment tensor, as will be described in the next.

1.4 Interpretation of RDCs. Determination of the alignment tensor and fit of experimental RDCs to candidate structures

Along this thesis and in the specialized literature cited, the procedure to interpret RDC is usually referred to with the term "RDC fit" and comprises:

i) to propose one or more candidate structures of the molecule, typically by molecular modeling (with or without restraints);

ii) to set up a system of equations that relates descriptors of the molecular geometry, the unknown alignment tensor components, and the experimental RDC values;

iii) to solve such system of equations, which is equivalent to find the best fit of the model (candidate geometry and alignment tensor) to the experimental RDC values.

The first item is not discussed in this Introduction but examples are given in the results chapters. This section focuses on items ii) and iii).

Along this thesis dissertation, as well as in our group, we tend to use two different methods to determine the alignment matrix, both furnishing equivalent results. Singular Value Decomposition, is one of the most used methods for this task, and will be presented in the following. Additionally, we use the Powell's conjugate direction method for the same task.

1.4.1 Order matrix analysis

Following Losonczi *et al.*,[41] RDC can be analyzed by setting up a linear system of the type $\vec{d} = \hat{\mathbf{M}}\hat{\mathbf{A}}$. The time-averaged dipolar coupling of a spin pair j ($j = IS$), can be expressed in matrix notation as

$$\begin{aligned} \overline{D}_j &= \frac{\kappa}{r_j^3}\vec{r}_j{}^T\hat{\mathbf{A}}\vec{r}_j, \text{ and} \\ \overline{d}_j &= \frac{\overline{D}_j}{\left(\frac{\kappa}{r_j^3}\right)}, \text{ then} \\ \overline{d}_j &= \vec{r}_j{}^T\hat{\mathbf{A}}\vec{r}_j \ , \end{aligned} \tag{1.53}$$

in which $\overline{d}_j$ is the so-called *reduced* coupling and $\vec{r}_j$ is the *unitary* internuclear vector, hence $\vec{r}_j$ elements are its director cosines. The last equation can be expressed in matrix form

$$\overline{d}_j = \begin{pmatrix} x_j & y_j & z_j \end{pmatrix} \begin{pmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{xy} & A_{yy} & A_{yz} \\ A_{xz} & A_{yz} & A_{zz} \end{pmatrix} \begin{pmatrix} x_j \\ y_j \\ z_j \end{pmatrix} \ , \tag{1.54}$$

by multiplying the $\hat{\mathbf{A}}$ tensor by its right

$$\overline{d}_j = \begin{pmatrix} x_j & y_j & z_j \end{pmatrix} \begin{pmatrix} A_{xx}x_j + A_{xy}y_j + A_{xz}z_j \\ A_{xy}x_j + A_{yy}y_j + A_{yz}z_j \\ A_{xz}x_j + A_{yz}y_j + A_{zz}z_j \end{pmatrix} \ . \tag{1.55}$$

By further multiplying by its left we arrive at

$$\begin{aligned} \overline{d}_j = {} & A_{xx}x_j^2 + A_{xy}x_jy_j + A_{xz}x_jz_j \\ & + A_{xy}x_jy_j + A_{yy}y_j^2 + A_{yz}y_jz_j \\ & + A_{xz}x_jz_j + A_{yz}y_jz_j + A_{zz}z_j^2 \ , \end{aligned} \tag{1.56}$$

that can be reordered to obtain the following expression

$$\overline{d}_j = A_{xx}x_j^2 + 2A_{xy}x_jy_j + 2A_{xz}x_jz_j + A_{yy}y_j^2 + A_{yz}y_jz_j + A_{zz}z_j^2 \ . \tag{1.57}$$

Now, if we recall the properties of the alignment matrix we can substitute in the previous expression $A_{xx} = -A_{yy} - A_{zz}$, to get

$$\overline{d}_j = 2A_{xy}x_jy_j + 2A_{xz}x_jz_j + A_{yy}(y_j^2 - x_j^2) + 2A_{yz}y_jz_j + A_{zz}(z_j^2 - x_j^2) \ , \tag{1.58}$$

in which only five elements of the alignment matrix are retained. Now we can re-write these expression back to matrix form

$$\overline{d}_j = \begin{pmatrix} 2x_jy_j & 2x_jz_j & (y_j^2 - x_j^2) & 2y_jz_j & (z_j^2 - x_j^2) \end{pmatrix} \begin{pmatrix} A_{xy} \\ A_{xz} \\ A_{yy} \\ A_{yz} \\ A_{zz} \end{pmatrix} \ . \tag{1.59}$$

If n experimental RDC values are available, n equations like Equation (1.59) can be written ($j = 1, 2 \ldots n$). This defines a system of equations

$$\vec{d} = \hat{\mathbf{M}}\hat{\mathbf{A}} \tag{1.60}$$

$$\begin{pmatrix} d_1 \\ d_2 \\ \ldots \\ \ldots \\ d_n \end{pmatrix} = \begin{pmatrix} 2x_1y_1 & 2x_1z_1 & (y_1^2 - x_1^2) & 2y_1z_1 & (z_1^2 - x_1^2) \\ 2x_2y_2 & 2x_2z_2 & (y_2^2 - x_2^2) & 2y_2z_2 & (z_2^2 - x_2^2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 2x_ny_n & 2x_nz_n & (y_n^2 - x_n^2) & 2y_nz_n & (z_n^2 - x_n^2) \end{pmatrix} \begin{pmatrix} A_{xy} \\ A_{xz} \\ A_{yy} \\ A_{yz} \\ A_{zz} \end{pmatrix} \ , \tag{1.61}$$

therefore, a new row on the $\hat{\mathbf{M}}$ matrix will be added for every experimental RDC, thus building the matrix with n rows for n experimental RDCs.

So the linear equation system ($\vec{d} = \hat{\mathbf{M}}\hat{\mathbf{A}}$) is composed of a vector containing the reduced experimental RDCs ($\overline{d}_{IS}$, Equation (1.53)), the alignment tensor ($\hat{\mathbf{A}}$, Equation (1.46)) represented by a column vector containing its unknowns, and a molecular coordinates matrix ($\hat{\mathbf{M}}$) composed by director cosines products, which derives from the known candidate structures.

1.4.2 SVD fit

The singular value decomposition method consists in a factorization of a $m \times n$ (symmetric, positive semi-definite) matrix $\hat{\mathbf{M}}$ of the form

$$\hat{\mathbf{M}} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T = \mathbf{U} \begin{pmatrix} \Sigma_{11} & 0 & \cdots & 0 \\ 0 & \Sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_{mm} \end{pmatrix} \mathbf{V}^T \ , \tag{1.62}$$

where $\mathbf{U}$ is an $n \times m$ matrix, $\mathbf{\Sigma}$ is a $m \times m$ diagonal matrix with non negative real diagonal, and $\mathbf{V}^T$ is an $n \times n$ unitary matrix. The diagonal entries Σ_{ii} are the *singular values* of $\hat{\mathbf{M}}$. The m columns of $\mathbf{U}$ and the n columns of $\mathbf{V}$ are the left-singular vectors, and the right -singular vectors of $\hat{\mathbf{M}}$, respectively. According to this decomposition, the inverse of $\hat{\mathbf{M}}$ is calculated as

$$\mathbf{M}^{-1} = \mathbf{V} \left[\text{diag}\left(\frac{1}{\Sigma_{mm}}\right) \right] \mathbf{U}^T \ . \tag{1.63}$$

Geometrically, SVD decomposes the matrix $\mathbf{M}$ intro three simple transformations: a rotation ($\mathbf{V}$), a scaling along the rotated coordinate axes ($\mathbf{\Sigma}$), and a second rotation ($\mathbf{U}$).

Going back to our linear system $\vec{d} = \hat{\mathbf{M}}\hat{\mathbf{A}}$, $\hat{\mathbf{A}}$ can be determined by the latter decomposition of the form

$$\hat{\mathbf{A}} = \mathbf{V} \left[\text{diag}\left(\frac{1}{\Sigma_{mm}}\right) \right] \mathbf{U}^T \vec{d} \ . \tag{1.64}$$

Note from Equation (1.61) that m is always 5 (the number of elements of vector $\hat{\mathbf{A}}$). If the equation system has no exact solution, as in the case of overdetermined systems ($n > m$), the SVD method will still provide a solution that will be the best in a least squares sense. However, if the system is not well conditioned, i.e. there are no five independent vectors available, small changes in the input will translate into large changes in the computed alignment tensor. The system conditioning can be evaluated in terms of the *condition number* (*c.n.*), defined as

$$c.n. = ||\hat{\mathbf{M}}||\,||\hat{\mathbf{M}}^{-1}|| \ . \tag{1.65}$$

The condition number can also be expressed as the ratio between the maximum and minimum singular values of the SVD decomposition. Large condition numbers ($c.n. > 100$) usually indicate lack of sufficient vectors to determine the tensor.

Along this thesis, SVD will be done with the program Mspin.[42,43]

1.4.3 Comparison of alignment tensors

Two different tensors $\hat{\mathbf{A}}^{(1)}$ and $\hat{\mathbf{A}}^{(2)}$ can be compared in terms of, mainly, two parameters, namely, the Global Degree of Order (*GDO*) and the generalized angle β. Following Kramer *et al.*,[26] they are defined as

$$GDO = \sqrt{\frac{3}{2}|\hat{\mathbf{A}}|} \ , \tag{1.66}$$

$$\cos\beta = \frac{\left\langle \hat{\mathbf{A}}^{(1)}|\hat{\mathbf{A}}^{(2)} \right\rangle}{|\hat{\mathbf{A}}^{(1)}|\,|\hat{\mathbf{A}}^{(2)}|} \ . \tag{1.67}$$

The *GDO* reflects the degree of orientation on the molecule in the aligned sample, while the generalized angle β represents the global rotation between the tensors. The scalar product and the norm of the matrices are defined as follows

$$|\hat{\mathbf{A}}| = \sqrt{\sum_{k,l=\{x,y,z\}} A_{kl}^2} \ , \tag{1.68}$$

$$\left\langle \hat{\mathbf{A}}^{(1)}|\hat{\mathbf{A}}^{(2)} \right\rangle = \sum_{k,l=\{x,y,z\}} A_{kl}^{(1)}\, A_{kl}^{(2)} \ . \tag{1.69}$$

1.4.4 Quality of the fit

Once the alignment tensor is determined, all dipolar couplings can be back-calculated and compared with the experimental ones. The quality of the fit among them can be used as a merit function to select the correct structure among a set of candidate geometries.

Usually, the quality of the fit is expressed in terms of the Cornilescu Q_C factor[44]

$$Q_C = \sqrt{\frac{\sum_j^n (D_j^{calc} - D_j^{exp})^2}{\sum_j^n (D_j^{exp})^2}} \ . \tag{1.70}$$

The fact that the Q_C factor is scaled by the experimental values allows the comparison of fits between different media possessing different generalized degrees of order. Note however, that the Q_C factor was developed in the context of protein structure determination, where enough RDC vectors are available to completely define the 5D space of the tensor. In contrast, fits of small molecules can be hampered by a lack of exploration of all possible spatial orientations of the experimental RDC vectors.

On condition that good quality candidate geometries are chosen, it can be assumed that the "correct" structure is the one that best fits the experimental dipolar couplings. So far, this approach is valid only for a rigid molecule, as the assumption of complete molecular rigidity has been made to build the time-dependent dipolar coupling constant equation (Equation (1.38)). Nevertheless, there is not a common agreement in the field about which merit function expresses best the agreement between the experimental and back-calculated RDCs.

In recent years, other functions have been presented aiming at better expressing the fit between D_j^{exp} and D_j^{calc}. Some of them try to incorporate the effect of experimental errors, like the n/χ^2 proposed by Hofmann and co-workers,[45,46] where n is the number of experimental couplings available and the denominator χ^2 of the function is defined as

$$\chi^2 = \sum_j^n \left(\frac{D_j^{exp} - D_j^{calc}}{\Delta D_j^{exp}} \right)^2 , \tag{1.71}$$

where ΔD_j^{exp} is the experimental error of each D_j^{exp}.

Likewise, the Q_σ[47] function proposed by Baltzar and co-workers includes an experimental error term, σ_j, which weights each RDC value

$$Q_\sigma = \sqrt{\frac{\sum_j^n \left(\frac{D_j^{calc} - D_j^{exp}}{\sigma_j} \right)^2}{\sum_j^n \left(\frac{D_j^{exp}}{\sigma_j} \right)^2}} , \tag{1.72}$$

note that Q_σ is equivalent to Q_C when all experimental errors σ_j are the equal.

1.4.5 Error treatment

Introduction of functions Equation (1.71) and Equation (1.72) raises the question of how experimental error is treated. Both functions attempt to include in the fit the individual error of each RDC.

In general, errors on the RDC values are estimated from peak shapes, using among others

i) the *peak center error* method proposed by Bax and co-workers.[48] For well-resolved doublets, the lower limit for the determination of the center of the peak can be given by LW/SN, where LW is the linewidth (down-scaled by κ, if a

κ-scaled J-evolution module is used in the experiment) and SN is the signal-to-noise ratio in the weaker of the two 2D spectra, which is usually the one acquired for the aligned sample

ii) the *peak displacement* method proposed by Luy and co-workers.[49] When extracting the couplings from 1D slices of the 2D experiments, usually the α and β components are shifted and the difference in Hz is taken as the coupling. The common choice is to align the center of the peaks, this method consists in the alignment of the leftmost and rightmost sides of the peak, these left- and right-couplings are averaged and 1/2 of their difference is taken as the error.

However, it is worth noting that other sources of error may affect the experimental values, like strong-coupling.[50]

Estimated errors can be taken into account by using the functions presented previously (Equation (1.71) and Equation (1.72)) or by resampling methodologies.

Bootstrapping. The experimental RDC data set has the form

$$\vec{\mathbf{D}}^{exp} = \begin{pmatrix} D_1^{exp} \\ D_2^{exp} \\ \vdots \\ D_i^{exp} \\ \vdots \\ D_n^{exp} \end{pmatrix} ;\ n \text{ experimental RDC values: } D_i^{exp};\ i = 1, 2, \ldots, n\ . \tag{1.73}$$

First, P simulated data sets or *decoys* are generated by sampling a Gaussian distribution centered on each D_i^{exp} experimental value with a given σ_i. This σ_i can have a different value for each D_i^{exp}, or a global uniform value can be used instead

$$\vec{\mathbf{D}}_k^{sim} = \begin{pmatrix} D_{1,k}^{sim} \\ D_{2,k}^{sim} \\ \vdots \\ D_{i,k}^{sim} \\ \vdots \\ D_{n,k}^{sim} \end{pmatrix} ;\ \text{where: } k = 1, 2, \ldots, P\ . \tag{1.74}$$

Each $\hat{\mathbf{D}}_{\mathbf{k}}^{\mathbf{sim}}$ is fitted to the molecular geometry $\hat{\mathbf{M}}$, resulting in an alignment tensor $\hat{\mathbf{A}}_k$ and a back-calculated RDC vector $\tilde{\mathbf{D}}_k^{back}$

$$\hat{\mathbf{D}}_k^{back} = \begin{pmatrix} D_{1,k}^{back} \\ D_{2,k}^{back} \\ \vdots \\ D_{i,k}^{back} \\ \vdots \\ D_{n,k}^{back} \end{pmatrix} \text{; where: } k = 1, 2, \ldots, P \ . \tag{1.75}$$

Sometimes *all* simulated datasets are analyzed. In such case, we refer to the experiment as *Bootstrapping* to remark that no rejection filter has been imposed. The Bootstrapping test reflects how severe the propagation (impact) of experimental errors can be. Analysis of these fits gives an idea of the distribution of alignment tensors (hence of back-calculated RDCs) that are accessible within the given experimental error. A drawback of Bootstrapping is that many of the back-calculated datasets do not fit the experimental data. In the following, the Montecarlo filter of Bootstrap *decoys* is presented.

Montecarlo. This is a refinement of the Bootstrapping, following Prestegard.[41] $\vec{\mathbf{D}}_k^{back}$ is accepted if every element satisfies

$$|D_{i,k}^{back} - {D_k^{exp}}_i| < \sigma_i^{pass}, \ \forall \ i = 1, 2, \ldots n \ , \tag{1.76}$$

where σ_i^{pass} is the allowed deviation between experimental and back-calculated couplings. The value of σ_i^{pass} can be different for each experimental value D_i^{exp} (σ_i^{pass}) or a uniform value can be used (σ^{pass}). Typically, σ^{pass} is equal to the σ_i of the Gaussian distribution used to generate the simulated $\vec{\mathbf{D}}_k^{sim}$. This results in a distribution of alignment tensors of each sample that are compatible with the experimental data within the allowed tolerance σ^{pass}. Accepted decoys are used to calculate the statistics of the fit, e.g. the distribution of values of degree of order (GDO), tensor orientation (angle between tensors, β) or Q_C. Graphical representation of the principal axes of tensors (see Figure 2.9) is also helpful to assess the error propagation in the particular system. Note that we do not rely only on the value of the quality factor to accept or reject a dataset. Instead, the ability of the tensor $\tilde{\mathbf{D}}_{\mathbf{k}}^{\mathbf{back}}$ to reproduce the experimental data (within the predefined error limits) is given the prominent role.

1.5 Treatment of the averaging of RDCs from equivalent C–H

There is an additional complication in the analysis of conformationally rigid small molecules, related to the concept introduced in Figure 1.5, namely the lack of enough independent vectors to sample appropriately the three dimensional space. It often happens that, due to experimental feasibility, only D_{CH} are available (one-bond C–H couplings, like the ones color-shaded in Figure 1.5). One source of vector scarcity is that some of the C–H bonds may be parallel in some cyclic moieties. Therefore, it will be desirable to have a methodology for the inclusion of RDCs from the most common freely rotating groups, namely methyl and phenyl groups.

Methyl. The rotation-averaged ${}^{1}D_{\mathrm{CH}}$ RDC of a methyl group is given by[51]

$$\langle D_{methyl} \rangle = \frac{3\cos^2\eta - 1}{2} D_{\parallel}, \tag{1.77}$$

where, $D_{\parallel}$ is D of a *virtual* C–H vector pointing in the direction of the rotation axis and η is the angle between the C–H vector and the rotation axis. For an ideal tetrahedral angle , this results in $\langle D_{methyl} \rangle = -\frac{1}{3} D_{\parallel}$. Many authors introduce the coupling of methyl groups in their calculations with the artifice of transforming it into the equivalent homonuclear ^{13}C–^{13}C coupling

$$D_{CC} = -\frac{1}{3}\frac{\gamma_C r_{CC}^3}{\gamma_H r_{CH}^3} = -0.22 D_{\parallel} \ . \tag{1.78}$$

Phenyl groups. Similarly, the averaged RDC of fast-exchanging equivalent positions in phenyl groups is

$$\langle D_{phenyl} \rangle = \frac{1}{4} D_{\parallel} + \frac{3}{4} D_{\perp} \ , \tag{1.79}$$

in which $D_{\parallel}$ represents the coupling of a *virtual* C–H vector pointing along the rotation axis, and $D_{\perp}$ represents the coupling of a *virtual* vector orthogonal to the axis and lying in the plane of the ring.

Methylene groups. In most of applications, C–H couplings are measured from proton-coupled spectra, ranging from the simple proton-coupled (gated-decoupled) ^{13}C experiment to the most commonly used ^{13}C–^{1}H HSQC experiments coupled either in the direct[52] (^{1}H, F_2) or indirect[53] (^{13}C, F_1) dimensions. In some cases, the individual protons of methylene groups can be stereospecifically assigned and their individual C–H couplings often can be individually extracted from gated-decoupled ^{13}C or F_2-coupled HSQC experiments. More often, the individual C–H couplings of

methylene groups are not accessible either because they can not be assigned or because they are not well resolved in the 1D or 2D spectra (due to insufficient signal dispersion).

Given that methylenes occur frequently in molecules, we devised a method to introduce them in the RDC fits even in the cases when the individual couplings are not accessible. The resulting, averaged coupling of a methylene is[54]

$$\langle D_{methylene} \rangle = \frac{D_{C-H\alpha} + D_{C-H\beta}}{2} \ . \tag{1.80}$$

This definition also facilitates the extraction of the couplings from F_1-coupled experiments, in which only the external peaks of the multiplet are clearly observed. In such experiments the extraction of the two $^1J_{CH}$ sum (from the outer lines) is more convenient.

1.5.1 Construction of the geometry matrix $\hat{\mathbf{M}}$ to account for averaged C–H couplings

It was seen in Section 1.4.1 that the Residual Dipolar Coupling between nuclei **I** and **S** is given by Equation (1.53)

$$\overline{D} = \frac{\kappa}{r^3} \vec{r}^T \hat{\mathbf{A}} \vec{r} \ . \tag{1.81}$$

Since $\hat{\mathbf{A}}$ is a symmetric and traceless tensor, Equation (1.53) can be expanded leading to the following linear system, which is equivalent to the one shown in Equation (1.61) but for the explicit weighting factors w_i

$$\begin{pmatrix} w_1 d_1 \\ \cdots \\ w_i d_i \\ w_j d_j \\ w_k d_k \\ \cdots \\ w_p d_p \\ w_q d_q \\ \vdots \\ w_N d_N \end{pmatrix} = \hat{\mathbf{M}} \begin{pmatrix} A_{xy} \\ A_{xz} \\ A_{yy} \\ A_{yz} \\ A_{zz} \end{pmatrix} \ , \tag{1.82}$$

where $d_i = D_i \frac{r_i^3}{\kappa}$ is the reduced RDC and w_i is a weighting factor for each matrix element, which is equal to 1 by default. The matrix $\hat{\mathbf{M}}$ in Equation (1.82) contains

director cosines products and represents the molecular geometry

$$\hat{\mathbf{M}} = \begin{pmatrix} 2w_1x_1y_1 & 2w_1x_1z_1 & w_1(y_1^2 - x_1^2) & 2w_1y_1z_1 & w_1(z_1^2 - x_1^2) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 2w_ix_iy_i & 2w_ix_iz_i & w_i(y_i^2 - x_i^2) & 2w_iy_iz_i & w_i(z_i^2 - x_i^2) \\ 2w_jx_jy_j & 2w_jx_jz_j & w_j(y_j^2 - x_j^2) & 2w_jy_jz_j & w_j(z_j^2 - x_j^2) \\ 2w_kx_ky_k & 2w_kx_kz_k & w_k(y_k^2 - x_k^2) & 2w_ky_kz_k & w_k(z_k^2 - x_k^2) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 2w_px_py_p & 2w_px_pz_p & w_p(y_p^2 - x_p^2) & 2w_py_pz_p & w_p(z_p^2 - x_p^2) \\ 2w_qx_qy_q & 2w_qx_qz_q & w_q(y_q^2 - x_q^2) & 2w_qy_qz_q & w_q(z_q^2 - x_q^2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 2w_Nx_Ny_N & 2w_Nx_Nz_N & w_N(y_N^2 - x_N^2) & 2w_Ny_Nz_N & w_N(z_N^2 - x_N^2) \end{pmatrix} , \quad (1.83)$$

where x_i, y_i, z_i are the direction cosines of the unit vector $\vec{r}_i$.

Each methyl group has three entries in matrix $\hat{\mathbf{M}}$ (one per C–H vector), indexed with i, j, k in Equation (1.83). Averaging of the fast-rotating methyl is achieved by averaging the corresponding elements in vector $\vec{d}$ and matrix $\hat{\mathbf{M}}$, and setting their weights to $w_i = w_j = w_k = \frac{1}{3}$. Therefore, $\hat{\mathbf{M}}$ elements average as follows

$$M_{i1} = M_{j1} = M_{k1} = w_i(2x_iy_i + 2x_jy_j + 2x_ky_k) \ , \quad (1.84)$$

and the corresponding averaged reduced RDC

$$\langle d_i \rangle = \langle d_j \rangle = \langle d_k \rangle = w_i(d_i + d_j + d_k) \ . \quad (1.85)$$

Phenyl groups entries are treated in a similar way. Equivalent phenyl C–H are represented by entries p and q. The weighting factors are adjusted to $w_p = w_q = \frac{1}{2}$. The averaged $\hat{\mathbf{M}}$ matrix elements are

$$M_{p1} = M_{q1} = w_p(2x_py_p + 2x_qy_q) \ , \quad (1.86)$$

and the reduced couplings are re-defined as

$$\langle d_p \rangle = \langle d_q \rangle = w_p(d_p + d_q) \ . \quad (1.87)$$

Averaging of methylenic protons will be discussed in Chapter 3.

1.6 Treatment of Molecular Flexibility

Quite a number of the marketed drugs are based on natural products, typically metabolites extracted from plants, marine organisms or microorganisms, which provide chemical complexity and diversity to the drug development pipelines.[55] In principle, the optimization of the efficiency of a natural product benefits from knowledge of the bioactive (bound) conformation of the small molecule. Unfortunately, the 3D structure of the bound drug/receptor complex is not always available. In these cases, structural information of the free small molecule in solution may be helpful on the assumption that the bioactive conformation is expected to be favored in the free state, as this minimizes unfavorable energetic requirements for binding.[56,57] The best solvent to study the conformational preference of biologically active natural products is aqueous solution, where the molecule retains its flexibility as well as its protonation state.

With the tools presented previously, the analysis of the RDC of rigid molecules is almost straightforward,[58–61] provided there are enough experimental couplings, which is aided by the possibility of including the fast rotating groups or the unassigned methylenic groups in molecules with insufficient methine C–H vectors. However, when there is more than one conformation with an appreciable population, the analysis rapidly becomes unmanageable as, in principle, each conformation can have a different alignment tensor (i.e. a different orientational probability).

Early approaches to the flexibility problem mostly explored single bond rotations in very simple organic molecules. Among other approaches, two of them were based on the additive potential (AP),[62] or in the the maximum entropy (ME)[35] models. The main drawback of the AP methodology was the requirement of previous detailed information of the potential surface. The ME model does not require such an extensive previous information, but it furnished too flat potential surface distributions. A more modern approximation is the APME model, which overcame the *a priori* requirements of the AP method as well as the flatness of the potential surfaces from the ME model.[63,64] The hybrid APME approach also permits the combination of RDCs with other NMR restraints such as NOE-derived distances or dihedral angles derived from ^3J couplings. Nevertheless, application of these methods to more complex systems became clumsy.

The flexibility problem can be greatly simplified with the introduction of the *single-tensor approach*. This approach relies on the calculation of a single common tensor for all conformers, adding to the fit a parameter to account for the population weight (i.e.

molar fraction). For applying the single-tensor approach, two requirements must be assumed:

i) a jumping model between the possible conformers, with no intermediate states,

ii) decoupling of the internal and alignment dynamics, i.e. the conformational change does not change the global orientational preference of the molecule.

Based on the single-tensor approach, RDCs have been successfully combined with NOE and ^3J as restraints in molecular dynamics (MM) calculations of conformational equilibrium problems.[65–68] They have also been analyzed on their own using ensembles derived from MM calculations and other methods.[51,54,69]

The main controversy in the recent years regarding the use of the single-tensor approach was the problem of the common reference frame. When using a single tensor for several structures, it is needed to define the orientation of the different conformers with respect to a common frame of reference.

Based on early approaches to the problem by de Lange and Burnell,[70] Thiele and co-workers introduced the Eckart[71] transformation of atomic coordinates to superimpose conformers. When fulfilling the Eckart conditions, this superimposition decouples the global tumbling movement from the internal conformational exchange.[47] Note, however, that this decoupling stands for infinitesimal movements and, therefore, its applicability to large-scale conformational changes can be questioned.

Another possible transformation is the least-squares superimposition of atomic coordinates of the conformers. This is mathematically equivalent to the Eckart transformation if mass-weighted coordinates are used.[72] In practice, this means that least-squares superimposition of atomic-weighted coordinates as compared with the Eckart transformation results in negligible differences for molecules with homogeneous mass distribution.[73]

Using the single-tensor approximation and understanding that the experimental RDCs ($\langle \overline{D}_i \rangle$) are actually the weighted average of all conformers in solution, the inclusion of a population weighting factor p_i permits the simultaneous fit of more than one conformer

$$\langle \overline{D}_{ki} \rangle = \sum_{k=1}^{n} \sum_{i=1}^{N} p_i \frac{\kappa}{R_{ki}^3} \vec{r}_{ki}^{\,T} \hat{\mathbf{A}}_i \vec{r}_{ki} \ , \tag{1.88}$$

where k is the k^{th} RDC, n is the total number of RDCs available, p_i is the population (molar fraction) of the i^{th} conformer and N is the total number of conformers in the solution ensemble.

In the general case, the alignment tensor $\hat{\mathbf{A}}_i$ is different for every single structure, rising the number of unknowns (five per tensor) to a unmanageable problem.[74] By using the single-tensor approach, the number of unknowns reduces to $5 + (N - 1)$, where N is the number of conformers included in the fit and 5 stems from the number of independent components of the *sole* alignment tensor $\hat{\mathbf{A}}$ [67,68,75]

$$\langle \overline{D}_{ki} \rangle = \sum_{k=1}^{n} \sum_{i=1}^{N} p_i \frac{\kappa}{R_{ki}^3} \vec{r}_{ki}^{\,T} \hat{\mathbf{A}} \vec{r}_{ki} \ . \tag{1.89}$$

Following the derivation of Equation (1.61), Equation (1.89) transforms to

$$\begin{pmatrix} d_1 \\ d_2 \\ \dots \\ \dots \\ d_n \end{pmatrix} = \sum_{i=1}^{N} \left(p_i \hat{\mathbf{M}}_i \hat{\mathbf{A}} \right)$$

$$= \sum_{i=1}^{N} \left(p_i \hat{\mathbf{M}}_i \right) \hat{\mathbf{A}} \ , \tag{1.90}$$

where p_i is the population weight (molar fraction) of the i^{th} conformer and $\mathbf{M}_i$ is its geometry matrix. The alignment tensor $\hat{\mathbf{A}}$ is common to all conformers. Note that the equation system is not linear, given that p_i are also unknowns. This system is solved by a combination of grid search and constrained least-squares procedure (Levenberg-Marquardt) in the MSPIN software.[42,43]

1.7 Experimental methods to measure $^1J_{CH}$ and $^1D_{CH}$

In principle, any 1D or 2D NMR experiment used to measure one-bond C–H couplings may be useful for the measurement of $^1D_{CH}$ RDCs, as the $^1D_{CH}$ RDC manifests as a contribution to the total coupling $^1T_{CH}$. The most widely used experiments have been recently reviewed by Thiele.[33]

RDCs appear as an additional contribution, ($^1T_{CH} = {}^1J_{CH} + {}^1D_{CH}$), to the corresponding scalar coupling when the measurement is done in weakly aligned media. Dipolar couplings can be extracted as the difference $^1D_{CH} = {}^1T_{CH} - {}^1J_{CH}$. In weak-alignment conditions, $|{}^1D_{CH}| < |{}^1J_{CH}|$. As $^1J_{CH}$ is known to be positive, if the total coupling $|{}^1T_{CH}|$ is larger than the $^1J_{CH}$, then the RDC is positive. Conversely, if $|{}^1T_{CH}|$ is smaller than $^1J_{CH}$, the RDC $^1D_{CH}$ must be negative. Care should be taken if the sizes of the dipolar couplings are equal or larger ($|{}^1D_{CH}| \geq |{}^1J_{CH}|$) than the corresponding scalar

couplings, as the sign of $^1T_{CH}$ is usually unknown and, therefore, two values of $^1D_{CH}$ would be compatible with the observed splitting.

As can be easily followed, two sets of experiments have to be acquired. An isotropic experiment for extracting the scalar couplings, and an additional experiment in aligned media that provides the $^1T_{CH}$ couplings.

1.7.1 ^{13}C detected experiments. 1D ^{13}C gated-decoupled and 2D ^{13}C–^{1}H HETCOR

The simplest experimental approach that can be used for the measurement of RDCs is a 1D ^{1}H-coupled ^{13}C experiment. Usually, a gated decoupling scheme is used as this provides signal enhancement due to ^{1}H–^{1}H NOE during the relaxation delay.

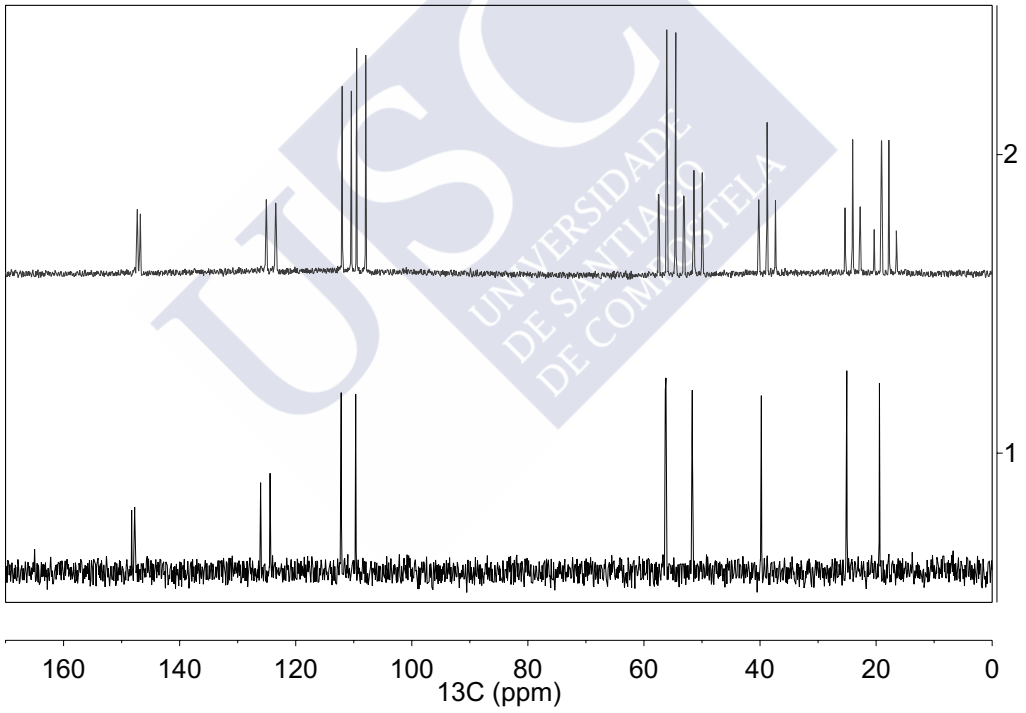


FIGURE 1.6: Comparison of 1D ^{13}C experiments. (1) bottom, regular, fully decoupled experiment; (2) top, gated-decoupled experiment displaying the $^1J_{CH}$ couplings. Note that the $^1J_{CH}$ splittings may generate signal overlap in crowded regions. Sample: salsolidine in isotropic D_2O solution. 500 MHz.

As coupling is observed in the direct acquisition dimension, a high digital resolution is achieved simply by using most of the relaxation time for FID acquisition. This high experimental resolution is impaired by the poor sensitivity (compared to proton-detected experiments) of room temperature probes for ^{13}C detection, even if a direct

probe is available (with the ^{13}C coil inside, like the BBO series), which results in long experimental times. One way to solve the inherent lack of sensitivity of direct ^{13}C detection is to use a cryoprobe.[76] A caveat about the cryogenically cooled probes —at least with Bruker cryo platforms— is that not all generations have a cooled carbon coil or dedicated preamplifiers for cool carbon channel.[3]

The main drawback of the gated decoupled experiment is the signal overlap typical of 1D spectra. Figure 1.6, shows the comparison of a regular decoupled-^{13}C experiment with a coupled (gated-decoupled) one. Notice that even with a simple molecule, several signals overlap in the coupled version of the experiment, difficulting the extraction of the information, especially with the anisotropic samples, which suffer from broader linewidths.

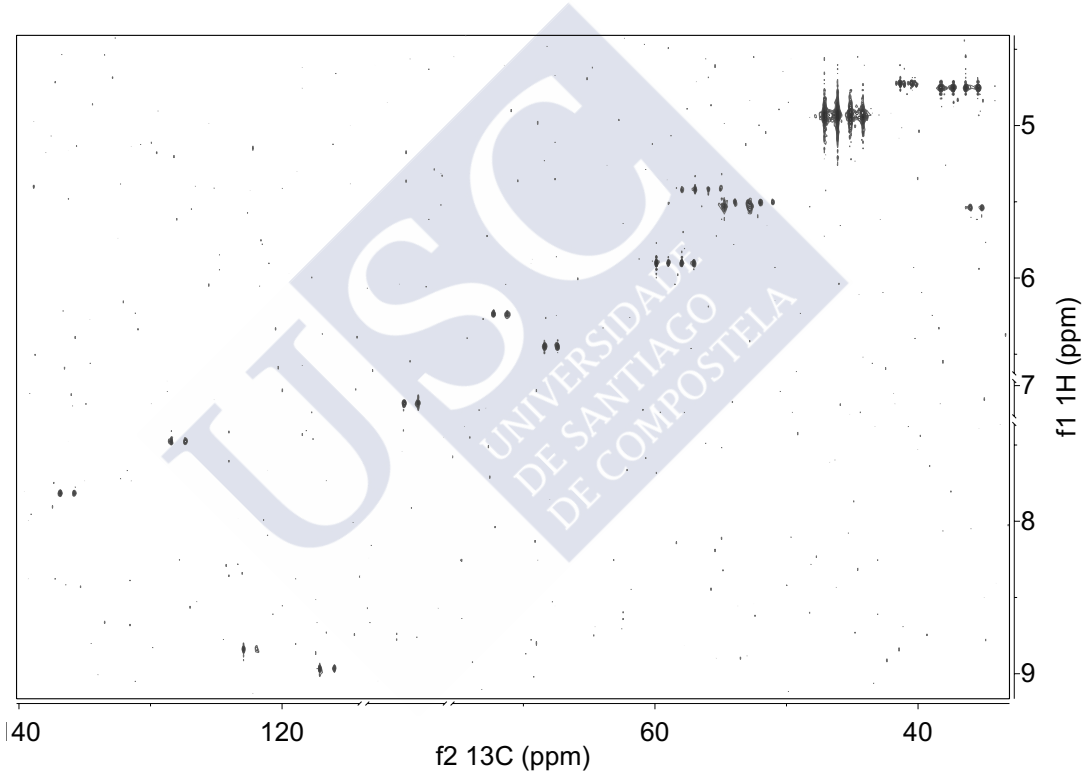


FIGURE 1.7: ^{13}C–^{1}H HETCOR anisotropic spectrum. Sample: methyl-codeine in isotropic D_2O solution. 600 MHz.

The problem of signal overlap can be alleviated by the use of the ^{13}C–^{1}H HETCOR 2D (Figure 1.7) experiment at the cost, however, of long experimental times.

[3]Specific information for every model can be found in the documentation. As a guideline, the reader is referred to the issue 157 of Bruker Spin Reports.

1.7.2 Heteronuclear Single Quantum Coherence (HSQC) experiments

Due to the lack of sensitivity associated with direct ^{13}C detection, inverse detected HSQC experiments are nowadays widely used. The HSQC experiment combines the sensitivity enhancement of ^{1}H detection with the dispersion of signal in two dimensions. There are many variants of the HSQC experiment, in which resonances can be proton-coupled in any of the two dimensions, in both or in none.

The Figure 1.8 shows a regular decoupled HSQC overlapped with the F_2-coupled version. $^{1}J_{CH}$ couplings can be measured from the ^{1}H-dimension (F_2) splittings.

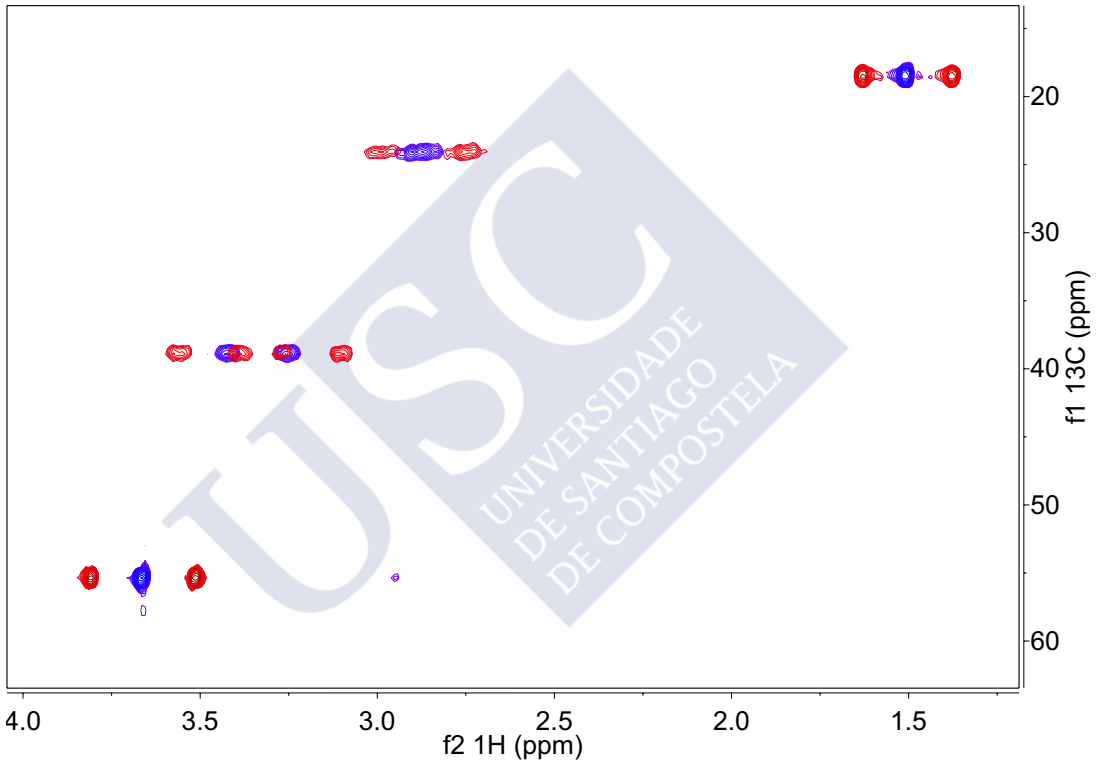


FIGURE 1.8: Comparison of a regular, decoupled, HSQC experiment (blue) and a F_2-coupled HSQC (red), in which $^{1}J_{CH}$ couplings can be measured from the ^{1}H dimension splitting of the peaks. Sample: salsolidine in isotropic D_2O solution. 500 MHz.

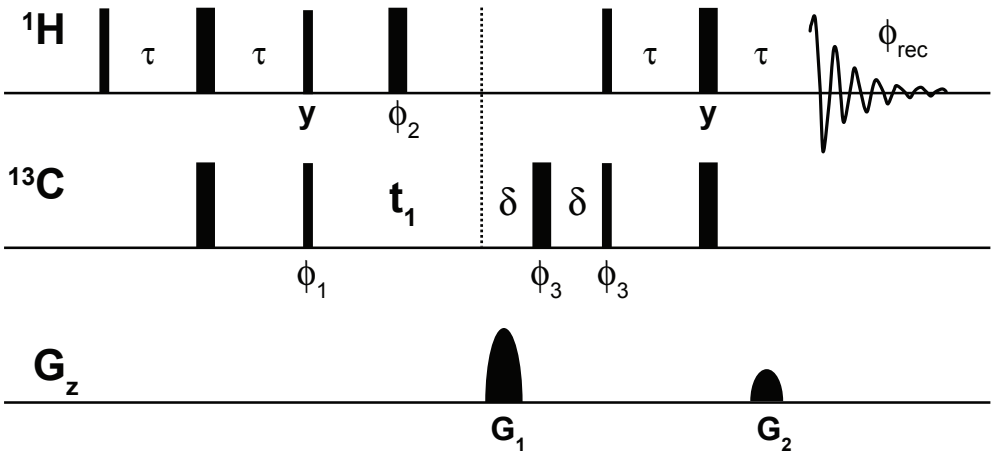


FIGURE 1.9: F_2-coupled HSQC pulse sequence. ^{13}C dimension pulses can be either adiabatic or hard pulses. $\tau = 1/(4J)$, INEPT delay adjusted usually to $J = 145$ Hz; $\delta = 1/(8J)$, multiplicity selection. Phase cycling: $\phi_1 = x, -x$; $\phi_2 = x, x, -x, -x$; $\phi_3 = x, x, -x, -x$; $\phi_{rec} = x, -x, -x, x$. Gradients power, as % of maximum power: $G_1 = 80$, $G_2 = 20$.

In an analogous way as in the 1D ^{13}C gated-decoupled experiment, proton decoupling must be switched off to maintain the scalar coupling information in the final detected Hamiltonian. The way in which the decoupling is performed originates two different coupled-HSQC experiments: F_1-coupled or F_2-coupled spectra. In the former, the $^1J_{CH}$ splitting occurs on the indirect (^{13}C) dimension, whilst in the latter, coupling evolves in the direct (^1H) dimension. Figure 1.14 shows an F_2- and an F_1-coupled HSQC experiments overlaid.

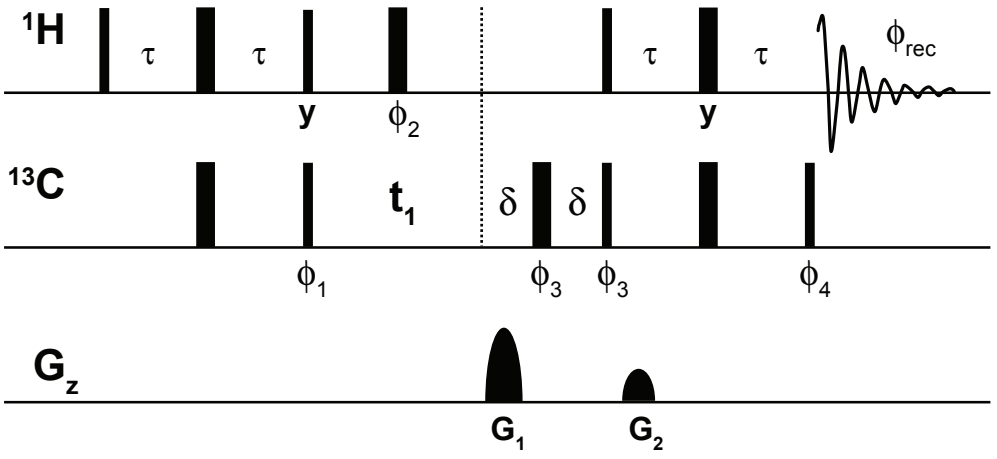


FIGURE 1.10: Clean inphase (CLIP) HSQC pulse sequence. ^{13}C dimension pulses can be either adiabatic or hard pulses. $\tau = 1/(4(J+D))$, INEPT delay adjusted usually to $(J+D) = 130-160$ Hz; $\delta = 1/(8J)$, multiplicity selection. Phase cycling: $\phi_1 = x,-x$; $\phi_2 = x,x,x,x,-x,-x,-x,-x$; $\phi_3 = x,x,-x,-x$; $\phi_4 = x,-x$; $\phi_{rec} = -x,x,x,-x$. Gradients power, as % of maximum power: $G_1 = 80$, $G_2 = 20$.

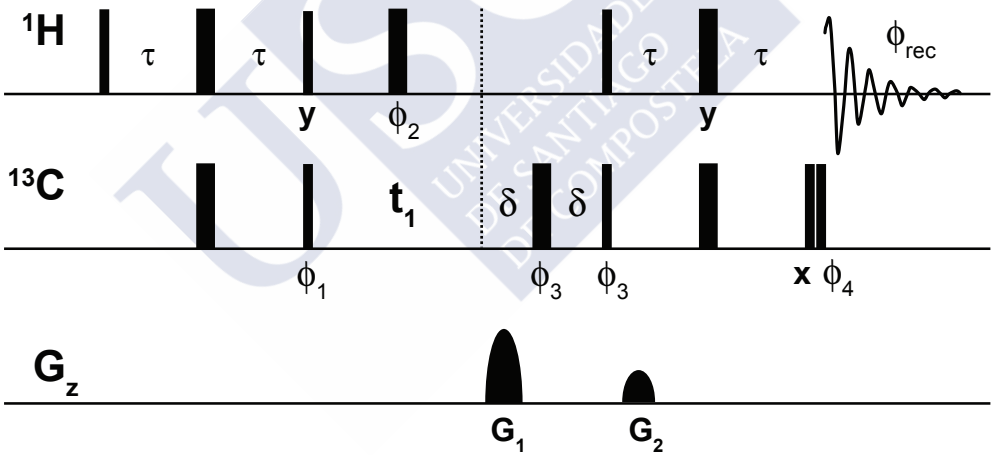


FIGURE 1.11: Clean antiphase (CLAP) HSQC pulse sequence. ^{13}C dimension pulses can be either adiabatic or hard pulses. $\tau = 1/(4(J+D))$, INEPT delay adjusted usually to $(J+D) = 130-160$ Hz; $\delta = 1/(8J)$, multiplicity selection. Phase cycling: $\phi_1 = x,-x$; $\phi_2 = x,x,x,x,-x,-x,-x,-x$; $\phi_3 = x,x,-x,-x$; $\phi_4 = x,x,-x,-x$; $\phi_{rec} = y,-y$. Gradients power, as % of maximum power: $G_1 = 80$, $G_2 = 20$.

One advantage of F_2-coupled HSQC is that absence of decoupling during acquisition permits longer detection times, resulting in good to excellent digital resolution in the coupling measurement dimension. Additionally, as this experiment can be acquired by "spectral sectors", and there is no need of high F_1 resolution, this very limited number of required increments results in short experimental times.

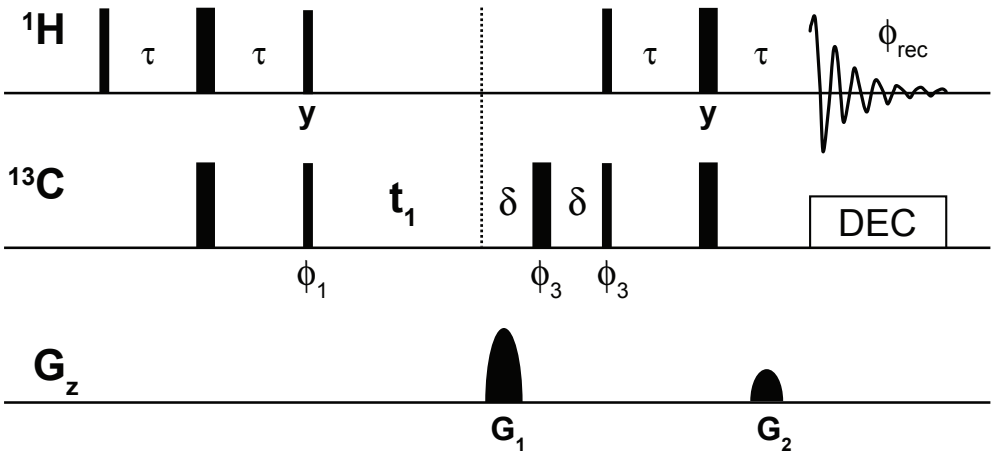


FIGURE 1.12: F_1-coupled HSQC pulse sequence. ^{13}C dimension pulses can be either adiabatic or hard pulses. $\tau = 1/(4J)$, INEPT delay adjusted usually to 145 Hz; $\delta = 1/(8J)$, multiplicity selection. Phase cycling: $\phi_1 = x, -x$; $\phi_2 = x, x, -x, -x$; $\phi_3 = x, x, -x, -x$; $\phi_{rec} = x, -x, -x, x$. Gradients power, as % of maximum power, $G_1 = 80$, $G_2 = 20$.

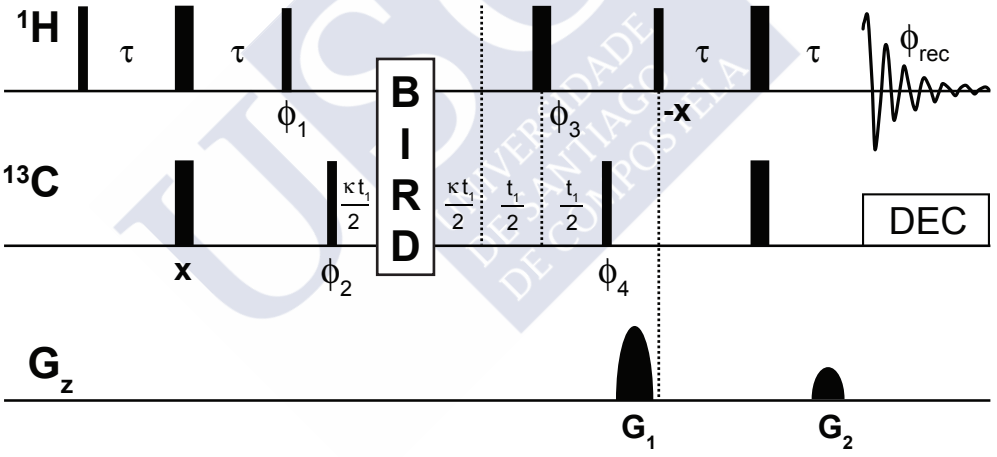


FIGURE 1.13: JE-BIRD^{d,X} HSQC pulse sequence. ^{13}C dimension pulses can be either adiabatic or hard pulses. $\tau = 1/(4J)$, INEPT delay adjusted usually to 145 Hz; κ, J evolution multiplication factor; BIRD^{d,X} delay adjusted to $1/(2J)$; $\delta = 1/(8J)$, multiplicity selection. Phase cycling: $\phi_1 = y$; $\phi_2 = x, -x$; $\phi_3 = x, x, -x, -x$; $\phi_4 = x, x, x, x, -x, -x, -x, -x$; $\phi_{rec} = x, -x, x, -x, -x, x, -x, x$. Gradients power, as % of maximum power, $G_1 = 80$, $G_2 = 20$.

It is important to notice that C–H couplings of non-equivalent methylenic protons can be individually extracted. The main drawback of this proton-dimension coupled experiment is the linebroadening in the measurement dimension, especially in the aligned sample. This becomes really problematic when several methylene groups are present in the molecule, due to the extensive dipole-dipole homonuclear coupling that evolves during acquisition, and the frequent occurrence of strong-coupling artifacts.

D_{CH} couplings can be either positive or negative, producing thus a great spreading in the values of T_{CH}, which should be matched by the INEPT transfer delays (τ, or $\Delta/2$) in the experiment. This is obviously not possible in a single experiment, and hence the mismatched delays lead to dispersive antiphase contributions. Experimentally, these antiphase terms can be removed by converting them to multi-quantum terms prior to detection, with a $\pi/2$ pulse in either the ^{1}H or ^{13}C dimensions with the same phase as the in-phase detected magnetization (CLIP-HSQC, Figure 1.10). Additionally, the acquisition of the corresponding anti-phase version (CLAP-HSQC, Figure 1.11) even allows to extract the individual couplings of overlapped diastereotopic protons in an IPAP manner.

An alternative solution for the above stated problems is the use of F_1-coupled HSQC experiments (Figure 1.13), which are specially useful in the case of not-so-weakly aligned samples in which homonuclear dipole-dipole coupling can cause severe line broadening.

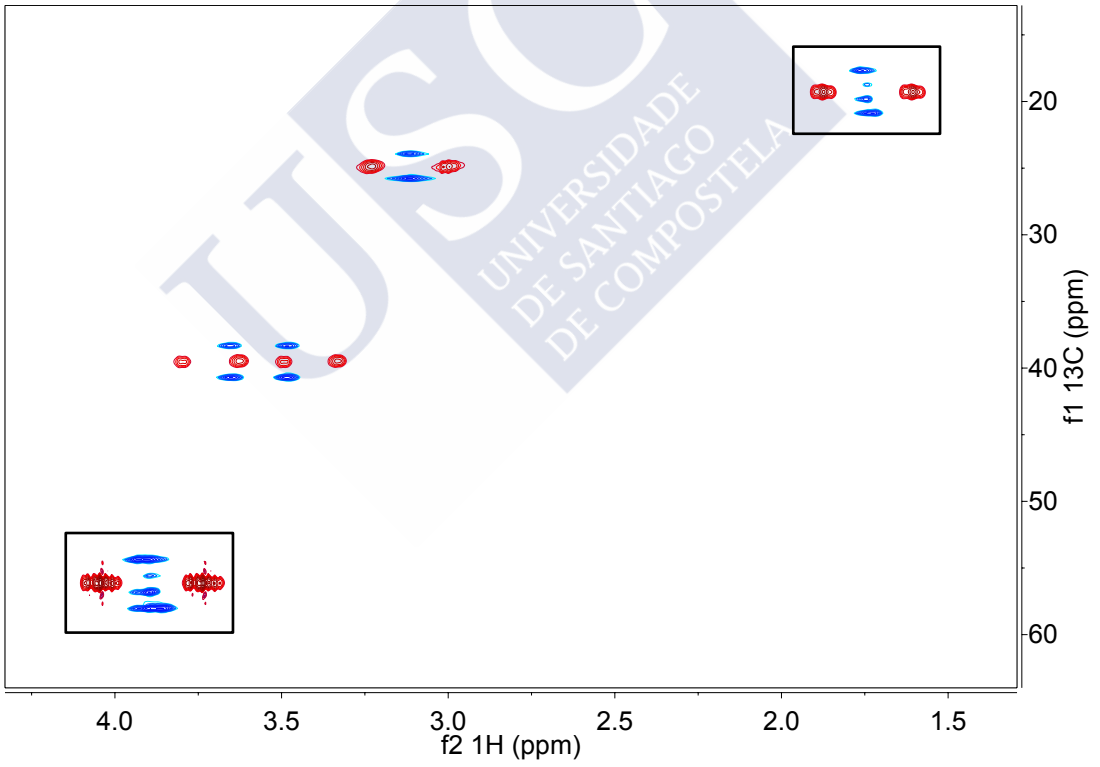


FIGURE 1.14: Overlay of a F_2-coupled HSQC (in red), and a F_1-coupled HSQC (in blue) of the same sample used in Figure 1.8. The same region is shown. Differences of linewidth in the coupled dimension can be appreciated. Note the deformed central lines of the more intense multiplets in the F_1-coupled HSQC (black boxes). Sample: salsolidine in isotropic D_2O solution. 500 MHz.

Along this thesis, and in most of works done in our group, the alternative F_1-coupled HSQC is employed. In general, the F_1-coupled version requires longer acquisition times as it needs high resolution in the indirect dimension. Due to the antiphase term introduced by evolution of passive coupling, only the external peaks from the methylene multiplet are finally observed, since the internal peaks pretty nearly cancel out. As a result, it is more convenient to extract the *sum* of the two $^1J_{CH}$ couplings as the total splitting of the outer lines. Chapter 3 describes how this is dealt with in the calculations with no further need of stereospecifically assigning the geminal protons. Likewise, the outer peaks of methyl multiplets are much more intense than the inner ones, facilitating the extraction of the coupling as $3\times$ $^1J_{CH}$. The main benefit from F_1-coupled experiments is the reduced line broadening of the signals when extensive dipolar coupling is present.

More sophisticated coupled experiments, like the *J*-scaled BIRD-HSQC experiment can be used (JE-BIRD^{d,X} HSQC, Figure 1.13).[77] The *J*-BIRD experiment contains a *J*-evolution module that can be scaled by a factor κ ($= 3$, typically) and a bilinear rotation (BIRD) element during the *J* evolution. *J* scaling permits an increment on the actual resolution without extending the experimental time, and the BIRD module reduces the linebroadening in the measurement dimension due to long-range heteronuclear $^nJ_{CH}$ couplings, which can be sizable in anisotropic media.

This is a summary of the typical experimental setup for these experiments:

i) 1D-gated decoupled ^{13}C usually is acquired to a total of 16K digital complex points and to a total of, at least, 1K scans. Typical acquisition time for the experiments shown in this dissertation was between 18 and 24 hours.

ii) F_2-coupled HSQC experiments were usually divided in sectors of 2 ppm ^1H spectral width, usually acquiring between 4K and 8K complex points in the direct dimension, with approximately $2-10$ points per ^{13}C ppm, and between 4 and 8 transients per t_1 increment. Typical acquisition time per experiment was between $40-60$ minutes, needing up to 4 quadrants for obtaining the full spectrum.

iii) F_1-coupled HSQC experiments needed up to 4K increments or $1.2-2$K increments when the *J*-scaling module was used. Typically, 1K points were acquired in the ^1H dimension, with $4-8$ scans per t_1 increment. Typical experimental time was between 3 and 12 hours per spectrum.

1.8 Long-range ^{13}C–^{1}H RDCs

Before RDCs were available, the determination of the configuration and conformation of small molecules by NMR experiments in solution was principally achieved on the basis of "classical NMR parameters" such as NOE correlations and ^{3}J coupling constants. If the chain of short-range spin-spin interactions is interrupted —for example by a magnetically inactive linker—, often there is no way to determine the relative configuration of the two disconnected molecular fragments by conventional NMR observables. As RDC provide information on the relative configuration of remotely located moieties, they are a very valuable structural restraint in such systems.

Most of the published studies on the application of RDCs to the structural analysis of small molecules involved the use of only one-bond ^{13}C–^{1}H RDCs ($^{1}D_{CH}$), measured at natural isotopic abundance. These $^{1}D_{CH}$ values are relatively easy to measure, from either F_2 or F_1 ^{1}H-coupled HSQC spectra, where the observed splitting in the aligned sample $^{1}T_{CH}$ contains the dipolar coupling as an addition to the scalar coupling ($^{1}T_{CH} = {}^{1}J_{CH} + {}^{1}D_{CH}$). These $^{1}D_{CH}$ couplings can suffice to solve configurational problems in small rigid or semirigid molecules.

Dipolar couplings scale with $1/r^3$, where r is the internuclear distance, hence $^{2}D_{CH}$ and $^{3}D_{CH}$ couplings cover a range that is about an order of magnitude smaller than $^{1}D_{CH}$ values and, therefore, require a measurement precision that is much higher than that required for $^{1}D_{CH}$ values. Although the value of long-range $^{n}D_{CH}$ couplings is well recognized, and methods for measurement of long-range ^{13}C–^{1}H RDCs have been described,[69,78] their accurate measurement has proved challenging,[79] therefore limiting their general application.

In Chapter 5, a more complete introduction to Long-Range ^{13}C–^{1}H RDCs —including available experimental methods— will be given.

1.9 Weak Alignment Media

Measurement of RDCs of practical use for structural analysis requires that the solute molecule is weakly oriented (or aligned) in solution, such that only a residual part of the dipole-dipole interaction shows up as a small contribution to the splitting due to the $^{n}J_{CH}$ coupling. This weak orientation is induced by the so-called weak-alignment media, the size of the observed RDC being proportional to the degree of order induced.

After the initial introduction of water-compatible weak-alignment media for studies of biomolecular compounds by Tjandra and Bax,[34] RDCs were not used as extensively for the analysis of small molecules, mainly because of the aqueous nature of the alignment media developed for biomolecules. Development of new alignment media compatible with organic solvents was needed in order to extend the use of RDCs to the structural analysis of organic compounds.

The weak alignment of a molecule can be produced by the intrinsic properties of the molecule or can be induced by external means (the so-called alignment media). Some molecules have a sizeable magnetic susceptibility and orient in the external magnetic field. However, most of the molecules lack this property and require an alignment medium, i.e. something external to the molecule that hampers its isotropic reorientation in the liquid state. Weak-alignment media can be classified in terms of the underlying alignment mechanism as auto-aligned media (lyotropic liquid crystals) and strained aligning gels. A liquid crystal is called lyotropic if it produces some degree of molecular order upon the addition of the solvent. Usually, this term was used to define phases composed by amphiphilic molecules. Strained aligning gels are polymeric gels in which the formation of anisotropic cavities can be induced.

Auto-aligned, liquid crystal media were first introduced in the late 1960s by Panar and Phillips in the form of poly-γ-benzyl-L-glutamate (PBLG).[80] In liquid crystals, the molecules that constitute the alignment media orient preferably respect to the external static field driven by their magnetic susceptibility. The same PBLG liquid crystal was used in the mid 1990s by Courtieu and co-workers for the enantiomeric discrimination of small molecules.[81] More recently, Griesinger and co-workers,[78] and Thiele and Berger,[82] made important contributions to the small molecule structural elucidation methodology by RDCs using the PBLG liquid crystal prepared in $CDCl_3$, and other organic solvents.

Thiele and co-workers presented the enhacement of the orienting properties of PBLG by means of additives.[83] Since the development of the PBLG liquid crystal, other liquid crystalline phases compatible with organic solvents have been developed, such as poly-γ-ethyl-L-glutamate (PELG)[58] and 4-*n*-penthyl-4′-cyanobiphenyl-d_{19} (PCBP).[84] Furthermore, liquid crystal systems compatible with water have been used for structural determination of biological organic molecules by RDCs, such as CPCl / hexanol / NaCl, employed for the alignment of lactose[85] and the small metabolite salsolidine[54] (see Chapter 3). The bicelles formed by DMPC / DHPC were employed by Martín-Pastor and Bush for the structural determination of the pentasaccharide lacto-*N*-fucopentaose 2.[86]

Strained aligning gels (SAGs) were developed by Ishii and co-workers[87] and Grzesiek and co-workers[88] in year 2000 based on the previous work by Deloche and Samulski[89] in the early 1980s. Strained aligning gels for the alignment of small molecules have not been developed until recent years. The first example of a polymeric gel applied to small molecules was presented by Luy and co-workers in 2004,[90] demonstrating that small molecules can be aligned in organic solvents using stretched polystyrene (PS) sticks. In a clever approach, very different from the one applied with proteins, a small cross-linked PS stick was introduced on the bottom of an NMR tube together with $CDCl_3$. Once the polymer reached the tube walls, the swelling continued in the axial direction only, resulting in anisotropic cavities that originated the orientation of the solvent and of the molecule under study. In the following years, several publications of the same group reported a wide variety of polymeric gels: polydimethylsiloxane (PDMS),[91] poly(vinyl acetate) (PVAc),[92] and polyacrylonitrile (PAN).[93] In recent years, other groups also presented aligning gels suitable for small molecules and compatible with organic solvents, such as the polymethylmethacrylate (PMMA) gel by Gil and co-workers,[60], and the acrylamide-based PMMA-acrylamide gels, by Griesinger and co-workers.[94] In addition, devices that reversibly stretch[95] and compress[96] such swollen polymer gels have recently been developed, thus permitting the collection of rapidly tunable RDC values. The introduction of water-compatible alignment gels was very much needed to take full advantage of RDC for the structural determination of water-soluble molecules, such as natural products and bioactive compounds. Water-compatible aligning gels have been developed, such as the popular german sweet "gummibärchen", presented as chiral alignment media by Luy and co-workers,[97] or the acrylamide-based gel PMMA-acrylamide[94], which was not probed in water solutions until 2012.[98] In principle, the structure of aqueous gels is less sensitive to interactions with solutes, making them, at first sight, more promising as wide-purpose alignment media.

Besides being widely used in macromolecular structural studies,[99] residual dipolar couplings (RDCs) have also proved to be a very powerful tool for the structural investigation of small organic molecules, either for constitutional,[45] configurational,[58,68,69,92,100] or conformational problems.[47,101]

Both acrylamide-based and cross-linked polystyrene polymeric gels used the strained aligning gel (SAG) approach, in which the gel is polymerized in a chamber diameter equal or larger than that of the NMR sample tube. After polymerization, most of the solvent is removed from the gel piece, e.g. by drying. The resulting dried stick is then placed in the bottom of an NMR tube and allowed to swell upon addition of solvent. First, the gel swells in both dimensions (axial and radial). Once the swelling gel reaches the walls of the tube, it can only grow further in longitudinal direction,

thus generating an *anisotropic* shape of the cavities in the direction (on average) of the tube main axis (uniaxially anisotropic system), which is collinear with the external $\mathbf{B_0}$.

Interestingly, the direction of alignment (relative to $\mathbf{B_0}$) can be changed by mechanically compressing the gel rather than stretching it.[96] Accordingly, the degree of alignment is independent of the field strength, and can be modulated by modifying the swelling conditions (e.g. solvent or additives) or the synthesis of the gel (e.g. concentration of cross-linker or dimensions of the cast chamber). The possibility of *scaling* the alignment strength and direction is one of the main advantages of the SAG approach over liquid crystals (for a discussion on liquid crystals see Chapter 3).

More detailed information on the used alignment media can be found in the corresponding chapters. Table 1.1 is composed as a quick reference for the reader, gathering the key features of selected alignment media.

Table 1.1: Weak-alignment media most used for the measurement of RDCs.

alignment media	solvent[a]	swelling[b]	$\lvert\Delta\nu_Q\rvert$	stability (months)[c]
Lamellar Phases	D_2O	Axial	18	$3-6$
	DMSO-d_6	Radial		
	$CDCl_3$	Plunger		
PAA gel	D_2O	Axial	20	$6-12$
	DMSO-d_6	Radial	6	>6
	$CDCl_3$	Plunger		
PMMA gel	D_2O	Axial		
	DMSO-d_6	Radial		
	$CDCl_3$	Plunger	30	>12
PDMS	D_2O	Axial		
	DMSO-d_6	Radial		
	$CDCl_3$	Plunger	20	>6

Each class of alignment media is divided in an horizontal block showing its solvent compatibility, swelling conditions, average solvent $\lvert\Delta\nu_Q\rvert$ splitting, and average stability in our hands.
[a] Solvent compatibility refers only to D_2O, DMSO-d_6 and $CDCl_3$, as the most representative solvents; for detailed information about solvent compatibility, please see the corresponding chapter.
[b] Swelling conditions are classified as *Axial* (for a fully axial growth gel; in the case of PDMS, this is shaded in gray as most of its growth is radial) or *Radial* (mostly for axially-restrained gels); *Plunger* indicates that it is mandatory (black text) or strongly recommended (dim-gray text) the use of a shigemi plunger to avoid excessive axial growth.
[c] *Stability* refers to the period in which a gel, with a test molecule diffused into it and stored at room temperature, maintains invariant the $\lvert\Delta\nu_Q\rvert$ of the solvent, the chemical shifts and the linewidths of the molecule resonances.

1.9.1 Alignment gels that swell in hydrophobic organic solvents

Water compatible polymeric gels have been used for orienting biomolecules in water since year 2000 after their introduction by Tycho *et al.*[87] and Grzesiek and co-workers,[88] based on the initial work by Deloche and Samulski on elastomeric polymers.[89] Luy and co-workers boosted the development of gels compatible with organic solvents, such as polystyrene,[90] PDMS,[91] or PAN gels.[93]

1.9.1.1 Cross-linked polystyrene, PS

Stretched polystyrene gel was introduced by Luy and *et al.* in 2004.[90] This gel represented the first example of the SAG approach applied to small molecules for the measurement of anisotropic NMR parameters in aligned media. Cross-linked polystyrene gels are compatible with a wide range of organic solvents, including chloroform, benzene and tetrahydrofuran (THF).[102]

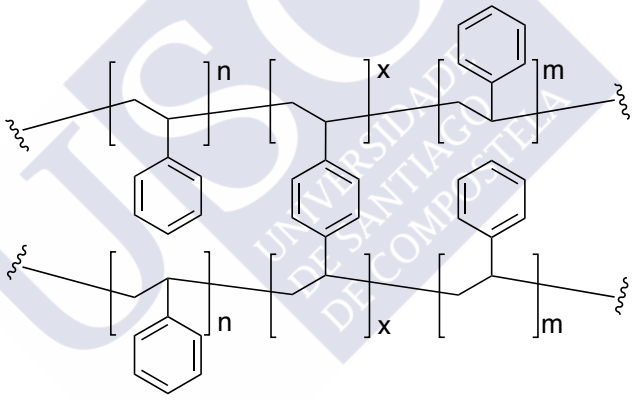


FIGURE 1.15: Chemical structure of the cross-linked Polystyrene polymer. [vsp]cambia analogamente todos los caption; no confundamos polimero con gel; el gel es mezcla liquido/malla

Cross-linked polystyrene (PS) is obtained by polymerization of appropriate amounts of styrene *monomer* and divinylbenzene (DVB) *cross-linker*, adding dibenzoylperoxide (DBP) or 2,2′-azobis(2-methylpropionitrile) (AIBN) as *radical initiator*. By tweaking DVB concentration, the total solvent ^{2}H quadrupolar splitting ($|\Delta\nu_Q|$) can be scaled from less than 20 Hz up to almost 600 Hz. Interestingly, the ^{2}H splitting can also be scaled by modifying other polymerization variables, like the radical initiator (DBP) concentration or the temperature.

Cross-linked polystyrene gels swell up in few days and molecule diffusion needs only few hours. Unfortunately, this gel is microscopically very heterogeneous, which results in some inconsistency in the alignment behavior of gel pieces polymerized as

part of the same batch and in equivalent conditions.[102] Thanks to the chemical stability of PS gels, they have recently been applied to the alignment of organolithium compounds by Stalke and co-workers.[103]

1.9.1.2 Cross-linked poly(dimethylsiloxane), PDMS

Cross-linked poly(dimethylsiloxane) (PDMS) was introduced as alignment media for orienting small organic molecules by Freudenberger *et al.* by cross-linking PDMS with β-radiation.[91] PDMS is recurrently used as a model for elastomers and it is known that stretched PDMS gels cause some degree of anisotropy.

Equilibration of these gels occurs in less than two weeks and sample diffusion needs only two days. Due to its high stickiness to glassware, $N_2^{(L)}$ freezing of the polymer stick is needed previously to placing it in the bottom of an NMR tube. It is crucial to push the gel to the bottom of the tube until it defrosts and sticks to the tube wall.

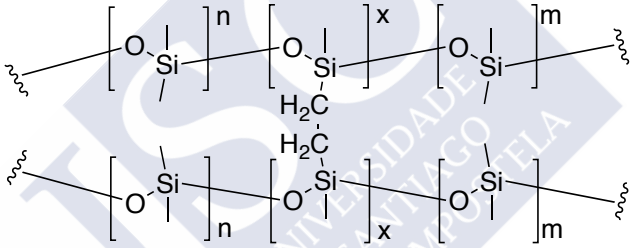


FIGURE 1.16: Chemical structure of the cross-linked PDMS polymer.

As stated by Luy and co-workers,[91] the main advantage of cross-linked PDMS gels is the almost exclusive steric origin of the alignment. Furthermore, the degree of alignment induced can be easily scaled by the dose of radiation applied for the cross-linking. Interestingly, the alignment induced by PDMS is stronger with less polar solvents, hence it is ideally suited for molecules soluble in apolar solvents.

1.9.1.3 Cross-linked poly(vinyl acetate), PVAc

Stretched poly(vinyl acetate) was introduced as a gel compatible with polar organic solvents by Luy and co-workers.[92] This gel can swell in a wide variety of organic solvents, as dimethyl sulfoxide (DMSO), methanol, THF, dioxane, acetone, acetonitrile, and chloroform. Polymerization is conducted in an similar way as with cross-linked PS gels. Appropriate amounts of *monomer* vinyl acetate and *cross-linker* adipic acid divinyl ester are mixed in the presence o a *radical initiator*, resulting in a matrix showed in Figure 1.17.

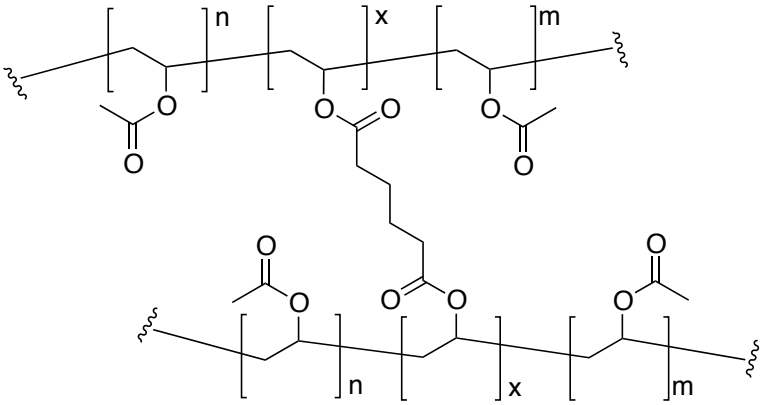


FIGURE 1.17: Chemical structure of the cross-linked PVAc polymer.

This gel equilibrates completely in less than two weeks, and only two days are needed for the diffusion of the molecule. As in the previous examples, alignment degree and $|\Delta\nu_Q|$ scale with the concentration of cross-linker used. Interestingly, the authors reported a significant difference in the alignment tensors of the same molecule in different solvents.[92] This was ascribed to specific interactions of the polymer-solvent or molecule-solvent pairs.

1.9.1.4 Cross-linked poly(acrylonitrile), PAN

The demand of uncharged alignment media compatible with cyclic peptides and, consequently, compatible with a wide range of molecules, induced Kummerlöwe *et al.* to start a systematic search of polymers that swell in DMSO with the least possible NMR signals.[93] Poly(acrylonitrile) showed the desired solubility and almost lack of ^{1}H NMR signals. Polymerization is carried out in the same way as PDMS, by accelerated electrons cross-linking (Figure 1.18). The irradiative cross-linking avoids the presence of undesired NMR signals from the cross-linking reagents or by-products.

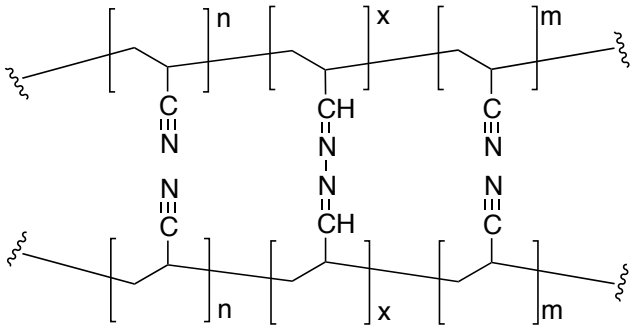


FIGURE 1.18: Chemical structure of the cross-linked PAN polymer.

In the same way as the other polymeric gels, cross-linking modification results in PAN gels with different alignment strengths. $|\Delta\nu_Q|$ splittings between 3 and 40 Hz have been reported. The main downside of the PAN gels are shared with PDMS, which is the need of an electron irradiation source. Furthermore, irradiation of the PAN is conducted in the NMR tube and the subsequent re-equilibration of the gel takes several weeks. Anyhow, sample diffusion takes only two days.

1.9.1.5 Cross-linked polymethylmethacrylate, PMMA

Polymethyl methacrylate gels were developed by Gil and co-workers.[60] Some characteristics were pursued over the development of these gels:

i) Cheap and easy to prepare

ii) Avoid the presence of aromatic rings to minimize chemical shift disturbance (see PS gels in Section 1.9.1.1)

iii) Rigid polymer backbone

Polymer rigidity is always advantageous to attain fast T_2 relaxation rates of the polymer protons. This minimizes the residual resonances from the gel masking the solute resonances. Fast T_2-relaxing polymer residual resonances can be further suppressed with a CPMG block (T_2 relaxation filter) with very short delays ($30 - 50$ ms).[104,105]

Non-compressible PMMA gels. The PMMA gel was initially developed to be used as an axially stretched gel (SAG method).

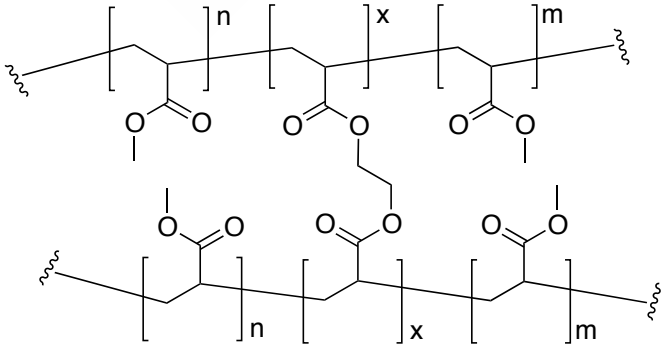


FIGURE 1.19: Chemical structure of the cross-linked PMMA polymer (*Figure kindly provided by Prof. Roberto R. Gil*).

Preparation of the gel is straightforward. Followed by the mix of appropriate amounts of *monomer* methylmethacrylate (MMA) and *cross-linker* ethylene glycol dimethacrylate

(EGDMA), the polymerization is carried out in a disposable NMR tube initiated by V-70 (2,2′-azobis(2,4-dimethyl-4-methoxyvaleronitrile)). The reaction is accomplished at 50 °C using acetone-d_6 as solvent to prevent the residual solvent peak in the gel (see Figure 1.19 for the final polymer structure).

The typical procedure for swelling PMMA gels is to place a dry gel stick (10 mm long) in the bottom of an NMR tube in which one drop of chloroform was previously added. When the gel starts to swell and reachs the walls of the tube, the rest of the solvent is added, dropwise as this polymer floats in chloroform. Following this procedure, PMMA gels reach their maximum length in about *30* days. The total ^{2}H quadrupolar splitting ($|\Delta\nu_Q|$) can be tuned from 20 to almost 140 Hz by altering the EGDMA molar fraction. The equilibration time can be reduced by starting the swelling with a volume of chloroform such that the gel floats at half the tube active volume (i.e. about 8 mm from the bottom of the tube). This alternative procedure reduces the swelling time to $7-10$ days.

Even though the swelling procedure can reduce the equilibration time of the gel, the main drawback of PMMA gel is not overcome. This is the presence of residual monomer in the gel, which results in some peaks that can overlay with the ones from the molecule of interest (Figure 1.20).

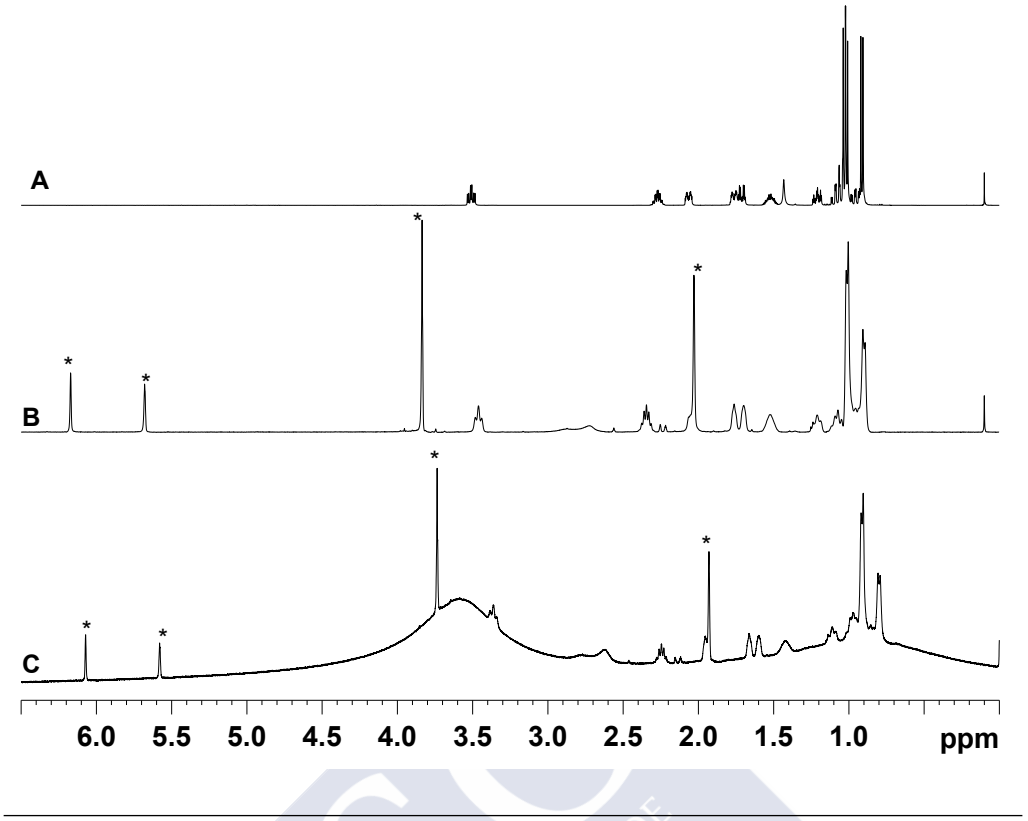


FIGURE 1.20: ^{1}H NMR spectra of menthol showing the residual MMA monomer peaks (labeled with *) of PMMA gels. Solvent is $CDCl_3$. (A) Menthol in neat $CDCl_3$. (B) Same sample as in (C) recorded with 30 ms CPMG filter, the broad polymer proton resonances vanish while the sharp monomer resonances remain. (C) Menthol dissolved in the gel; no CPMG filter (*Figure kindly provided by Prof. Roberto R. Gil*).

The residual monomer can be eliminated by extensive washing of the freshly made gel with chloroform. Swelling rate in pure chloroform outside an NMR tube is so fast that drives the gel to crack and disintegrate. This can be avoided by swelling in acetone/chloroform mixtures, increasing the chloroform ration as the gel swells up. When completely swollen, it can be extensively washed in neat chloroform. To dry the gel, first chloroform is replaced with acetone in an inverse procedure as the swelling. Most of the gel pieces break up during this procedure.

Compressible PMMA gels. PMMA gels show very good alignment properties for organic (small) molecules, despite the drawbacks presented in the previous section regarding their use with the SAG approach:

i) long swelling times , which can result in sample decomposition

ii) longer than desirable diffusion time of solute into the gel (72 h)

iii) persistence of residual MMA monomer

A slight variation of the SAG method consists in controlling the axial expansion of the gel with a plunger. The best setup comprises a few changes with respect to the SAG setup, mainly related with the dimensions of the dry polymer sticks. For the SAG embodiment, PMMA was polymerized in 5 mm diameter tubes, and trimmed to 10 mm length. In contrast, for the compressible gel approach dry polymer pieces are cut to match the full length of the final NMR sample (typically 25 mm) such that the polymer will be allowed to swell only in the *radial* dimension due to the confinement imposed by the plunger. Furthermore, PMMA is polymerized in casts of narrower diameter (3 mm disposable NMR tubes), resulting in 2 mm diameter dry sticks.[96] The gel swells upon addition of solvent ($CDCl_3$) until it fills all the available space (Figure 1.21).

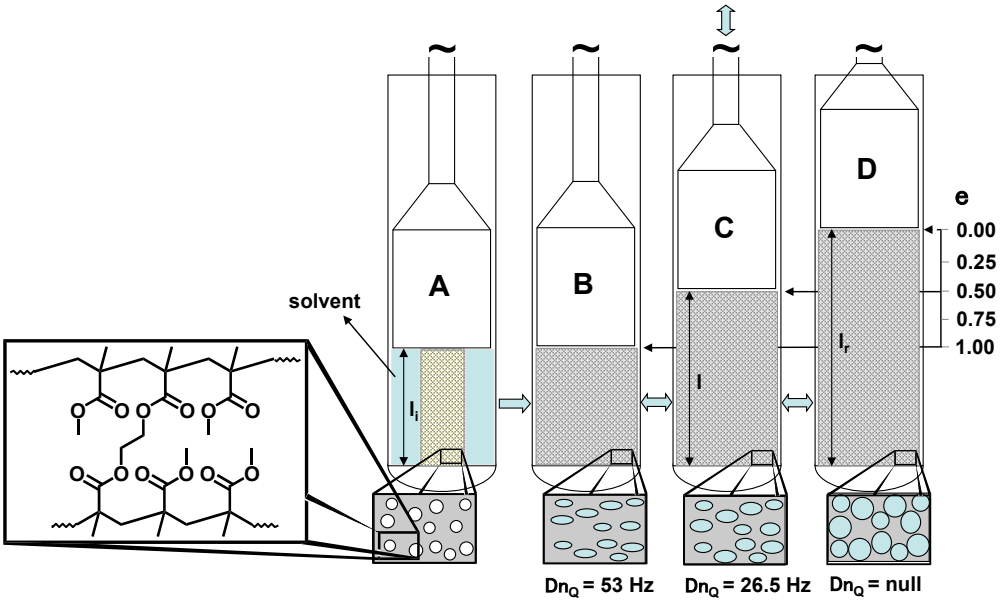


FIGURE 1.21: Compressible PMMA gel, showing the effect of the compression over the gel matrix cavities (*Figure kindly provided by Prof. Roberto R. Gil*).

A key feature of this method is that it makes possible to wash unwanted by-products from the gel once swollen in the NMR tube. By gently pushing the shigemi plunger back and forth, and successively replacing the chloroform in the tube with fresh portions , the gel can be completely washed out from the MMA residual monomer. Likewise, this method makes possible to diffuse the sample inside the gel in a few minutes. Furthermore, the solute can be recovered by the same procedure, which opens the way to use the same gel sample to measure RDCs of several molecules. Typically, the gel is ready for measurement in only 48 hours, including swelling, washing and sample diffusion.

Another useful feature of this embodiment is that the alignment strength of the gel can be easily tuned by moving the plunger up and down (Figure 1.21). Interestingly, the solvent quadrupolar splitting $|\Delta\nu_Q|$ depends linearly on the degree of compression.[96][4] In addition, the possibility of easily modifying the alignment strength permits the measurement of the same molecule inside the same readily prepared gel at various degrees of compression (different degrees of alignment) to determine the coupling sign. This behavior reflects that molecular alignment at the microscopic scale is directly related with gel cavity shape.

1.9.2 Water-Compatible gels

In Section 1.9.1, some of the most widely used polymeric gels were introduced. In the following, some alignment gels suitable for aqueous samples are presented. This section is shorter due to the less intense effort exerted by the community for the development of water-compatible alignment media for small molecules.

1.9.2.1 Gelatin

Focusing on the problem of enantiomeric discrimination, for which only liquid crystal were available, Luy and co-workers presented gelatin as a chiral alignment media in the form of the popular German sweets *gummibärchen*.[97] Proof of concept was done with rod-like cut-dried-extensively washed *gummibärchen* resulting in a $|\Delta\nu_Q|$ in the order of $\approx$ 20 Hz. After proving the concept, the procedure was optimized by the preparation of 10% w/v gelatin solutions. Gelification of this heated solution was done on pipette tips and furnished a solvent $|\Delta\nu_Q|$ in the order of $\approx$ 120 Hz.

1.9.2.2 AMPS-Acrylamide Gels

The acrylamide-based copolymer gel AMPS-acrylamide was introduced by Haberz *et al.* for the measurement of RDCs of small molecules in DMSO and aqueous solutions,[94] offering an opportunity to determine RDC of polar molecules (e.g. natural products) not soluble in hydrophobic organic solvents.[66–68,75,94,106,107]

Cross-linked acrylamide gels have been extensively used in biochemistry laboratories for decades. Based on the initial approach by Deloche and Samulsky,[89] the

[4] Trigo-Mouriño, P., Troche-Pesqueira, E., Gil, R.R. manuscript under preparation

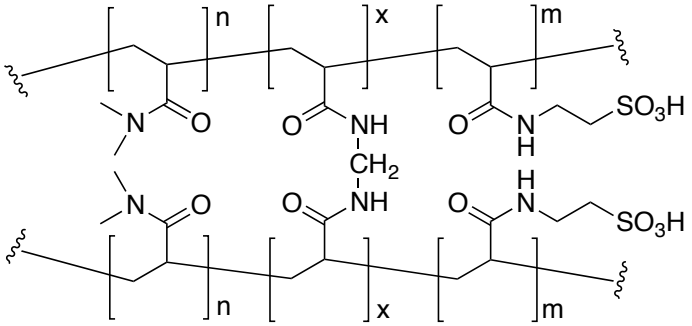


FIGURE 1.22: Structure of the cross-linked AMPS-acrylamide copolymer . Sulfonate groups of AMPS are shown in their non-dissociated form.

groups of Tycko and Grzesiek introduced the application of cross-linked acrylamide-bisacrylamide gels —in concentration ranges similar to those used for protein electrophoresis— to the protein NMR field.[87,88] These gels provide enough alignment strength for proteins but not for small molecules. An option to provide acrylamide-based gels with stronger alignment strength would be to increase cross-linker concentration. Instead, based on work by Liu *et al.*,[108] Griesinger's group envisaged the possibility of increasing the alignment strength by adding a second co-monomer with a strong electrolyte side chain, namely 2-(acrylamido)-2-methylpropanesulfonic acid (AMPS). Dissociation of sulfonate groups was proposed to lead to the formation of ion pairs in this kind of gels. These ion pairs can further interact with each other forming *ion-pair multiplets* that establish additional cross-linking contacts in the gel matrix, thus providing enhanced alignment properties for small molecules.[94] More recently, Griesinger and co-workers presented a modification with AMPS-based chiral monomers (*R*)- and (*S*)-APhES furnishing enantiomer-discriminating gels compatible with water.[109]

1.10 Computational Methods

The software tools used in this thesis will be presented in the next.

NMR spectra analysis. NMR spectra for the extraction of NMR restraints were analyzed in TopSpin and MestReNova.[42]

Molecular Modeling. Molecular Mechanics modeling has been done either with Maestro interfaced Macromodel[110] or PCmodel.[111] DFT computations were done using the resources of Centro de Supercomputación de Galicia (CESGA), on the *Finis*

Terrae and *SVGD* clusters. Analysis of the output files was done with Openbabel[112,113] and in-house developed software.

Analysis of RDC and other NMR data. *J* couplings and RDCs were analyzed with the program Mspin .[42,43] Additionally, RDCs were also fitted with RdcFit. RdcFit is an in-house written program that performs the most important tasks related with the fit of RDC data. RdcFit has been coded by Dr. A. Navarro-Vázquez in Python language makes use of SciPy,[114,115] NumPy,[116] and Openbabel[112,113] libraries. Optimization of the alignment matrix function is done by means of the Powell minimization algorithm.[117]

RdcFit also performs Bootstrapping simulations, which were further analyzed with a set of *Bash*, *Python* and *AWK* scripts. Montecarlo Bootstrapping calculations were done at Centro de Supercomputación de Galicia (CESGA), on the *SVGD* cluster.

TEXt editor. This thesis dissertation was written in LaTeX$_{2\epsilon}$, using TEXmaker as editor. Bibliography and other additional files needed for the correct compilation of the main document were generated with TEXShop. The bibliography file in bibTEX format was edited and managed with BibDesk program.[5] The bibliographic style used was the original Angewandte Chemie style, compiled by Patrick W. Daly and obtained from the Comprehensive TEX Archive Network.

Figures. All bar plots, linear plots and scatter plots presented in this dissertation were generated in *Python* using matplotlib[118][6] and *SciPy* libraries.[114,115][7] All the NMR spectra figures were generated with with the program MestReNova.[42] All 2D chemical structures have been drawn with CS Chemdraw Ultra using the standard ACS templates. All 3D representations of the molecules have been generated with UCSF Chimera and MolMol,[119] using their POV-ray tracing options. Final conversion of figures and additional edition have been done with Adobe Illustrator and Photoshop CS5 .

[5]The reader can find a sample bibtex file and a LaTeXthesis template in the author's GitHub.
[6]The reader can fork the last version of the libraries from the project's GitHub repository.
[7]The reader can find sample scripts in the author's GitHub.

1.11 Objectives

This thesis illustrates the use of residual dipolar couplings measured in aligned media for the solution of structural problems with small molecules. In the Introduction we have seen that progress is needed in several directions: better aligment media, better NMR spectroscopic tools to obtain experimental restraints, and better data analysis tools. The three topics are addressed in this thesis and examples of applications to solve chemical problems are presented. The objectives are listed with cross-references to the chapter where related results are presented.

1. **Properties of gels as alignment media**. To study the alignment properties of gels with the aim of controlling the alignment of solutes for practical applications. The effect of experimental variables on the alignment of the model compound *N*-methylcodeine in AMPS-acrylamide gels is presented in Chapter 2.

2. **NMR spectroscopy tools**. To enhance the structural usability of RDCs by making accessible the measurement of more experimental restraints from spectra. More precisely, a protocol to use long-range RDCs will be presented.

   - in Chapter 5 a new experiment SJS-HSQC for the measurement of long-range C–H couplings is discussed and applied to determine the relative configuration of five stereogenic centers
   - in Chapter 6, long-range RDCs are used to study the solution conformational equilibrium of a new anti-obesity drug

3. **RDC analysis tools.** To devise protocols to introduce more restraints in the calculations with RDC data. There were limitations to the introduction of RDC data from fast-rotating averaging groups, as methyl or phenyl. Likewise, use of RDC from methylenes was also problematic. Additionally, fit of RDC to conformers in exchange equilibrium was also an issue. These topics are addressed in the following chapters:

   - inclusion of RDCs form averaging groups and unassigned methylenes
     - in Chapter 3, a new method for the inclusion of unassigned methylene RDCs is presented. This development is also used in Chapter 2, Chapter 4, Chapter 5 and Chapter 6
   - inclusion of long-range RDCs with proper weighting
     - in Chapter 5, a convenient scaling procedure is proposed
     - in Chapter 6, the scaling procedure of long-range couplings is further refined by the inclusion of angular penalty functions for the RDC fit

- description of conformational equilibria with the single-tensor approximation
  - conformational analysis in terms of conformational ensembles is the objective of Chapter 3 and Chapter 6
  - the method is also tested in Chapter 2, Chapter 4 and Chapter 5 to ensure the absence of conformational averaging

4. **To solve structural problems in chemistry taking advantage of RDC and other spectroscopic data**

   - Relative configuration
     - to determine the relative configuration of two natural product alkaloids (Chapter 4 and Chapter 5)
   - Absolute configuration
     - to determine the absolute configuration of a new isolated vinca alkaloid (Chapter 4) with RDCs in conjunction with chiroptical measurements
   - Conformational analysis
     - to determine the conformational state of two biological-active molecules in water solutions (Chapter 3 and Chapter 6)
     - in Chapter 6, the utility of long-range couplings to address conformational analysis is evaluated for the first time

Chapter 2

Characterization of the alignment properties of acrylamide gels

2.1 AMPS-acrylamide gel aligns small molecules in water

As explained in Chapter 1, AMPS-acrylamide gels were presented as NMR alignment media by Haberz *et al.*, providing a suitable gel for the structural determination of small polar molecules in aqueous solutions.[94] An attractive feature of AMPS-acrylamide gels is that they swell in water and in DMSO solutions. As these gels can be prepared easily in a laboratory with standard equipment, they can become a tool of general use also among non-specialists in NMR spectroscopy.

At present, it is no possible to predict which of the existing alignment media will be suitable to weakly orient a given molecule under study, hence trial and error is unavoidable (i.e. see Chapter 3 and Chapter 6). A deeper knowledge on the alignment process inside the alignment media is needed in order to choose the correct one, as well as to control how the alignment degree and/or other characteristics can be varied once the aligned sample is prepared. To attain this goal, deeper characterization of the properties of these gels is required. Therefore, we decided to explore the properties of AMPS-acrylamide gels.

The structure of the AMPS-acrylamide polymer is shown in Figure 2.3. The molar ratio of co-monomers *N*,*N*-dimethylacrylamide (DMAA) and 2-(acrylamido)-2-methylpropanesulfonic acid (AMPS) is in the range 1 : 3 to 3 : 1 DMAA:AMPS.[108,120,121] Bisacrylamide (BIS) is present in a small proportion (ca. $1-2$ %) and acts as a covalent cross-linker between polymer chains.

The strong electrolyte AMPS contains a sulfonate functional group that provides the gel with the capability of establishing cross-links between polymer chains *via* ion pairing. Therefore, the ratio between the co-monomers modifies the swelling behavior of the gel by affecting the charge density.[108] The charged nature of the BIS cross-linked DMAA-AMPS chains results in repulsive forces between chains with the same charge as well as binding of the (neutralizing) counterions to the gel matrix. The counterion-polyion association has been identified as the most essential factor driving the properties of these hydrogels.[120]

Due to their ionic nature, the electrostatic interactions within the gel matrix play a crucial role in the swelling behavior of the gel, thus influencing the alignment properties of the swollen gel. Two factors affecting the swelling properties of AMPS-acrylamide gels had been described by Liu *et al.*, namely, the dissociation of the sulfonate groups and the association of the counterions.[108,120,121] Solvent polarity was identified as a key factor modulating the dissociation process of the sulfonate groups. The impaired dissociation of the sulfonate groups with low polarity solvents leads to the formation of ion-pairs. Ion pairing reduces the concentration of mobile counterions, thus reducing the repulsive interactions inside the gel matrix and promoting the formation of ion-pair multiplets. Ion-pair multiplets generate attractive non-covalent forces between the polymer chains, promoting the shrinkage of the gel. When swelled in polar solvents (i.e. high dielectric constant), the dissociation of the sulfonate groups is enhanced, leading to strong repulsive forces in the gel matrix, thus promoting swelling of the gel.[108,121] Additionally, due to the ionic character of this hydrogel, the nature of the counterions is also important. Particularly, in a series of investigations on the effect of the counterion nature over swelling —testing Na^+, H^+, K^+ and Ca^{2+}—, Na^+ cation showed strong affinity for the dissociated sulfonate groups. Cation coordination with the dissociated sulfonate groups promote the formation of anion/cation aggregates. Therefore, the swelling properties of AMPS-acrylamide gels can be modulated by additional cross-linkings depending on the presence of counterions and on the polarity of the solvent.[108,120,121]

These factors have to be considered when using AMPS-acrylamide gels to align molecules for NMR measurements, as molecular alignment depends very much on the average size and shape of the gel cavities (see Section 1.9). The geometry of the gel cavities —in these gels— is a consequence, primarily, of the extent of swelling (causing the shape anisotropy due to constraints imposed by the NMR tube) and the degree of cross-linking between polymer chains (which defines the size of the cavities). Cross-linking can be covalent (due to BIS) and non-covalent (formation of ion-pair multiplets as described above). It is worth noting the twofold effect of formation of stronger ion multiplets: defining smaller cavities and impairing overall swelling. Furthermore,

molecular alignment does not depend only on the geometry of the alignment medium, but also on electrostatic interactions (its charge distribution). In other words, molecular alignment can be seen as the resultant of two components, steric and electrostatic. Therefore, the solvent dielectric constant and ion concentration are expected to modulate molecular alignment in AMPS-acrylamide gels by the two mechanisms, i.e. by influencing the geometry of the gel (steric component) and by changing the spacial charge distribution (electrostatic component).

Based on previous findings of electrostatic modulation of the orientational properties of alignment media, we foresaw an opportunity to modulate small-molecule alignment by taking advantage of the sensitivity of AMPS-acrylamide gels structure to electrostatic forces.

Herein, our objective is to explore the effect of factors that are known to modulate the swelling properties of AMPS-acrylamide gels by acting on the formation / dissociation of ion-pairs (from the sulfonate groups), namely, the addition of counterions (NaCl) and the polarity of the solvent (using water, DMSO, or mixtures thereof). We will study not only the aligning properties of the swollen gels, but also the whole swelling process —evolution of alignment over time in relation to the evolution of swelling itself—. *N*-methylcodeinium iodide (**1**, Figure 2.1) was used as model compound to test the gels and working conditions.

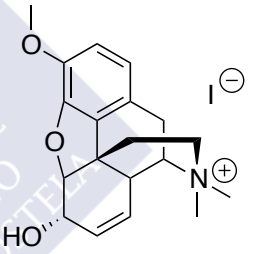


FIGURE 2.1: *N*-methylcodeinium iodide (**1**).

2.1.1 An improved synthesis of AMPS-Acrylamide gels at room temperature

The original publication by Haberz *et al.* described a gel with convenient properties for the alignment of small molecules in water.[94] Briefly, the gel was cast in a polymerization chamber where the pre-gel mix and the initiator were added. After polymerization, the gel rod was pushed out with a piston and washed thoroughly.

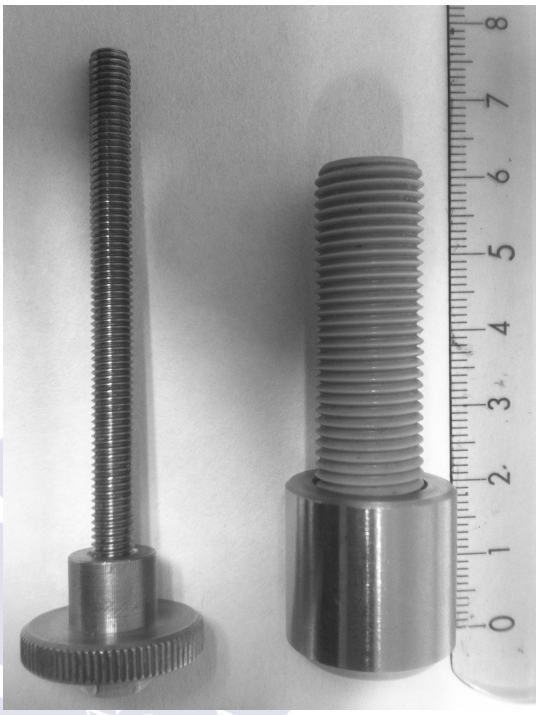


FIGURE 2.2: Gel casting apparatus made in PEEK and kindly donated by Prof. Christian Griesinger. The gel is polymerized inside the chamber and extracted carefully. The scale is in cm.

However, in our hands the extraction of the gel from the polymerization chamber is the most problematic part of the process, as these gels are rather stiff and tend to break when pressure is applied with the piston to extract them. This difficulty had been already described by the authors.[94] Due to this problem, we improved the preparation protocol, as described in the following.

The polymer gel presented by Haberz *et al.* was formed from AMPS : DMAA : BIS with molar ratio of 1 : 1 : 0.026, where BIS acts as a covalent cross-linker between polymer chains. Liu *et al.* considered ratios between 1 : 3 and 3 : 1 for the co-monomers with a BIS molar fraction of 0.013. Polymerization was initiated by APS (in low concentration, *ca.* 0.15%, w/w) by heating the pre-gel solution containing the monomers, cross-linker and the radical initiator, to about $79 - 90$ °C for two hours. Following the polymerization, the gel was thoroughly washed with ultrapure water and charges were neutralized in dilute solutions of the desired cation. We considered that gel stiffness could be reduced by either reduction of the amount of polymer (i.e. total concentration of monomers) or by modification in the ratio of BIS cross-linker.

As first modification, we diluted the pre-gel mix to a total monomer concentration of 0.5 M instead of the 0.75 M of the original protocol.[120] This made the extraction easier, but the gel was still very stiff and this method resulted in very small dry gel sticks (as expected, because of the reduction of total polymer mass).

In the protocol published by Haberz *et al.*, the ratio between AMPS : DMAA : BIS was 1 : 1 : 0.026,[94] i.e. quite a different amount of the *cross-linker* BIS than in the original reference by Liu *et al.*[108,120,121] Therefore, we increased the *cross-linker* molar fraction up to a ratio 1 : 1 : 0.034 (AMPS : DMAA : BIS) with a total monomer concentration (AMPS+DMAA) of 0.75 M. The gel stiffness was reduced with this modification, although the extraction from the chamber was still problematic. Additionally, the gel pieces usually presented bubbles inside, most likely caused by the high polymerization temperature. Furthermore, this kind of polymerization makes impossible the use of the whole length of the polymerized gel piece, as its ends became partially dehydrated and tended to break during the extraction or in the following washes.

We interpreted that the poor consistency of these gels was a consequence of the polymerization conditions and not (a consequence) of their co-monomer ratio. Moreover, we did not want to change its composition, given that the gel with 1 : 1 ratio had been reported to successfully align organic small molecules.[94] One of the main weaknesses —in our hands— of Haberz's protocol was the high temperature (about 2 hours between 70 °C and 90 °C) needed to complete the polymerization reaction, which resulted in air bubbles and dry (polymer) ends.

Polyacrylamide gel is one of the most commonly used gels for protein electrophoresis in biology laboratories dedicated to protein expression and purification. Basically, it is the same kind of polymer but for the *co-monomer* AMPS. When APS is used as the radical initiator of the polymerization of acrylamide, the reaction is very slow and can take several days at room temperature to complete. The reaction can be accelerated by using UV light or by increasing the temperature. Another way to accelerate the gel polymerization is the addition of the radical initiator *N*,*N*,*N*′,*N*′-tetra-methylethylenediamine (TEMED), which makes feasible the polymerization of acrylamide in the presence of APS at room temperature.

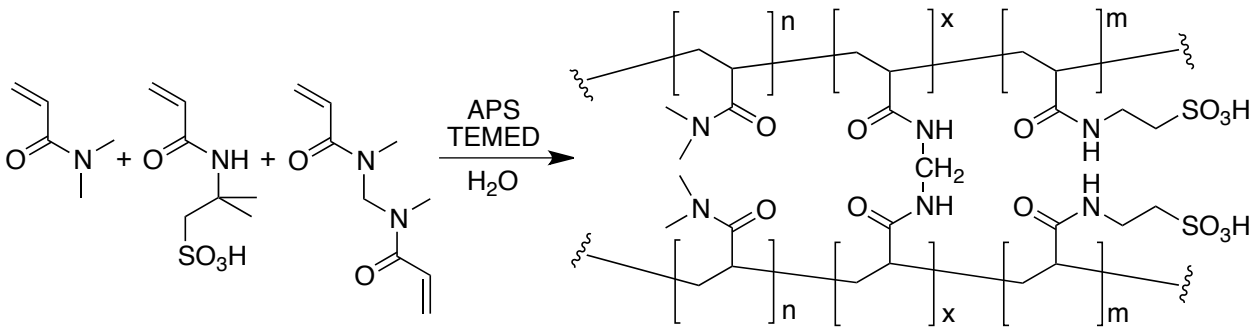


FIGURE 2.3: Synthesis of the AMPS-acrylamide gel. APS: ammonium persulfate. TEMED: tetramethylethylenediamine.

Accordingly, we performed the polymerization at room temperature in the presence of TEMED. The *pre-gel solution* contained AMPS : DMMA (1 : 1 mol) to a total monomer concentration of 1.5 M in milli-Q water. This solution is stable for more than 12 months when stored at 4 °C and carefully protected from light. The cross-linker BIS was prepared as a 65 mM stock in milli-Q water and is stored at 4 °C until used, for a maximum of one month. APS was freshly prepared in small aliquots at 0.5% (w/v) in milli-Q water and kept on ice. The APS stock solution should not be stored.

Moreover, proper mixing and timing are critical to attain reproducible results. Appropriate amounts (500 μL of AMPS/DMMA stock, 200 μL of BIS solution plus 300 μL of 0.5 % APS) of the stock solutions are mixed by vortexing in a 1.5 mL eppendorf tube *immediately* after the addition of TEMED. The critical step is the complete mixing of the monomers and cross-linker. This solution is transferred to the polymerization chamber. The chamber is sealed with parafilm tape and allowed to polymerize undisturbed for 30 minutes.

When the polymerization is finished, the gel piece is easily extracted from the chamber by gently pushing with the piston. Once extracted, the gel rod is washed for one hour in a neutralization solution. In our hands, 200 mM KOH, NaOH, or HCl furnished good quality gels with no appreciable air bubbles or any other major defect.

After neutralization, gels are thoroughly washed with abundant milli-Q water (about 100 mL each time per gel piece): 1 $\times$ overnight followed by 3 $\times$ 1 hour. Once washed, gels are dried in an oven at 70 °C for 2$-$3 days on Petri dishes to remove most of their water content. It is recommended to cover the dishes with parafilm tape to prevent sticking to the plastic dish. The resulting dry polymer sticks are stored at room temperature and protected from light until used.

Following this protocol, it is possible to reproducibly prepare gels with little defects and suitable to measure RDCs. It is also possible to make ready-to-use gels with no need to cut out the long polymer piece. Instead of adding 1 mL of pre-gel mix, we add (into the gel-casting chamber) just the volume of the desired final gel. In our lab, gels are prepared from 550 μL of pre-gel solution, which leads to gels of the desired length ($\approx$ 45 mm). This length fills quite precisely the active coil volume, which minimizes the magnetic field inhomogeneity caused by the Shigemi plunger.

2.2 *N*-methylcodeinium iodide (MCI, 1) is used as the model compound to probe the aligning properties of AMPS-acrylamide gels

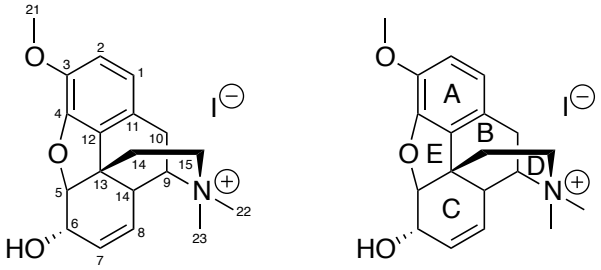


FIGURE 2.4: *N*-methylcodeinium iodide (MCI, **1**). Atom numbers and ring names are shown.

N-methylcodeinium iodide (**1**, MCI) was chosen as a model compound because it is soluble both in DMSO and water. Importantly, the morphinane skeleton (Figure 2.4) provides a rigid frame and a good number of C–H vectors (14) to determine the alignment tensor. Methylation of the amine nitrogen of codeine was required to prevent chemical exchange due to inversion of its configuration.[122]

2.2.1 Conformational Search of MCI

A MMFF94s molecular mechanics conformational search was performed in order to determine the rigidity of the MCI model compound.[123] DFT refinement of the MM computed coordinates resulted in three different conformations within a 8 kcal/mol energy threshold. DFT refinement was performed *in vacuo* and in implicit solvent (water or DMSO), starting from the MM computed coordinates. Solvation was modeled with the PCM continuum method in its integral equation formalism.[124] The same three geometries **1A**-**C** (Figure 2.5) were found in the three media. As no major structural differences can be attributed to solvation, only water-optimized coordinates were used from here on. Major conformational differences are located at rings C and D (Figure 2.4). The conformational space of MCI is represented by three conformers: **1A** (twist-boat / chair), **1B** (twist-boat / boat), and **1C** (half-chair / boat, Figure 2.5).

Conformer **1A** corresponds to the crystallographic structure,[122] has the lowest DFT computed energy and is the only one expected in solution, as the others are at least 7 kcal/mol higher in energy (Table 2.1).

Even though DFT-computed energies and Boltzmann statistics predict a negligible population of conformers **1B** and **1C** in solution, we introduced them in the RDC

fits as a quality control to illustrate the power of the methodology at differentiating structurally close conformers.

TABLE 2.1: DFT computed energies of MCI conformers

	$\Delta G_{298.15K}$, kcal/mol	
Structure	Water	DMSO
1A	0.0	0.0
1B	7.0	6.9
1C	7.6	7.7

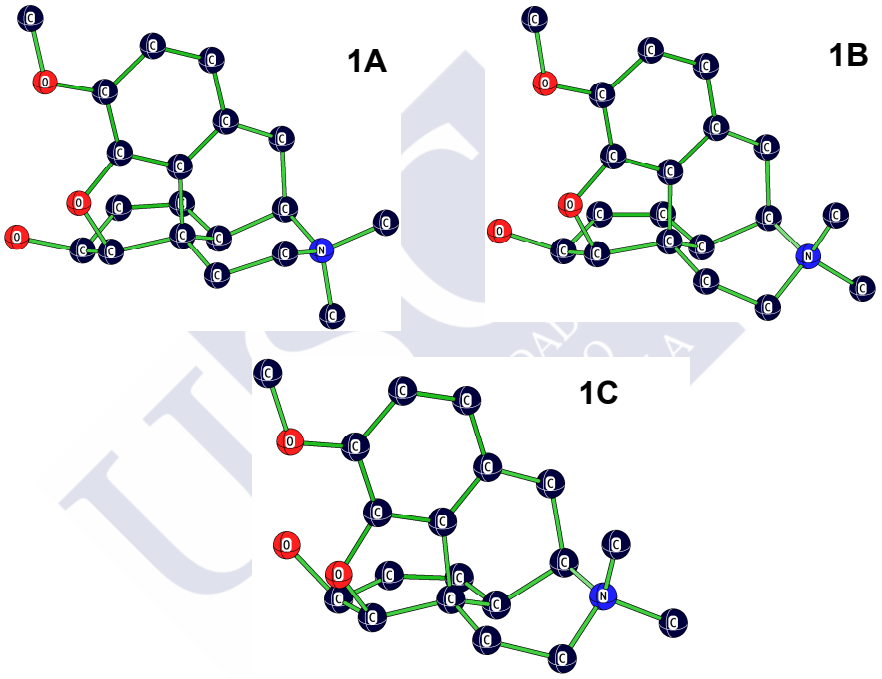


FIGURE 2.5: Structures found in the conformational search of MCI.

2.3 Swelling of gels

We wished to study the variables that most likely influence the alignment properties of AMPS-acrylamide gels, namely: solvent polarity and presence of ions. As these gels are reported to swell in D_2O and DMSO-d_6, we used these neat solvents and a 50 : 50 mixture. Ions can be introduced (in the gel) in two ways: in the wash solutions after polymerization and in the swelling solutions. The post-polymerization neutralization soaking is done either with acidic (200 mM HCl) or alkaline (200 mM NaOH) solutions, introducing H^+ or Na^+ as sulfonate group counterions, respectively.

The concentration of H^+ or Na^+ cations was not determined,[120] but we can assume that it is equal to the concentration of sulfonate groups, as gels are extensively washed with milli-Q water afterwards. Gels neutralized with NaOH or HCl will be named as *Na*-gels and *H*-gels, respectively.

In addition, excess Na^+ ions were introduced by swelling gels in 500 mM and 1000 mM NaCl solutions. In total, 10 NMR samples were prepared, which are summarized in Table 2.2. The following solutions were used for swelling: *w* (D_2O), *s1* (500 mM NaCl in D_2O), *s2* (1000 mM NaCl in D_2O), *d* (DMSO-d_6), and *m* (1 : 1 *v*/*v* D_2O/DMSO-d_6) (see Section 2.8.2). The concentration of **1** was 50 mM in all samples. Sample composition is summarized in Table 2.2 and in Table 2.3 showing the short code of the sample used along this Chapter.

Table 2.2: Sample conditions used in this study.[a]

	NaCl concentration, mM		
Solvent	0	500	1000
D_2O	*Na*-gel *H*-gel	*Na*-gel *H*-gel	*Na*-gel *H*-gel
D_2O/DMSO-d_6[b]	*Na*-gel *H*-gel	*Na*-gel *H*-gel	*Na*-gel *H*-gel
DMSO-d_6	*Na*-gel *H*-gel	*Na*-gel *H*-gel	*Na*-gel *H*-gel

[a] AMPS-acrylamide did not swell in several conditions, shown in light gray.
[b] 1 : 1 *v*/*v* D_2O/DMSO-d_6 mixture.

Table 2.3: Codes of sample conditions.

	NaCl concentration, mM		
Solvent	0	500	1000
D_2O	*w*	*s1*	*s2*
D_2O/DMSO-d_6[a]	*m*	–	–
DMSO-d_6	*d*	–	–

[a] 1 : 1 *v*/*v* D_2O/DMSO-d_6 mixture.

Samples were compared in terms of evolution of

i) gel swelling —macroscopic level—, and

ii) efficiency of molecular alignment —microscopic level—.

Time evolution of the alignment was assessed from the size of couplings (solute RDC magnitude and solvent $|\Delta\nu_Q|$) and the characteristics of the alignment tensor.

2.4 Time Evolution of the Samples: solvent polarity and counterions influence the swelling properties of AMPS-acrylamide gels

Swelling of the gels to their maximum length took from a few days up to several weeks. Time evolution of swelling of the 10 samples is shown in Figure 2.6. Gel swelling was slower in solutions containing DMSO-d_6 ($50-100$ %), namely: *m* and *d* samples, than in the aqueous samples. Slower swelling was observed with increasing concentration of NaCl in the D_2O solutions (*s*1 and *s*2 samples, respectively) both in *H*-gels and *Na*-gels (Figure 2.6).

With respect to the effect of the two neutralization solutions —acidic or alkaline—, both *H*- and *Na*-gels swelled at similar rates in aqueous solutions (Figure 2.6) but it is worth noting that *Na*-gels always induced larger T_{CH} couplings than *H*-gels in all swelling solutions (Figure 2.10).

When comparing the effect of the two neutralization solutions, a remarkable difference was found in the case of the less polar solvent (DMSO-d_6, Figure 2.6 dotted line with circle labels). Swelling of *H*-gels in DMSO-d_6 was very slow and did not reach the maximum length even after 70 days (maximum length of 20 mm), whilst *Na*-gels eventually reached the maximum length within 70 days.

In summary, the neutralization solution —the counterion in the gel— is more important when the gels swell in the less polar solvent DMSO-d_6 than in the more polar water. This difference observed in swelling rates of *H*- and *Na*-gels in D_2O but, more importantly, in DMSO-d_6 solutions, agree well with the described effect of Na^+ counterions leading to higher swelling capacity than H^+.[108,120,121]

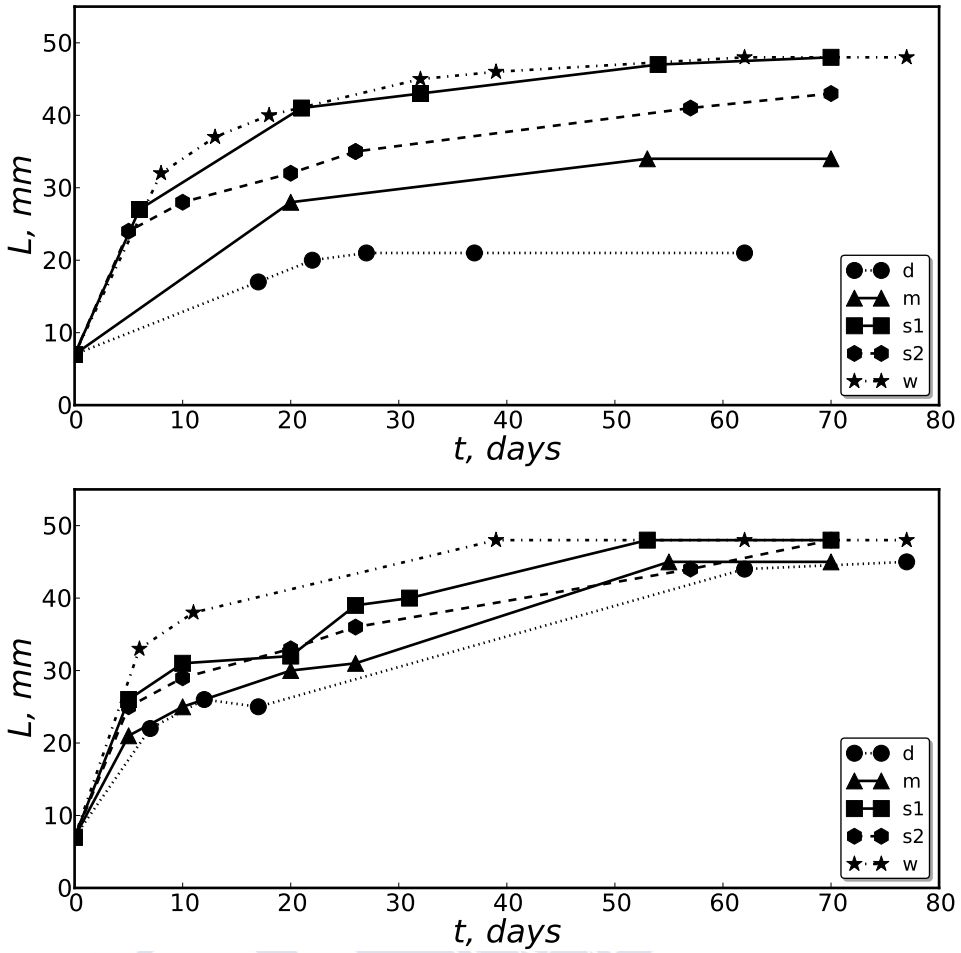


FIGURE 2.6: Evolution of gel swelling over time. Top, *H*-gels; bottom, *Na*-gels. Both *H*-gels and *Na*-gels swelled at similar rates in aqueous solution, while in the presence of DMSO-d_6, *H*-gels swelled much slower than *Na*-gels. Solvent: *w* (D_2O); *s1* (500 mM NaCl in D_2O); *s2* (1000 mM NaCl in D_2O); *d* (DMSO-d_6); and *m* (1 : 1 *v*/*v* D_2O/DMSO-d_6).

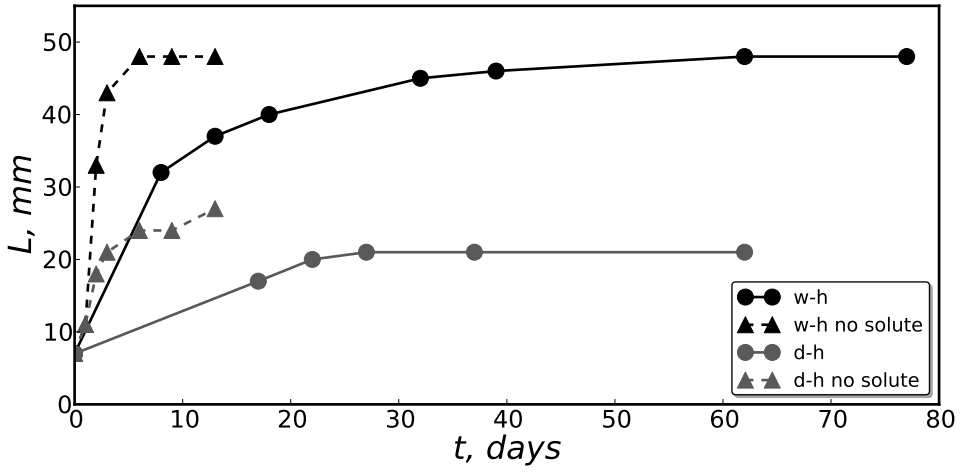
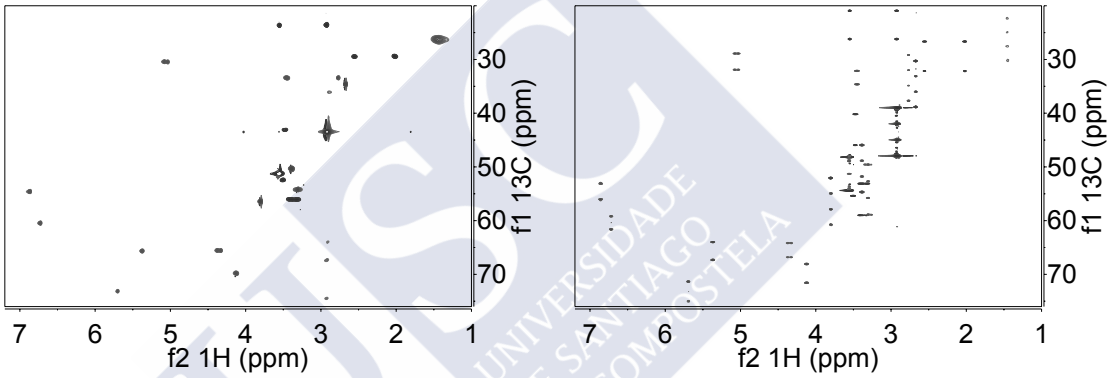


FIGURE 2.7: Evolution of *H*-gel length (*L*) over time (*t*) with and without solute **1**. Solvent: *w* (D_2O) and *d* (DMSO-d_6).

The presence of solute **1** in the swelling solution produced a decremented swelling rate of *H*-gels, (Figure 2.7) compared to the neat solvents D_2O or DMSO-d_6. This is attributed to the ionic nature of the solute, which can participate in the network of ion-pair interactions.

2.5 Time Evolution of the Samples: weak alignment of solvent and solute

Weak alignment induced by AMPS-acrylamide gels is the consequence of restricted molecular tumbling inside the anisotropic hollows formed upon gel swelling. Both the solute and the solvent molecules orient weakly in the stretched gel.



(a) Regular (decoupled) HSQC (b) F_1-coupled HSQC

FIGURE 2.8: Regular (decoupled) and F_1-coupled HSQC experiments of the completely swollen *Na*-gel/*w* MCI sample. Both HSQC are multiplicity edited. F_1-coupled HSQC contains a BIRD element during the *J* evolution, which is κ-scaled, see text.

Gel anisotropy and alignment strength can be assessed from the (residual) quadrupolar splitting ($|\Delta\nu_Q|$) of the solvent deuterium resonance. In addition, the use of a structurally more complex solute molecule (MCI, **1**) offers the opportunity to extract more information, like several RDCs and its alignment tensor.

One-bond C–H RDC ($^1D_{CH}$) couplings of **1** were determined from the difference in C–H splittings in F_1-coupled *J*-Scaled (JS) HSQC spectra (Figure 2.8) recorded in isotropic ($^1J_{CH}$) and anisotropic ($^1T_{CH} = {}^1J_{CH} + {}^1D_{CH}$) conditions. We used two different pulse sequences for the measurement of one-bond couplings: the JS-HSQC —written by Dr. J. Ying—, which includes a scaled *J*-evolution period, and the JS-BIRD-HSQC, which includes a BIRD element during the *J*-evolution period for refocusing heteronuclear long-range $^nJ_{CH}$ couplings that contribute to linebroadening.[52,77]

In these versions of the HSQC experiment that are coupled in the indirect dimension, the observed splitting of methylene groups in F_1 is the sum of their two one-bond C–H couplings. Their couplings were treated as presented in Section 3.3.4.3 (see also Chapter 1, Section 1.5.1) in the RDC fit calculations.

A reference isotropic sample of **1** was measured in each swelling solution to account for any potential solvent-dependent difference in the isotropic $^1J_{CH}$ values (Table 2.4).

Table 2.4: Isotropic $^1J_{CH}$ of MCI in each of the five solutions used for swelling.

	Isotropic $^1J_{CH}$, Hz				
Vector	*w*	*s1*	*s2*	*m*	*d*
C1–H1	160.3	165.6	160.2	163.8	159.2
C2–H2	158.7	163.4	161.0	161.2	157.2
C5–H5	160.5	160.5	160.6	160.2	158.8
C6–H6	143.4	143.4	143.3	141.7	139.9
C7–H7	165.0	165.1	164.9	164.6	163.4
C8–H8	166.0	166.1	166.5	166.2	165.0
C9–H9	151.4	151.3	150.5	151.5	151.4
C10–H10[a]	130.4	130.3	130.3	130.2	129.8
C14–H14	134.6	134.6	135.0	134.4	134.5
C15–H15[a]	132.7	132.6	132.8	132.6	132.1
C16–H16[a]	146.2	146.3	144.6	145.9	145.5
C21–H21	145.8	145.8	145.7	145.6	144.5
C22-H22	144.9	144.9	144.4	144.8	144.3
C23-H23	144.9	144.9	144.5	144.9	144.3

[a] There is only one entry per methylene group, which is the half-sum of the individual $^1J_{CH}$, as described in the text.
Solvent: *w* (D_2O); *s1* (500 mM NaCl in D_2O); *s2* (1000 mM NaCl in D_2O); *d* (DMSO-d_6); and *m* (1 : 1 *v*/*v* D_2O/DMSO-d_6).

2.5.1 Weak Alignment in the Fully Swollen Samples

First, we analyzed the RDC data of the completely swollen samples (Table 2.5). Experimental RDC recorded in each of the 10 samples were fitted to each one of the MCI structures **1A-C** found in the conformational search (Figure 2.5) with RDCFIT. As the stereo-specific assignment of the diastereotopic *N*-methyl groups 22/23 was not known, we tried both candidate assignments in the RDC fit.[82]

As expected based on the dramatic differences of their calculated energies, conformer **1A** gave the best fit with all the experimental RDC sets, scored in terms of the Cornilescu quality factor Q_C (Table 2.6).[44]

TABLE 2.5: $^1D_{CH}$ of MCI in the fully swollen samples.

vector	*H*-gels $^1D_{CH}$, Hz					*Na*-gels $^1D_{CH}$, Hz				
	w	*s1*	*s2*	*m*	*d*	*w*	*s1*	*s2*	*m*	*d*
C1–H1[a]	–	–	–	–	–	–	–	–	–	–
C2–H2[a]	–	–	–	–	–	–	–	–	–	–
C5–H5	−6.8	0.0	0.1	−1.7	−3.6	−8.8	1.8	3.9	−3.1	−3.1
C6–H6	−3.9	−0.3	0.7	1.6	5.0	−7.8	−2.0	−1.1	1.7	2.1
C7–H7	14.9	11.7	11.4	1.5	–[c]	17.6	14.2	13.4	1.7	−0.8
C8–H8	−1.7	0.2	−0.1	−1.4	–[c]	0.4	4.0	4.2	−2.8	−2.8
C9–H9	16.2	11.5	10.3	5.9	–[c]	23.7	14.9	12.8	8.7	3.9
C10–H10[b]	1.7	1.8	1.7	−1.1	–[c]	2.6	3.6	3.7	−1.7	−2.6
C14–H14	−5.1	−0.9	−0.3	1.2	–[c]	−9.3	−2.0	−1.5	1.9	1.8
C15–H15[b]	5.2	5.7	5.4	3.5	–[c]	6.8	7.5	6.5	5.0	0.4
C16–H16[b]	2.7	5.0	5.2	1.1	–[c]	1.3	4.3	5.5	1.3	−0.6
C21–H21[a]	–	–	–	–	–	–	–	–	–	–
C22-H22	0.5	−0.3	−0.1	−0.8	−1.4	1.7	0.3	0.5	−1.2	−0.6
C23-H23	9.9	8.0	8.6	2.9	0.8	11.9	10.9	10.0	3.9	−0.1

[a] C1–H1, C2–H2 and C21–H21 $^1D_{CH}$ couplings were not determined due to poor peak shape and/or severe shape distortions.
[b] There is only one entry per methylene group, which is the half-sum of each individual $^1D_{CH}$, as described in the text.
[c] Value not determined due to line broadening in the aligned sample.
Solvent: *w* (D_2O); *s1* (500 mM NaCl in D_2O); *s2* (1000 mM NaCl in D_2O); *d* (DMSO-d_6); and *m* (1 : 1 *v*/*v* D_2O/DMSO-d_6).

Regarding the stereoassignment of the *N*-methyl groups 22/23, values of Q_C calculated with both possibilities are shown for the *Na*- and *H*-gels swollen in neat water (*w*, Table 2.6). The disposition with methyls $22_{pro-R}/23_{pro-S}$ ($Q_C = 0.065$, *Na*-gel; $Q_C = 0.064$, *H*-gel) fits considerably better than the opposite ($Q_C = 0.371$ and $Q_C = 0.446$, respectively), hence the former is the assignment we used thereafter. In conformer **1A**, $22_{pro-R}/23_{pro-S}$ is equivalent to $22_{axial}/23_{equatorial}$. This assignment is consistent with the one reported for MCI in acetonitrile.[122]

TABLE 2.6: Quality of fit (expressed as Q_C values) of experimental RDC to each of the MCI conformers **1A**-**C**.

gel	solvent	**1A**	**1B**	**1C**
Na-gel	*w*	0.065	0.573	0.671
	s1	0.084	0.628	0.728
	s2	0.109	0.600	0.713
	m	0.159	0.652	0.774
	d	0.271	0.331	0.544
H-gel	*w*	0.064	0.608	0.718
	s1	0.113	0.634	0.769
	s2	0.105	0.660	0.795
	m	0.123	0.665	0.798
	d	_[a]	_[a]	_[a]
Na-gel[b]	*w*	0.371	0.542	0.619
H-gel[b]	*w*	0.446	0.555	0.632

[a] RDC values not determined due to line broadening in the aligned sample, see Table 2.5, and text.
[b] Q_C values of the fits with the inverted stereoassignment ($22_{pro-S}/23_{pro-R}$) of the diastereotopic *N*-methyl groups. All other entries correspond to assignment $22_{pro-R}/23_{pro-S}$.
Solvent: *w* (D_2O); *s1* (500 mM NaCl in D_2O); *s2* (1000 mM NaCl in D_2O); *d* (DMSO-d_6); and *m* (1 : 1 *v*/*v* D_2O/DMSO-d_6).

For the sake of completeness, a multi-conformer analysis of the RDC data sets obtained from the fully swollen samples was performed in order to address the possibility of a low-populated conformer in fast-exchange equilibrium with **1A**. We applied the single-tensor approximation (Chapter 1, Section 1.6).[47,54] The three DFT structures were superimposed by minimizing the RMDS between the cartesian coordinates of all heavy atoms.[125]

All possible combinations of 2 (Table 2.7) and 3 conformers (Table 2.8), out of the **1A**-**C** set were considered in the fit with RDCFIT (see Chapter 1 and Section 2.8.5 for a detailed description of the computation methods). The best fits of the RDC experimental sets (lowest Q_C) were obtained for ensembles containing the lowest energy

TABLE 2.7: Fit of the experimental one-bond RDCs to the 3 possible two-membered ensembles of MCI. Conformer populations and Q_C factors are shown.

Ensemble	Q_C	Population
1A+1B	0.060	100:0
1A+1C	0.060	100:0
1B+1C	0.601	100:0

TABLE 2.8: Fit of the experimental one-bond RDCs to the ensemble composed by all MCI conformers. Conformer populations and Q_C factors are shown.

Ensemble	Q_C	Population
1A+1B+1C	0.060	100:0:0

conformer **1A**. Such ensembles converged in all cases to 100 % population of geometry **1A**, in agreement with the DFT computed energy differences. Therefore, only the lowest-energy conformer **1A** was considered in the following.

2.5.2 Comparison of molecule orientation between swelling conditions

Examination of the alignment tensor characteristics (*GDO* and β) resulting from the fit to **1A** at different swelling times and in different swelling conditions, shows a dependence on solvent and gel neutralization solution.[126] Relative tensor orientation was determined in terms of the generalized angle β[26] between pairs of tensors (Table 2.9 and Figure 2.9). The degree of alignment (i.e. the alignment strength) is expressed in terms of the generalized degree of order (*GDO*). Equations are summarized in Chapter 1.[26]

Differences between the generalized angle β between two calculations can not be graded as significant without considering the impact of the experimental uncertainty of RDC values. Therefore, the uncertainty of the alignment tensor of each sample (Table 2.3) was estimated by generating a distribution of tensors all compatible with the experimental data, following the method proposed by Prestegard and co-workers (see Chapter 1, Section 1.4.5).[41] For every experimental RDC set, 512 simulated RDC sets were generated by sampling Gaussian distributions centered on the experimental values, allowing a very conservative 1.5 Hz standard deviation. This results in a distribution of alignment tensors of each sample that are compatible with the experimental data (Figure 2.9). Results of the 10 samples are summarized in Table 2.9 and in the following. The s.d. of β of the aqueous samples lies between 1.7° and 2.6°. Tensor

orientation of samples swollen in the same solution are very similar ($\beta = 5.0 - 8.2°$, Table 2.9), regardless of gel type (*Na*- or *H*-), even though the degree of order *GDO* is $\approx$ 25% larger in *Na*-gels compared to the *H*-gels swollen in the same solution.

In contrast, large angular differences up to $\beta \approx 30°$ are observed between samples swollen in different solutions. For instance, 50% DMSO-d_6 (*m*) leads to $\beta = 20 - 21°$ relative to neat D_2O (Table 2.9, column $\beta^{[c]}$). Similarly, added NaCl (*s1* and *s2*) results in rotation of the tensor by $\beta = 20 - 21°$ relative to neat D_2O. Note that tensor rotation is also different between DMSO-d_6 (*m*) and NaCl (*s1* and *s2*) samples; this can be seen more clearly in column $\beta^{[d]}$, where angles are calculated relative to the *Na*-gel/*s2* sample. Figure 2.9 displays graphically the relative orientations of the distributions of tensors in the different sample conditions. This influence of NaCl and DMSO-d_6 content on tensor orientation (β) and size (*GDO*) parallels the efects observed on the swelling properties of the AMPS-acrylamide gels (Figure 2.6), thus suggesting a dominant role of the steric component on alignment. In other words, impaired swelling is the consequence of enhanced cross-linking between polymer chains which, at the microscopic level, results in different size and shape of the gel cavities. This consideration does not exclude that the electrostatic component also contributes to molecule alignment (in all swelling media).

2.5.3 Time evolution of alignment in the aqueous samples

The time evolution of the alignment in each sample was analyzed using a series of RDC sets acquired at different days as gels swelled up. Figure 2.10 shows an overlay of F_1-coupled HSQC experiments as the samples in *H*- and *Na*-gels swelled up in neat D_2O. Evolution of the alignment over time is proved by the total coupling ($^1T_{CH}$) observed for the CH9 resonance.

The observed $^1T_{CH}$ and, consequently, the magnitude of RDCs, increased with gel length and time in all sample time series (Table 2.10). Accordingly, *GDO* (Chapter 1) values resulting from RDC fits to structure **1A** became larger over time (Figure 2.11), indicating an increment in the molecular order as a consequence of the increasing anisotropic compression of the gel matrix upon swelling. As expected, ^2H quadrupolar splitting $|\Delta\nu_Q|$ of the solvent resonance follows the same trend as *GDO* (Figure 2.12). When gels are completely swollen, both *GDO* and $|\Delta\nu_Q|$ are larger in *Na*-gels than in *H*-gels.

Comparison of *H*-gels evolution with *Na*-gels evolution (Figure 2.11) supports the conclusion that the neutralization treatment with NaOH furnishes gels (*Na*-gel) that can become more anisotropic than those neutralized with HCl (*H*-gel) at the same

TABLE 2.9: Comparison of *GDO* and tensor orientation (β) between samples. Data are shown only for the fully swollen samples.

gel type	solvent	L, mm[a]	$GDO \cdot 10^6$	β[b]	β[c]	β[d]	$\langle\beta\rangle$[e]	*s.d.*[e]
Na-	*w*	48	1000	0.0	0.0	21.5	3.5	1.7
H-	*w*	48	780	5.0	5.0	20.5	6.9	1.9
Na-	*s1*	48	877	0.0	18.6	3.4	18.7	2.6
H-	*s1*	48	666	7.0	15.4	8.6	15.9	2.3
Na-	*s2*	48	813	0.0	21.5	0.0	21.7	2.6
H-	*s2*	48	657	8.2	17.2	8.2	17.7	2.5
Na-	*m*	45	326	0.0	21.2	30.0	23.8	4.3
H-	*m*	34	238	5.3	20.3	25.1	24.5	5.3

[a] L: length of the swollen gel at the time of the acquisition of the NMR spectra.
[b] β relative to the *Na*-gel swollen in the same swelling solvent (comparison between *H*- and *Na*-gels).
[c] β relative to the *Na*-gel swollen in D_2O (effect of the solution used for swelling).
[d] β relative to the *Na*-gel swollen in *s2* (effect of the solution used for swelling; alternative reference).
[e] Average $\langle\beta\rangle$ and standard deviation (*s.d.*) values of the distribution of tensors obtained with an estimated experimental error of 1.5 Hz. All β are calculated relative to the sample *Na*/*w* (i.e., *Na*-gel swollen in D_2O).
Solvent: *w* (D_2O); *s1* (500 mM NaCl in D_2O); *s2* (1000 mM NaCl in D_2O); *d* (DMSO-d_6); and *m* (1 : 1 *v*/*v* D_2O/DMSO-d_6).

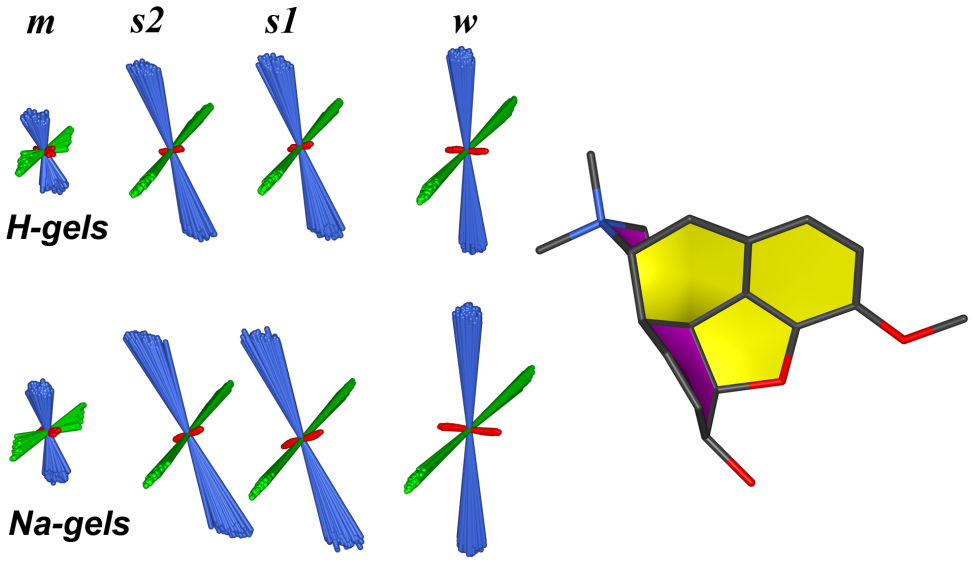
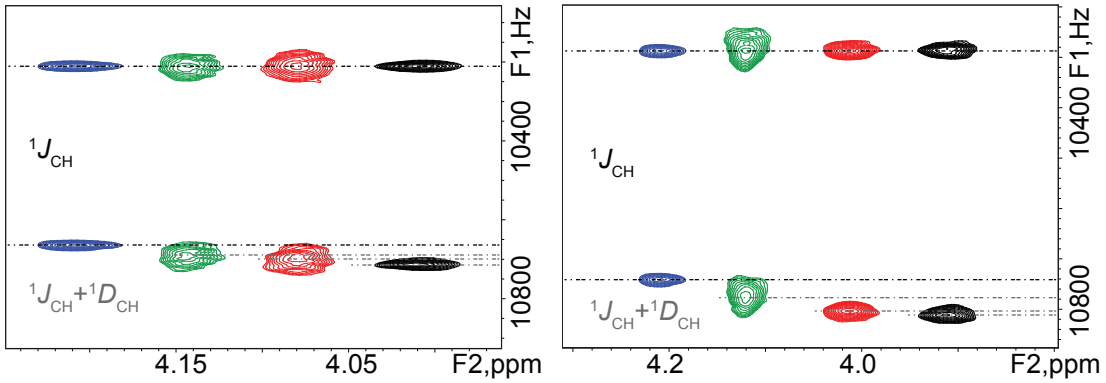


FIGURE 2.9: Comparison of tensor orientation between samples, showing the simulated tensor distributions in each alignment condition. Each tensor is represented by its three principal axes. We labeled the principal axes of the alignment tensor such that $A_z \geq A_x \geq A_y$, thus ensuring that the eigenvalues A_z and A_y are always positive and negative, respectively. Note that our choice deviates from the usual convention of labeling them such that $|A_z| \geq |A_y| \geq |A_x|$, that is less suited for the purpose of the present study, as it would make difficult to compare the preferred orientational probabilities. Principal axes colour code: blue, A_z; green, A_y; red, A_x. The length of each axis is proportional to the magnitude of its associated eigenvalue. Bundles are labeled with gel type (*Na-* or *H-*) and swelling solution (*w*, *s*1, *s*2, *m*). DMSO-d_6 samples are not shown. Solvent: *w* (D_2O); *s1* (500 mM NaCl in D_2O); *s2* (1000 mM NaCl in D_2O); and *m* (1 : 1 *v*/*v* D_2O/DMSO-d_6).

growth length. It can be speculated whether the stronger alignment in *Na*-gels is due to the steric (additional cross-lonking in the micro-cavities) or to the electrostatic components. The fact that both classes of gel swell at similar rates could suggest that the anisotropy and size of the cavities (steric component) is similar, thus supporting and influence of the electrostatic component (H^+ *vs*. Na^+ cations).

Interestingly, after the gel was fully swollen, the degree of molecular order (in terms of *GDO* value) still increased over time, indicating microscopic rearrangements of the gel matrix. This period, in which macroscopically the gel does not change (gel swelling is complete), but there are still microscopic changes in the gel matrix (that arise as modifications on the alignment tensor) can be referred to as *maturation period*. It is worth noting that, even during this maturation period, the relative generalized angle between tensors β remains practically unchanged. This is clearly appreciated if we take as reference the tensor of the same sample series at the point when the gel reached its maximal length, i.e. when it is completely swollen (Table 2.10 and Figure 2.13).



(a) *H*-gel (b) *Na*-gel

FIGURE 2.10: F_1-coupled HSQC of **1** showing the time evolution of coupling ($^1T_{CH}$) upon gel swelling. Overlay of the doublet of the C–H9 resonance in samples *H*-gel/D_2O and *Na*-gel/D_2O. Isotropic, blue; swollen gel, green ($L = 40$ mm, 18 days), red ($L = 45$ mm, 32 days), black, ($L = 48$ mm, 62 days).

The largest deviations of the relative β angle were observed at the initial stages of swelling. These larger deviations can be ascribed, at least partially, to the low degree of order present in the samples at that stage, which results in higher uncertainty of the experimental RDC values (as values are small, relative error is large). Nevertheless, evolution of the shape of the gel matrix cavities can not be excluded.

2.5.4 DMSO-d_6 samples

As discussed in the previous sections, gels swollen with 50 % or 100 % DMSO-d_6 have displayed slower swelling rates and reached much lower degrees of alignment than gels swollen with water. Actually, the *H*-gel barely swelled in DMSO-d_6 and the *Na*-gel reached only 45 mm length after 77 days giving, accordingly, very small RDC values (Table 2.5) and poor fit to MCI structures (Table 2.6). This contrasts with aqueous samples reaching the maximum length of 48 mm in 40 days or less (Table 2.10 and Figure 2.6). Samples swollen in 50 : 50 DMSO-d_6 : D_2O displayed an intermediate behavior in terms of swelling rate and degree of order.

RDC fit to geometry **1A** furnished a reasonably good $Q_C = 0.12 - 0.16$ (Table 2.6), although **1** was aligned more weakly than in neat aqueous solutions as reflected by the smaller *GDO* values (Table 2.9, Figure 2.11). Tensor orientation deviated $15 - 20°$ from the reference sample swollen in water (Table 2.9 and Figure 2.13).

The reduced gel swelling (and hence reduced sample alignment) in the presence of DMSO-d_6 is in agreement with the reported influence of solvent polarity on gel

TABLE 2.10: Time evolution of RDC couplings (in Hz) of MCI in D_2O solution. *Na-* and *H*-gels.

gel type	*Na-*	*Na-*	*Na-*	*Na-*	*Na-*	*H-*	*H-*	*H-*	*H-*	*H-*
t, days	6	11	39	62	77	18	32	39	62	77
L, mm[a]	33	38	48	48	48	40	45	46	48	48
$\lvert\Delta\nu_Q\rvert$, Hz	4.5	4.8	8.5	9.5	9.6	3.8	5.3	5.6	6.2	6.3
C–H5	−0.1	−3.2	−6.8	−8.2	−8.8	−1.8	−4.2	−5.2	−6.3	−6.8
C–H6	−2.8	−2.9	−6.4	−7.5	−7.8	−2.0	−2.7	−3.0	−3.7	−3.9
C–H7	6.0	9.1	15.6	17.8	17.6	9.1	8.6	12.7	14.9	14.9
C–H8	−0.2	0.3	0.6	0.4	0.4	−0.3	−0.9	−1.3	−1.8	−1.7
C–H9	6.2	10.8	20.7	24.3	23.7	7.5	9.8	14.1	16.4	16.2
C–H10[b]	1.4	1.9	2.4	2.5	2.6	1.3	1.3	1.5	1.5	1.7
C–H14	0.9	−3.7	−7.4	−8.6	−9.3	−3.0	−3.7	−3.9	−4.5	−5.1
C–H15[b]	3.8	4.1	6.4	6.4	6.8	2.7	4.5	4.6	5.3	5.2
C–H16[b]	1.2	0.3	1.7	1.6	1.3	1.6	2.0	2.5	2.6	2.7
C–H22	−0.3	0.3	1.3	1.7	1.7	−0.1	0.2	0.5	0.5	0.5
C–H23	4.9	5.7	10.4	11.9	11.9	4.8	7.9	8.4	9.8	9.9
Q_C	0.314	0.097	0.072	0.082	0.065	0.093	0.136	0.073	0.074	0.064
$GDO \cdot 10^6$	353	517	889	1000	1000	415	537	675	780	780
β, °[c]	13.7	8.2	1.9	1.9	0.0	6.0	7.5	1.5	1.0	0.0
β, °[d]	13.7	8.2	1.9	1.9	0.0	8.7	8.5	5.4	5.4	5.0

[a] L: length of the swollen gel at the time of the acquisition of the NMR spectra.
[b] There is only one entry per methylene group, which is the half-sum of their two $^1D_{CH}$.
[c] Generalized angle β is calculated relative to the completely swollen stage of each sample.
[d] Calculated relative to the completely swollen *Na*-gel/*w* sample.

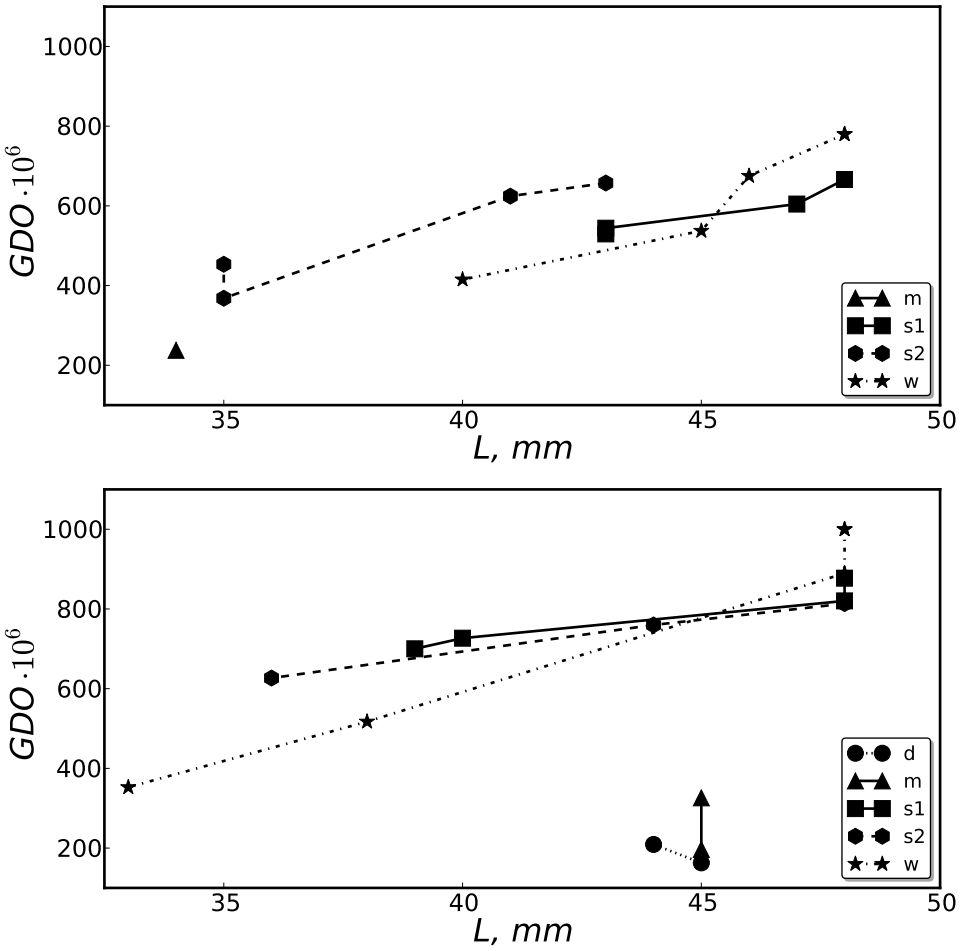


FIGURE 2.11: Evolution of alignment upon swelling. *GDO* reflects the total degree of order of the solute molecule. At equal gel length, *GDO* is larger in *Na*-gels (bottom) than in *H*-gels (top), reflecting the larger alignment strength of the *Na*-gels. Solvent: *w* (D_2O); *s1* (500 mM NaCl in D_2O); *s2* (1000 mM NaCl in D_2O); *d* (DMSO-d_6); and *m* (1 : 1 *v*/*v* D_2O/DMSO-d_6).

swelling, as lower solvent polarity favors the cross-linking among segments of the gel due to ion-pair aggregation.[108,121]

2.6 Discussion

It has been shown that solution composition influences gel swelling and hence its alignment strength. Optimal conditions to attain the maximum alignment of MCI comprise the use of *Na*-gel swollen in neat D_2O. The ion and solvent dependence of the alignment strength of AMPS-acrylamide gels is closely related to their swelling properties: alignment of MCI was higher with Na^+ than with H^+ counterions in the gel, and was reduced by high concentration of Na^+ ions in water or by the decreased solvent polarity of DMSO-d_6, both being factors that strengthen ion-pair aggregation and hence reduce the swelling capacity of AMPS-acrylamide gels. These observations

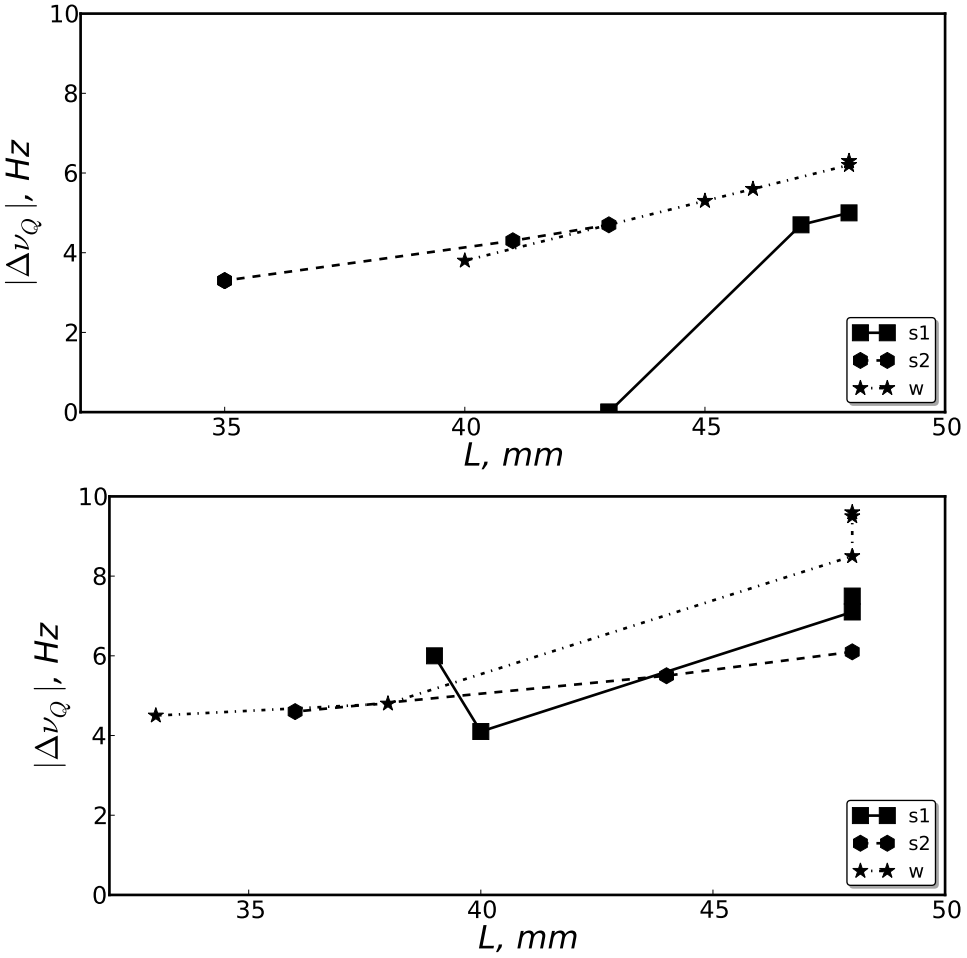


FIGURE 2.12: Evolution of alignment upon swelling. $|\Delta\nu_Q|$ reflects the degree of alignment induced by the gel. At equal gel length, $|\Delta\nu_Q|$ is larger in *Na*-gels (bottom) than in *H*-gels (top), reflecting the larger anisotropy of the *Na*-gels. Solvent: *w* (D_2O); *s1* (500 mM NaCl in D_2O); and *s2* (1000 mM NaCl in D_2O).

are in agreement with the model proposed by Liu *et al.*[108,120,121] to explain the response of these gels to solvent polarity and ions. According to this interpretation, Na^+ and H^+ counterions participate with the polymer sulfonate groups in the formation of ion-pair multiplets that strengthen the interaction between adjacent polymer chains. This cross-linking interaction is stronger with counterions that dissociate less efficiently (i.e. H^+). According to Liu *et al.*, affinity of the sulfonate group is larger for H^+ than for Na^+ ions, hence ionic cross-linking by this mechanism is stronger in *H*-gels than in *Na*-gels. This explains why *Na*-gels swell faster and better than *H*-gels.

It should be noted that extensive cross-linking interactions do not necessarily mean strong alignment, even though "cross-linking" may suggest the idea that cavities are smaller in size. Contrarily, what we observe is stronger alignment at lesser cross-linking degree. This is in agreement with the model of the SAG, i.e. that alignment is

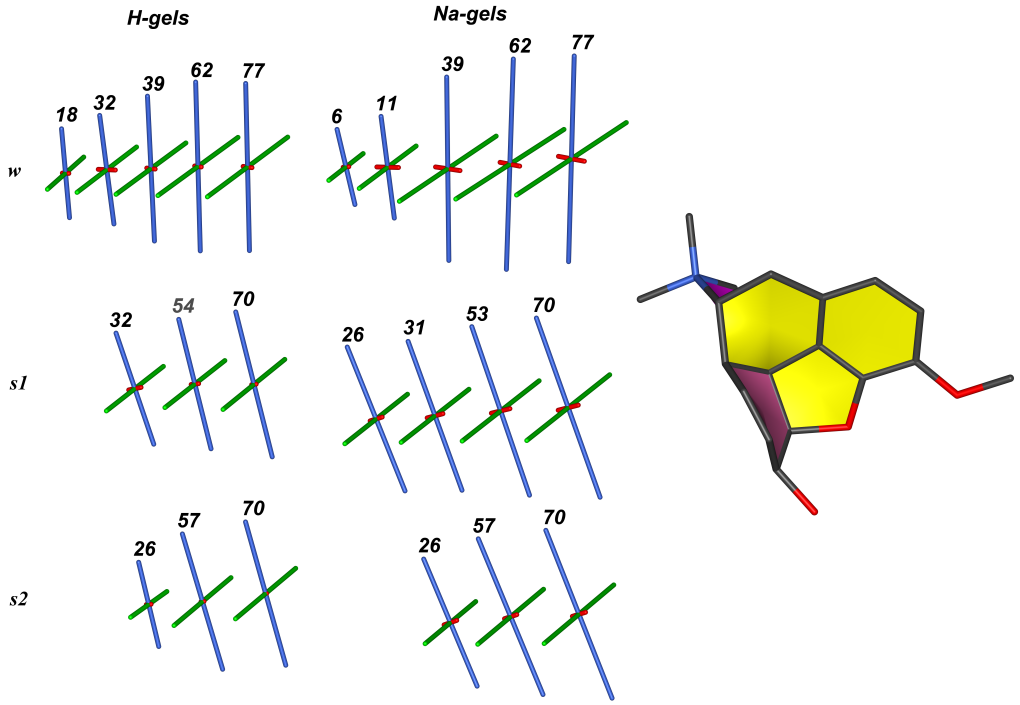


FIGURE 2.13: Time evolution of the relative orientation of the principal axes of the alignment tensor calculated in each sample condition. Axis length is proportional to the magnitude of its corresponding eigenvalue. We labeled the principal axes of the alignment tensor such that $A_z \geq A_x \geq A_y$, thus ensuring that the eigenvalues Az and Ay are always positive and negative, respectively. Blue, Az; green, Ay; red, Ax. Left, *H*-gels; right, *Na*-gels. From top to bottom: swelling solvent, w, $s1$, $s2$. Time evolution: each tensor is labeled with the value of the swelling time of the sample in *day* units. Solvent: w (D_2O); $s1$ (500 mM NaCl in D_2O); and $s2$ (1000 mM NaCl in D_2O).

the consequence of the anisotropic shape of cavities induced by the physical constrain to swelling imposed by the walls of the NMR tube.

Interestingly, high concentration of NaCl changed the orientation of the alignment tensor of MCI in aqueous solution. This observation opens the opportunity of modulating molecular alignment by the addition of ionic additives, given that molecular alignment is the consequence of steric and electrostatic forces.[127,128]

Previous work on electrostatic modulation of molecular alignment has been done in the biomolecular NMR field. Protein alignment in these examples was changed upon addition of salt, which can be ascribed to variations of the electrostatic component of the alignment mechanism, while the steric component remained practically unchanged. Sass *et al.* demonstrated the tuning on the alignment of Ubiquitin by calculating the $\hat{\mathbf{A}}$ tensor characteristics when adding NaCl to purple membranes.[128] In an analogous way, Zweckstetter and Bax showed that the alignment of protein G in Pf1 phages is dependent on the ionic strength. Interestingly, the orientation (and not only the magnitude) of the tensor varied, which was interpreted as a consequence of

changes in the electrostatic component while the steric component remained nearly unchanged.[127] In another work, Meier *et al.* presented a charged acrylate-acrylamide gel for the alignment of proteins. This gel has an enhanced electrostatic contribution as compared with the neutral acrylamide gels. Therefore, it presents different alignment properties. The electrostatic component of the alignment was probed by addition of NaCl and by modification of the alignment media pH, both resulting in the modification of the swelling and alignment properties of the gel.[129]

It is worth noting that the above-mentioned examples dealt with proteins, which have different charge distribution and are larger in size than organic small molecules. Influence of salt and other additives can be expected to be different in the case of alignment of small molecules, at least regarding the size of cavities that may lead to molecular alignment, which shall be quite different.

Indeed, an effect of salt on tensor orientation and size —modulation of the electrostatic contribution— is expected even if the geometry of the alignment medium —steric contribution— is not altered by the counterion, as described for purple membranes[128] and phages.[127] Nevertheless, Figure 2.6 illustrates that Na^+ ions have an effect on the geometry of the gel AMPS-acrylamide gels (hence on the steric contribution to alignment) that can be attributed to formation of ion-pair multiplets, supporting previous findings by Liu *et al.*[108,120,121] With our current data, we can not exclude that part of the reorientation of tensors is due to the electrostatic component and not to the steric component of the aligning interactions.

These observations confirm correlated differences of the swelling and *solute* alignment properties of AMPS-acrylamide gels depending on solvent polarity and counterion addition. Most likely, differences in molecule alignment originate from the different shape distribution of the gel cavities in the different swelling conditions, although an additional effect of the electrostatic component on the alignment can not be discarded.

If we admit that excess NaCl reduces gel swelling due to the ion-pair cross-linking mechanism, it is expected that excess NaCl also affects the shape and anisotropy of gel cavities, i.e. the steric component of alignment. The good correlation between swelling (gel length) and *GDO* of **1** supports an important role of the steric mechanism. In qualitative terms, this suffices to explain the observed changes in tensor size (increase or decrease) and/or orientation upon changes in NaCl content. Nevertheless, some influence of the electrostatic component of the alignment can not be discarded *a priori*, especially if we note that solute **1** has a net positive charge. With our current data, we can not exclude it.

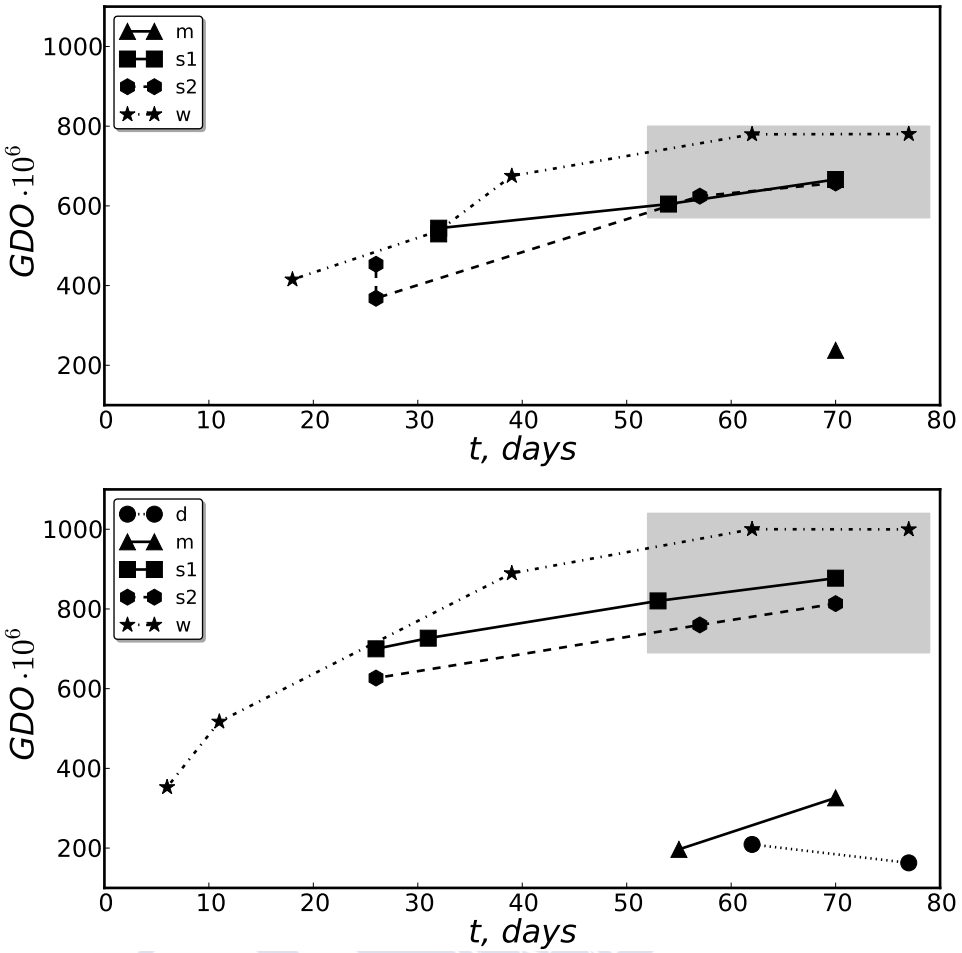


FIGURE 2.14: Time evolution of *GDO* in *H*- (top) and *Na*-gels (bottom). *GDO* reflects the total degree of order of the solute molecule. The shaded rectangle highlights the "maturation region". Modified from Figure 2.11. Solvent: *w* (D_2O); *s1* (500 mM NaCl in D_2O); *s2* (1000 mM NaCl in D_2O); *d* (DMSO-d_6); and *m* (1 : 1 v/v D_2O/DMSO-d_6).

Figure 2.15 and Figure 2.14 are modified versions of Figures 2.12 and 2.11, respectively . The "maturation period" is indicated by a shaded rectangle that highlights the time span when the gel is macroscopically fully swollen but is still experiencing changes in its microscopic structure, leading to changes in the degree of alignment of solvent (Figure 2.15, $|\Delta\nu_Q|$) and solute molecules (Figure 2.14, *GDO*).

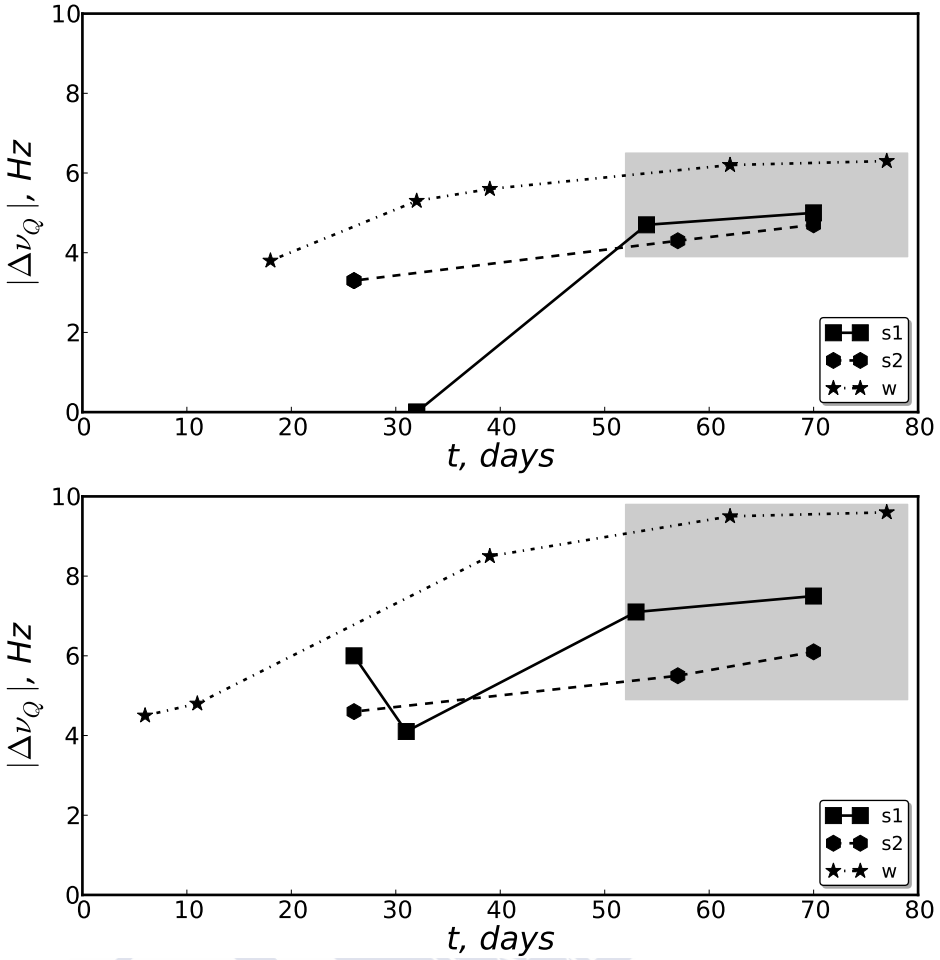


FIGURE 2.15: Time evolution of $|\Delta\nu_Q|$ in *H*- (top) and *Na*-gels (bottom). $|\Delta\nu_Q|$ reflects the degree of alignment induced by the gel on solvent molecules. The shaded rectangle highlights the "maturation region". Modified from Figure 2.12*(top)*. Solvent: *w* (D_2O); *s1* (500 mM NaCl in D_2O); and *s2* (1000 mM NaCl in D_2O).

This *matured* gel furnishes slightly different RDCs (Table 2.11) that translate into slightly different tensors, as reflected by the changes in β and *GDO*. It is worth noting that tensor evolution is more pronounced in the presence of Na^+ ions than in neat D_2O. The largest tensor rotation ($\beta = 17°$) is observed in DMSO-d_6 solution.

TABLE 2.11: Change in aligment upon gel maturation[a]. RDCs and fit results for *Na*- and *H*-gels.

gel type	${}^1\bar{D}_{CH}$[b]	$\lvert\Delta^1 D_{CH}\rvert$[c]	$\Delta^1 D_{CH}$, %[c]	ΔQ_C[d]	$\Delta\beta$, °[d]	$\Delta\left(GDO \cdot 10^6\right)$[d]
Na-gel						
w	8.4	0.3	4.7	< 0.01	1.13	10
s1	4.1	0.6	31.8	< 0.01	< 1.00	60
m	3.0	1.0	59.5	< 0.04	5.21	105
d	1.7	0.7	68.4	0.15	17.02	47
H-gel						
w	6.2	0.2	4.9	< 0.01	< 1.00	1
s1	6.0	0.3	4.4	< 0.03	2.43	47

[a] Differences from the first to the last day within the lapse defined as maturation period, see text.
[b] RDC mean in the last day of measurement (absolute value).
[c] RDC average change within the maturation period.
[d] Differences in fit results and alignment tensor parameters within the maturation period.
Solvent: *w* (D_2O); *s1* (500 mM NaCl in D_2O); *s2* (1000 mM NaCl in D_2O); *d* (DMSO-d_6); and *m* (1 : 1 *v*/*v* D_2O/DMSO-d_6).

2.7 Conclusions

The influence of solvent polarity and ionic force (salt concentration) on molecular alignment has been investigated for the strained aligning AMPS-acrylamide copolymer gel. The polymerization reaction was improved by means of TEMED addition, permitting the synthesis of homogeneous gels at room temperature. We tested AMPS-acrylamide gels in a variety of conditions (Table 2.2) aiming to investigate the influence of the different ionic additives on gel swelling. Therefore, RDCs were acquired for neat D_2O and DMSO-d_6, and various D_2O/DMSO-d_6 mixtures at different proportions, all containing different sodium chloride concentrations.

Experimental RDCs were fitted to the DFT-computed geometry of *N*-methylencodeinium ion with our RDCFIT software. Alignment tensors and several parameters, such as *GDO*, Q_C and generalized tensor β angle were obtained from the fit.

We observed decreased gel swelling as well as lower alignment degrees with increasing proportions of DMSO-d_6 (low polarity solvent, Figure 2.6 and Figure 2.7) and sodium chloride concentration (ionic force, Figure 2.6). This was expected based on previous findings by Lui *et al.*[108,120,121] that correlated conditions favoring ion-pairing

with gel shrinkage. Moreover, variation of the sodium chloride concentration was correlated to alignment tensor rotation (Table 2.9 and Figure 2.9).

The degree of alignment (*GDO*) as well as RDC differences reflects the larger maturation effect associated to the presence of high salt concentrations or apolar solvents.

2.8 Materials and Methods

2.8.1 Materials

MCI was synthesized starting from codeine following a literature procedure.[122]

2.8.2 Gel Preparation

Polymerization

A pre-gel solution containing AMPS, DMAA, and BIS (1 : 1 : 0.034 mol) with total monomer concentration of 0.75 mol/L was prepared in milli-Q water. The pre-gel solution (0.6 mL) was poured into the gel chamber (New Era Enterprises, 6 mm inner diameter) and polymerized at room temperature (30 min.), initiated by 1.5 mg/mL ammonium persulphate (APS) and 0.023 g/mL TEMED. After polymerization, gels were extracted from the chamber, washed in a neutralizing solution of 200 mM HCl or NaOH, as indicated, and extensively washed with milli-Q water (1 $\times$ overnight, followed by 3 $\times$ 1 hour).

The swollen gels were dried in a oven at 70 °C, typically for $2-3$ days, and the resulting dry sticks were stored at room temperature until use.

We prepared two variants of the AMPS-acrylamide gel that differed in the counterion introduced by the post-polymerization neutralization treatment, either with NaOH or with HCl solutions. We refer to them as *Na*-gels and *H*-gels, respectively. All gels used in this study were part of the same batch.

Swelling

A dry polymer stick ($\approx$ 3 mm diameter $\times$ 10 mm length) is placed on the bottom of an NMR tube. The swelling solution (550 μL) containing MCI is added dropwise to prevent the gel from floating. Samples were kept at 25 °C for swelling. The concentration of MCI was 50 mM in all NMR samples.

In order to test the effect of ion concentration and solvent polarity on alignment, stock solutions of MCI containing 22 mg/mL (50 mM) were prepared and labeled as follows:

- *w* (D_2O);
- *s1* (500 mM NaCl in D_2O);
- *s2* (1000 mM NaCl in D_2O);
- *d* (DMSO-d_6); and
- *m* (1 : 1 (v/v) D_2O/DMSO-d_6).

2.8.3 NMR spectroscopy

All experiments were carried out at 298 K on a Bruker Avance 600 spectrometer operating at 600.13 MHz for ^1H, 150.90 MHz for ^{13}C and 90.56 MHz for ^2H, equipped with a triple resonance inverse (TXI) room temperature probe with Z-only gradients.

^1H and ^{13}C resonances of MCI in DMSO-d_6 and D_2O solutions were assigned from HSQC, HMBC, COSY and NOESY (τ_{mix} = 600 ms) experiments. The assignment spectra can be seen in Appendix A, Section A.1.1 and the assignment in Appendix A, Section A.1.2.

One-bond CâĂŞ-H couplings were extracted from F_1-coupled JS-HSQC spectra. Spectra were recorded in isotropic ($^1J_{CH}$) and anisotropic ($^1T_{CH} = {}^1J_{CH} + {}^1D_{CH}$) conditions. One-bond ($^1D_{CH}$) RDCs were determined as the difference $^1D_{CH} = {}^1T_{CH} - {}^1J_{CH}$. Couplings of the methylene protons were taken as half the splitting in the HSQC, as the observed splitting of methylene groups is the sum of their two one-bond CâĂŞ-H couplings. All JS-HSQC spectra were acquired as 512* (^{13}C) × 512* (^1H) data matrices, where N* refers to N complex pairs, and a spectral width of 9057.97 × 4807.69 Hz, using 4 transients per FID and 1.5 s delay between scans, with a total acquisition time of 56 ms in the F_1 dimension and a *J*-amplification factor, κ, of 3.[52] Two version of the JS-HSQC were used. The one with the G-BIRD(r) module was the pulse sequence *hsqcbietgpjcsp.2* from the standard Bruker library.[77] The one lacking (not including) the BIRD module was kindly provided by Dr. J. Ying.

2.8.4 Conformational search

The conformational space of MCI was explored computationally by molecular mechanics MMFF94s force field[123] calculations using the mixed MMCM/low-mode

sampling algorithm[130] as implemented in Macromodel software.[110] The solvent (water) was taken into account with the semi-continuum dielectric model ($\epsilon = 79.0$). The energy cutoff was set to 16.3 kcal/mol to ensure complete coverage of the available conformational space. The resulting geometries and the available X-Ray geometry (CCDC 703371)[122] were subsequently optimized in G09[131] at the M06L/6-31G** level of theory[132] in the gas phase and in solution by using the Continuum Polarizable Model in its integral equation formalism (IEFPCM)[124] with G09 standard solvent parameters for water and DMSO. The pruned (75,302) default grid was employed for all DFT computations. Analytical frequencies were computed for each stationary point to ensure that true minima were reached. DFT derived energies are shown in Appendices (Appendix B, Section B.1.2).

2.8.5 RDC fit and analysis of molecular order

RDC fits were performed using RdcFit (see Chapter 1, Section 1.10). Optimization of the alignment tensor was done by means of the Powell minimization algorithm.[117]

Quality of fit was evaluated with the Cornilescu Q_C factor.[44] We compared the molecular alignment in the different sample conditions using the values of the parameters *generalized degree order* (*GDO*) and *generalized angle* β (see Chapter 1).[26]

Multiconformer fit

Multiconformer fit made use of the single-tensor approximation, which requires the definition of a common reference frame for all conformations. Therefore, we superimposed the three DFT structures **1A-C** by minimizing the RMSD between the cartesian coordinates of all heavy atoms.[51,125] All possible conformational ensembles containing combinations of 2 or 3 conformers among **1A-C** were generated and the experimental RDC were fitted to each of them by simultaneous optimization of populations and alignment tensors, using RdcFit. The output of the fit consisted of the alignment tensor components, conformer populations and Q_C factor.

Input and output data are reported in Appendices. Molecular coordinates of conformers **1A-C** (Appendix B, Section B.1.1), RDC input tables in RdcFit-ready format (Appendix B, Section B.1.3), and program outputs (Appendix B, Section B.1.4).

Chapter 3

Application of one-bond RDC to determine a fast conformational exchange

3.1 Salsolidine is an isoquinoline alkaloid with biological activity

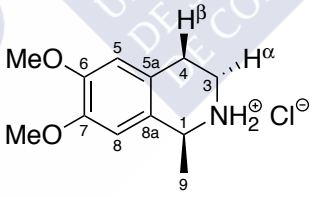


FIGURE 3.1: Salsolidine hydrochloride (**2**).

The isoquinoline alkaloid salsolidine (**2**, Figure 3.1), was first isolated as the (*S*)-enantiomer by Proskurnia and Orekhov from *Salsola richteri*.[133] Both enantiomers of the natural product, as well as the racemate, have been isolated from natural sources.

A number of endogenous tetrahydroisoquinolines (TIQ) have been identified in human and animal brain. Some of them, such as salsolinol and 1BnTIQ, are neurotoxic, while others —noticeably salsolidine, also abbreviated 1MeTIQ— show neuroprotective activity (Figure 3.2). MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine) and rotenone are known to induce *parkinsonian syndrome* humans and are often used to produce experimental animal models of Parkinson's disease by selective destruction of the nigrostriatal pathway.[134]

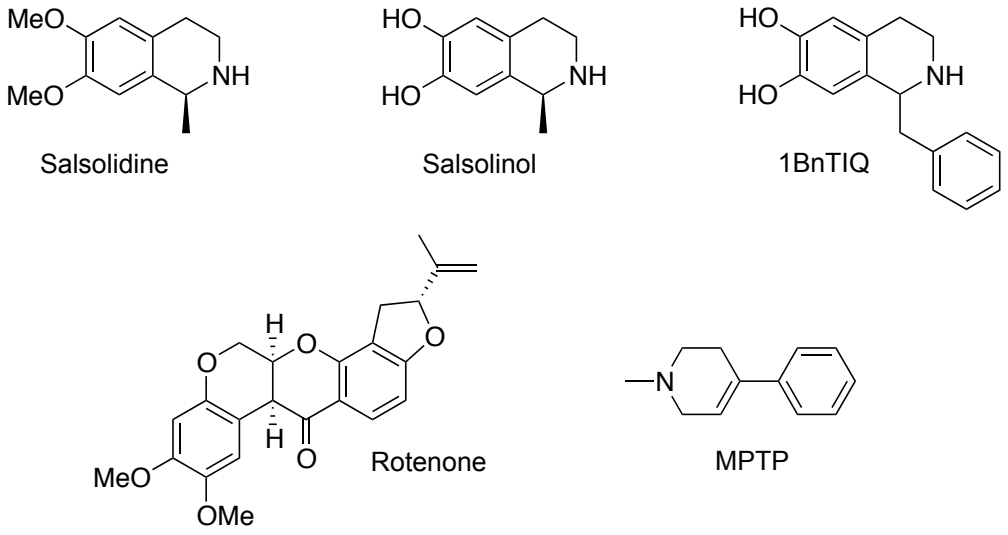


FIGURE 3.2: Parkinsonian syndrome related molecules (from top-left to bottom-right). Salsolidine (**2**); salsolinol (1-methyl-6,7-dihydroxy-1,2,3,4-tetrahydroisoquinoline); 1BnTIQ (1-benzyl-1,2,3,4-tetrahydroisoquinoline); rotenone ((2*R*,6a*S*,12a*S*)-1,2,6,6a,12,12a-hexahydro-2-isopropenyl-8,9-dimethoxychromeno[3,4-*b*] furo(2,3-h)chromen-6-one); and MPTP (1-Methyl-4-phenyl-1,2,3,6-tetrahydropyridine).

Contrary, salsolidine has been shown to counteract the neurodegeneration induced by endogenous (1BnTIQ) and exogenous neurotoxins such as MPTP and rotenone in experimental animals.[135–137] This makes salsolidine a promising molecule to become a neuroprotective drug or, at least, a leading compound of anti-Parkinsonian agents.[136,138] Both MPTP and rotenone are known to reduce the respiratory activity in the mitochondria, by the inhibition of ND1 (NADH-ubiquinone oxireductase chain 1, EC:1.6.5.3), thus causing a mitochondrial membrane depolarization and reactive oxygen species (ROS) liberation. This change in the oxidative state of the cell changes the dopamine catabolism to the MAO-mediated oxidative route, promoting the liberation of more ROS, and therefore causing oxidative stress. Interestingly, salsolidine promotes the (opposite) O-methylation route, reducing the generation of ROS during dopamine catabolism.[138] [1]

Other biological activities of **2** have also been reported. For instance, salsolidine inhibits the uptake of 5-hydroxytryptamine (serotonin) by human blood platelets.[143] It is also a competitive inhibitor of the methylation of 3,4-dihydroxybenzoic acid by the catechol-O-methyltransferase (COMT).[144]

The endogenous origin of salsolidine in human brain is yet unknown. In *Papaveraceae* plant tissue cultures, dopamine (3,4-dihydroxyphenethylamine) condenses with

[1] The reader can find a comprehensive Parkinson disease reference pathway on the **KEGG** database.

acetaldehyde to give salsolinol —through a Pictet-Spengler *non-enzymatic* condensation—, which is further metabolized to produce salsolidine.[139] This condensation is hypothesized to take place also *in vivo* in humans.[140–142]

Exposure to toxic substances of exogenous origin can raise the body levels of TIQ. For instance, it has been shown that products of condensation involving cigarette acetaldehyde and biogenic amines increase the addictive potential of tobacco. Salsolinol is one of such products formed from the condensation of acetaldehyde with endogenous dopamine.[145] Inhibition of the enzyme monoamine oxydase (MAO) has been proposed as the molecular basis of this activity. Also, antiseizure properties and supra-additive effect have been demonstrated in combination with the anticonvulsants valproate and carbamazepine.[146]

Conformational study of cyclic phenethylamines is an important subject due to their biological activity. Given the promising properties of salsolidine as neuroprotectant, we chose it as a representative of this class of compounds. The development of tools to precisely study their structure and conformational preference in solution may help us to understand the relationship between structure and function in this class of molecules. It is worth noting that classical NMR restraints like NOE or $^{3}J_{\mathrm{HH}}$ often do not suffice to define the structure of small molecules due to the scarcity of experimental data, especially in the case of conformational exchange. As discussed in Chapter 1, these classical parameters are hampered in the case of natural products as they are based in sequential information, and such molecules are very likely to have unsaturated bonds or oxidized moieties, thus breaking the sequential connectivity of the atoms (^{1}H) and making the structural analysis fail. Recently,[101] it has been shown that $^{1}D_{\mathrm{CH}}$ RDCs can be used to distinguish the preferred conformation of a 3-benzazepine (2,3,4,5-tetrahydro-*N*-methyl-3- benzazepine) in solution, without resorting to other NMR restraints. However, as aliphatic amines are (mostly) protonated at physiological pH it would be desirable to find aligning conditions compatible with amonium ions.

In this chapter we will show *first*, how the CPCL-LC medium can be used for the alignment of amonium ions and *second*, how the conformational distribution of salsoldine can be analyzed with the help of $^{1}D_{\mathrm{CH}}$ RDCs.

3.2 Salsolidine experiences extensive conformational exchange in solution

3.2.1 Previous ^{1}H-NMR Studies proposed a two-site exchange of salsolidine in aqueous solutions.

The conformational preference of several TIQs, including salsolidine, in acidic aqueous solution has been previously studied by ^{1}H NMR at 60 MHz.[147] Fast exchange between the two half-chair conformations was proposed based on the averaging of chemical shifts and scalar couplings of protons H3α/β and H4α/β as can be appreciated in Figure 3.5 and Figure 3.6 (no further computational or experimental data was presented on this compound). These proposed half-chair forms are equivalent to the **2A** and **2B** conformers found by us in the conformational search presented in Section 3.3.1.

3.2.2 Previous computational studies pointed to conformational exchange in 1-substituted TIQs.

Conformation of tetrahydroisoquinolines (TIQs) has been studied as part of drug discovery programs targeting dopamine D_1 receptors. Charifson *et al.* conducted conformational studies on a series of isoquinolines in the late 1980s, including a 1-benzyl-1,2,3,4-tetrahydroisoquinoline. By using Molecular Mechanics (MM) computations of both the free bases as well as the protonated forms, they found four possible isoquinoline ring conformations: two half-chair forms (in which only the H4 protons are truly axial or equatorial while the other substituents are in pseudo-ecuatorial or pseudo-axial dispositions) resembling **2A** and **2B** conformers, and two boat conformers along the C1–C4 axis (not like **2C**, Figure 3.3).[148]

In a more recent study, Olefirowicz and Eliel determined the conformational equilibrium of of salsolidine (1MeTIQ). 1MeTIQ was predicted to exist as a conformational average between two half-chair forms equivalent to **2A** and **2B** conformers (Figure 3.3) with *in vacuo* calculations. Additionally, NMR $^3J_{HH}$ analysis conducted in CD_2Cl_2 confirmed the previously determined equilibrium with populations approximately 30 : 70 for the conformers equivalent to **2A** and **2B**.[149]

We wished to address the conformation of salsolidine in solution, with the aid of RDC analysis and modern computational tools, to better define its conformational equilibrium in aqueous solution. We will present in the next advanced molecular modeling, *J* coupling analysis and RDC analysis of salsolidine in aqueous solution.

3.3 Results

3.3.1 Computational search identifies two accessible conformations

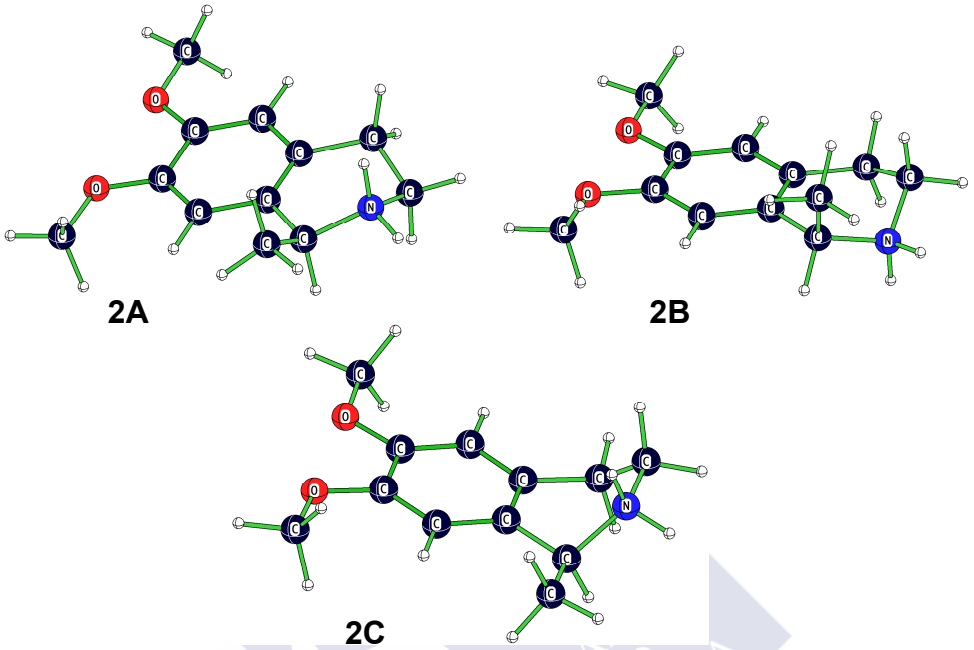


FIGURE 3.3: Low-energy conformations of Salsolidine (**2**).

The accessible conformations of salsolidine were explored by molecular mechanics MM3 force field calculations[150,151] by using the global GMMX algorithm as implemented in the PCmodel program.[111] Three stationary points were found within 10 kcal/mol cutoff limit: forms **2A** and **2B** mentioned above, and a half-boat form **2C**. These geometries were further minimized at the DFT level of theory taking solvation into account by the use of a implicit water Onsager model[152] with a solvating radium of 4.95 Å and a relative dielectric constant $\epsilon = 78.39$. The two half-chair forms **2A** and **2B** are very close in energy ($\Delta\Delta G_{298.15} < 0.1$ kcal/mol), while the half-boat conformer **2C** lies 1.9 kcal/mol over the basal **2A** form (Table 3.1).

DFT-derived Gibbs free energies ($\Delta G_{298.15}$) can be translated to expected populations with the use of the Maxwell-Boltzmann equation

$$p_i = \frac{\exp(-\Delta G_{298.15}/RT)}{\sum_{i=1}^{n} \exp(-\Delta G_{298.15}/RT)}, \tag{3.1}$$

where p_i is the population expressed as molar fraction, and $\Delta G_{298.15}$ the DFT-computed energy. The expected populations at T = 298.15 K, estimated on the basis of

Maxwell-Boltzmann statistics (Equation (3.1)), are shown in Table 3.1. Thus, according to DFT energies the **2C** conformer will present a very small population (2.1%), while the half-chair conformers **2A** and **2B** are close to 1 : 1 ratio. DFT energies are in agreement with the earlier interpretation of the averaging in the NMR spectra in terms of a fast two-site exchange between the two lowest-energy conformations (half-chair conformers).[147]

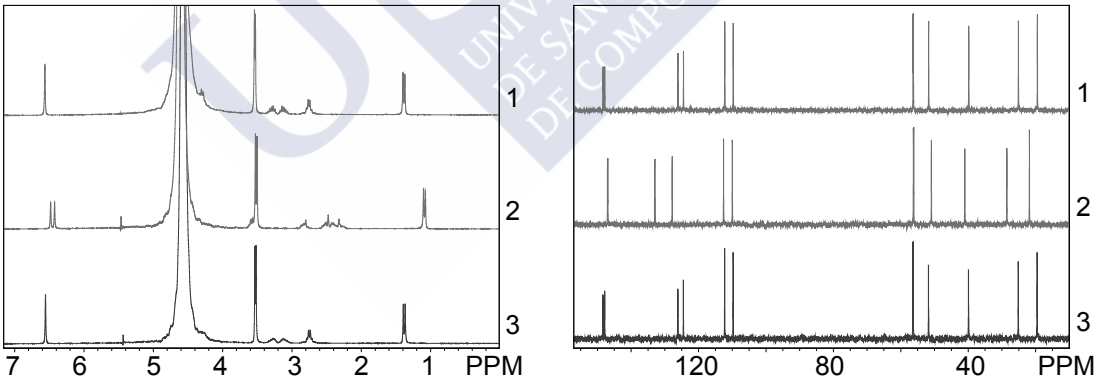
TABLE 3.1: DFT computed energies ($\Delta G_{298.15}$) and Boltzmann populations of the three low energy conformations of salsolidine.

Geometry	$\Delta G_{298.15}$, kcal/mol	p[a], %
2A	0.0	52.8
2B	0.1	45.1
2C	1.9	2.1

[a] p: Boltzmann population calculated from the -derived energies.

3.3.2 Evidence of conformational averaging in the ^{1}H spectrum

3.3.2.1 Protonation state of salsolidine in the samples



(a) ^{1}H (b) ^{13}C

FIGURE 3.4: Dependence of salsolidine NMR spectra on protonation of its *N* atom. (1) top, original solution of 200 mM **2**·HCl (**2** as ammonium salt); (2) center, after addition of 1 equiv. NaOH (**2** as free base); (3) bottom, upon addition of 1 equiv. HCl, the *N* is protonated again and the original spectrum is recovered.

Salsolidine was analyzed in its protonated form, i.e. as hydrochloride salt, which is the expected state at physiological pH. It is conceivable that a change in its protonation state changes dramatically its conformational equilibrium in solution, thus making necessary to probe the protonation state of salsolidine in the NMR sample conditions.

The isotropic sample consisted of 15 mg salsolidine hydrochloride (**2**·HCl) dissolved in H_2O:D_2O (85 : 15, v/v) to a final concentration of 200 mM. The protonation state of **2** was assessed by comparing the ^{1}H and ^{13}C spectra after successive addition of base and acid to the original sample (Figure 3.4). Addition of NaOH (1 equiv.) shifts resonances in both spectra, indicating that **2** deprotonates. Subsequent addition of HCl (1 equiv.) restores the original spectra, indicating reprotonation of the nitrogen atom. Addition of excess HCl does not produce further changes in the spectra. These observations prove that **2** is already protonated in the original isotropic sample. It is noteworthy the large shift of the H3 resonances, which shift > 0.4 ppm upon deprotonation of the *N* atom.

3.3.2.2 Averaging of vicinal couplings

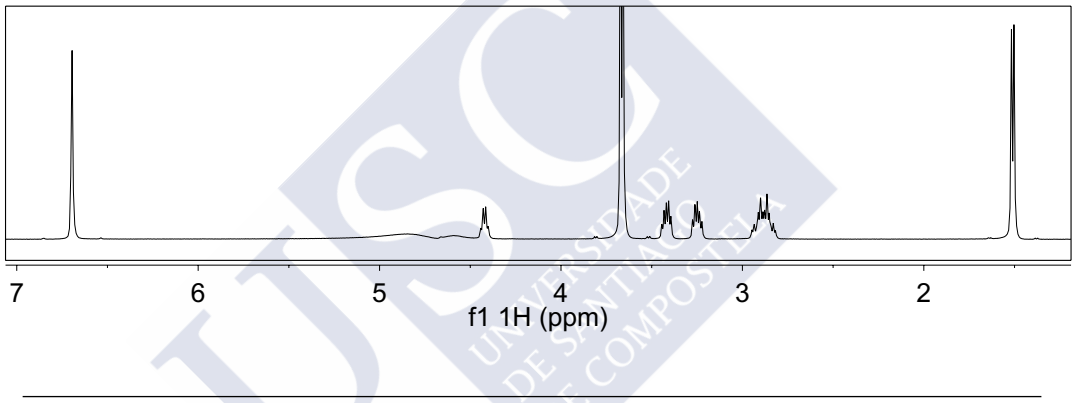


FIGURE 3.5: ^{1}H NMR spectrum of salsolidine hydrochloride. Spectrum acquired in H_2O:D_2O 90 : 10 (v/v) at 500 MHz field.

We measured the ^{1}H spectrum of salsolidine hydrochloride at 500 MHz. At this field, chemical shift averaging of the H3 and H4 α/β resonances is not as severe as that reported by Karimov *et al.* at 60 MHz[147] (see Figure 3.5, Figure 3.6 and Figure 3.8). Nevertheless, the values of the splittings and the shape of the multiplets are indicative of strong coupling and —possibly— conformational averaging.

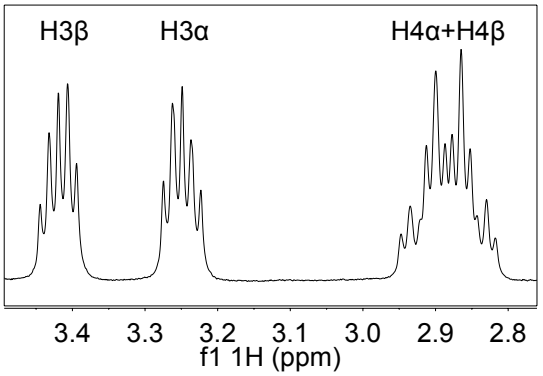


FIGURE 3.6: Aliphatic expansion of salsolidine hydrochloride ^{1}H NMR spectrum. Spectrum acquired in H_2O:D_2O 90 : 10 (v/v) at 500 MHz field

H3β and H3α protons were assigned based on the observation of a NOESY cross-peak between resonances Me9 and H3a, which is consequently assigned to H3β (Figure 3.7) and can be explained by the presence of **2B** conformer. This finding is supported by DFT-GIAO chemical shifts computed at the OPBE/pcS-2[153,154] level of theory and referenced to tetramethylsilane (TMS), which determine H3β resonance to be down-field for **2A** and **2B** conformers (Table 3.2).

TABLE 3.2: Experimental and computed chemical shifts of **2** referenced to TMS, in ppm.

Resonance	Experimental	**2A**	**2B**	**2C**
C1	51.67	61.39	60.23	58.08
C3	39.77	48.85	43.77	47.32
C4	25.04	29.59	28.43	28.66
C4a	124.67	113.62	112.49	119.76
C5	112.16	106.69	106.47	106.64
C6	148.28	149.48	149.45	150.43
C7	147.76	148.37	148.25	148.44
C8	109.67	101.87	103.20	101.70
C8a	126.25	113.61	112.61	114.90
C9	19.40	20.60	22.69	16.74
H1	4.42	4.92	4.77	4.39
H3β	3.42	3.71	3.86	3.07
H3α	3.25	3.65	3.55	3.77
H4β	2.91[a]	3.42	3.15	3.04
H4α	2.85[a]	3.20	3.37	3.16
H5	6.70	6.77	6.69	7.06
H8	6.69	6.54	6.36	6.85
H9	1.50	1.76	1.79	1.83

[a] Experimental chemical shifts from H4 α/β are not stereoasigned.

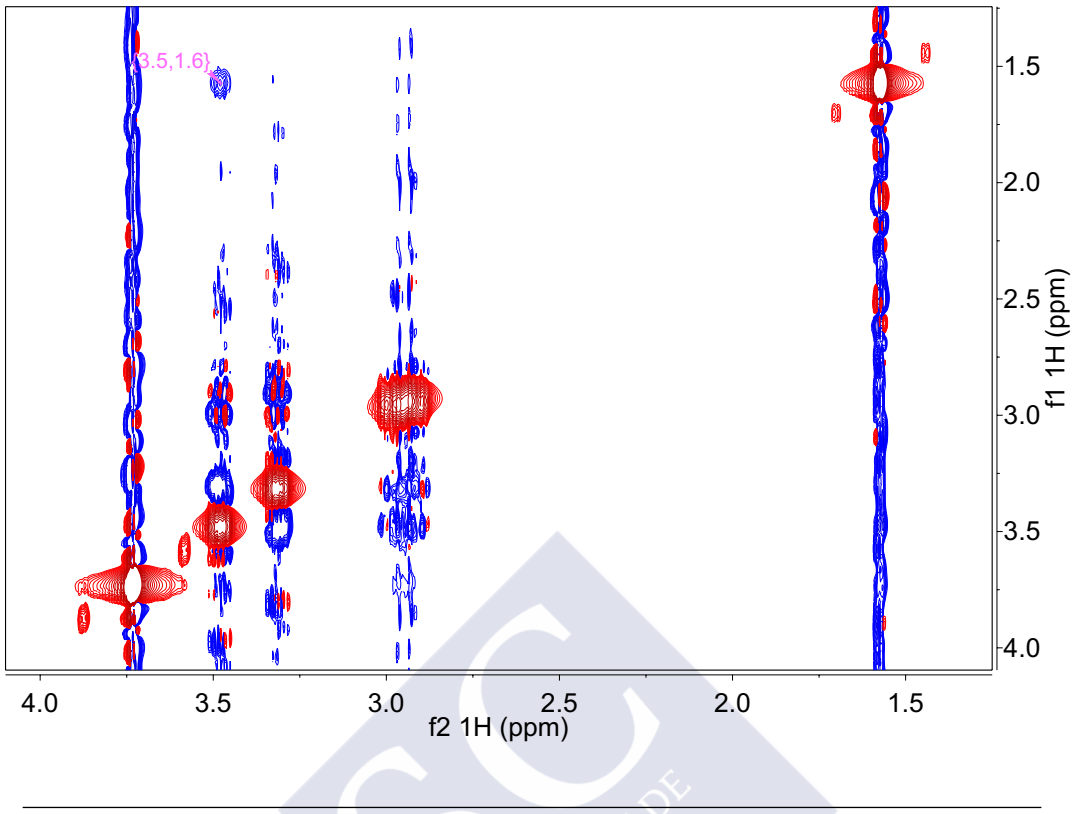


FIGURE 3.7: Aliphatic expansion of salsolidine 2D ^{1}H NOESY NMR spectrum. Cross-peak corresponding to Me9–H3a(β) correlation at $\approx$ 3.5/1.6 ppm is labeled in pink, see text for details. Spectrum geNOESY acquired at 500 MHz field with τ_{mix} (d8) $=$ 500 ms in H_2O:D_2O 90 : 10 (v/v).

An obvious way to experimentally quantify the ratio between conformations is to extract the $^{3}J_{HH}$ couplings of the protons at C3 and C4. However, direct extraction from the spectra was not possible due to resonance overlap. Spin simulation is a useful tool for the accurate analysis of multiplets of such strongly coupled spin systems. We applied this procedure to the spin system formed by *the ethylene fragment* H3α/β and H4α/β. We simulated the spectrum with the NUMMRIT module of the program Spinworks,[155] which uses the NUMARIT algorithm.[156] NUMARIT algorithm fits the simulated spectrum to the experimental one by iterating over scalar coupling and chemical shift values. Iteration stops when the RMSD between the spectral lines is below the given threshold or when it is no further minimized. Spectral fit needs a guess initial assignment of the resonances from the simulated spectra to the experimental one. For this particular ABMX spin system, the experimental and best-fit simulated spectra are shown in Figure 3.8. Frequencies and scalar couplings are shown in Table 3.4 and Table 3.3, respectively.

The vicinal $^{3}J_{HH}$ couplings of geometries **2A**-**C** were calculated with the Haasnoot-Altona equation[32] as implemented in MSPIN[42,43] (Table 3.5). At first sight, none of

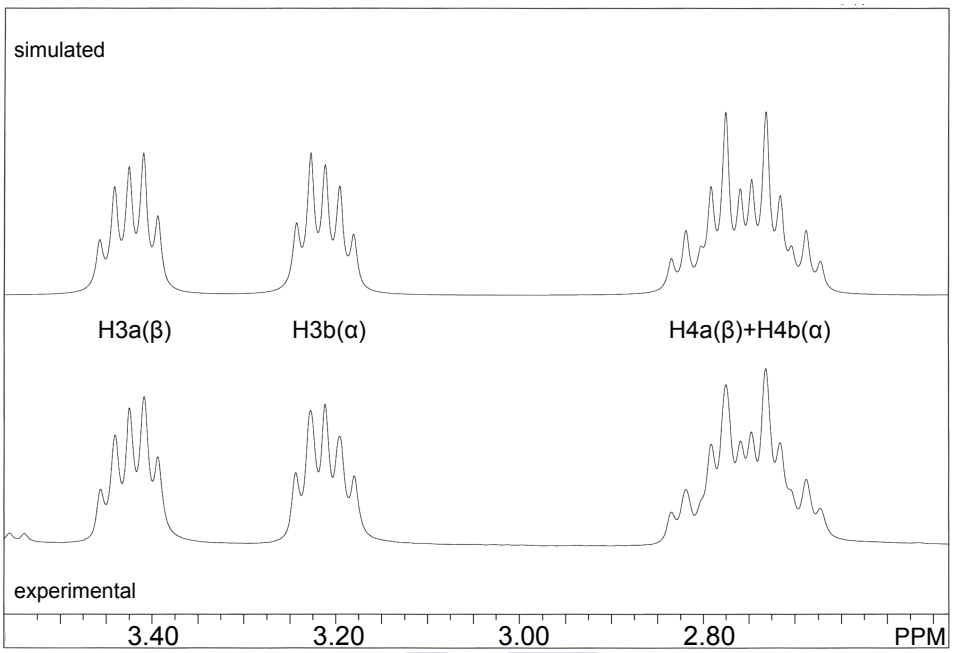


FIGURE 3.8: ^{1}H spectrum of Salsolidine; expansion of the H3/H4 region. Bottom, experimental; top, simulated with SPINWORKS. Values of frequencies and couplings are given in Table 3.3 and Table 3.4.

TABLE 3.3: Chemical shifts of protons H3/H4 determined from the ^{1}H spectrum by multiplet simulation with Spinworks NUMMRIT.

Resonance	ν, Hz			δ, ppm
H3a	1710.19	$\pm$	0.09	3.42
H3b	1625.66	$\pm$	0.09	3.25
H4a	1457.19	$\pm$	0.10	2.91
H4b	1426.77	$\pm$	0.10	2.85

the conformers matches the experimental $^3J_{HH}$ values (see Table 3.4 and Table 3.5). Furthermore, the fact that experimental values are all similar and close to the average value 6.3 Hz, indicates that the observed $^3J_{HH}$ values in the spectrum are conformational averages. Table 3.5 also shows the ensemble averaged values of a 50 : 50 equilibrium of the half-chair conformers **2A** and **2B**. Figure 3.9 shows the RMSD of experimental and computed $^3J_{HH}$, which results in deviations as high as 4 Hz . As expected, the 50 : 50 **2A**+**2B** ensemble fits clearly better (RMSD = 1.0 Hz) than any of the sole conformers.

TABLE 3.4: Scalar couplings of protons H3/H4 determined from the ^{1}H spectrum by multiplet simulation with Spinworks NUMMRIT.

	J_{HH} coupling, Hz		
Resonance	H3b	H4b	H4a
H3a	12.69[a]± 0.12	6.19± 0.14	6.33 ± 0.14
H3b		6.23± 0.14	6.48 ± 0.14
H4b			17.36[a]± 0.12

[a] Absolute value, the sign of the $^2J_{HH}$ was not determined.

TABLE 3.5: Haasnoot-Altona[a] predicted $^3J_{HH}$ coupling for the **2A**-**C** conformers as well as ensemble averaged values of a 50 : 50 equilibrium of **2A** and **2B**.

Coupling	**2A**	**2B**	**2C**	**2A**+**2B**[b]
H3β-H4β	5.9	3.8	4.1	4.9
H3β-H4α	1.2	12.2	12.1	6.7
H3α-H4β	12.3	1.2	1.2	6.7
H3α-H4α	3.6	6.1	6.0	4.9

[a] Haasnot-Altona equation as implemented in MSPIN.[42,43]
[b] Ensemble-averaged values of 50 : 50 **2A**+**2B**.

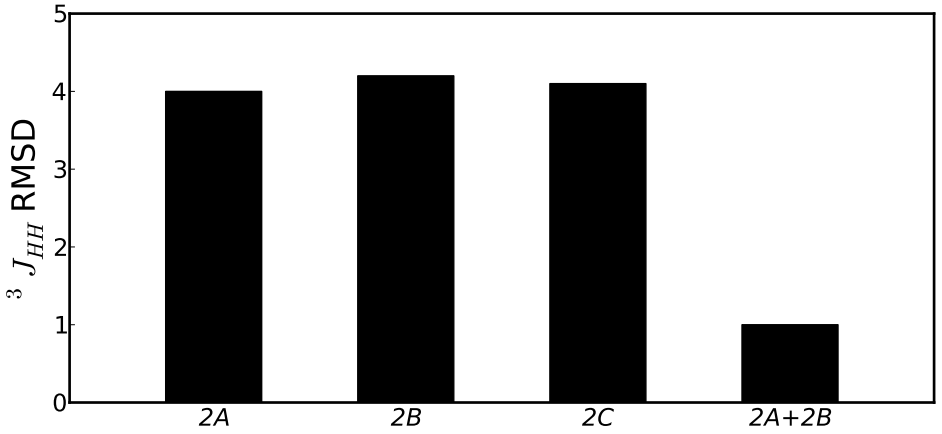


FIGURE 3.9: Agreement of experimental $^3J_{HH}$ and computed values of Salsolidine conformers **2A**-**C**, expressed as RMSD, in Hz. Computed $^3J_{HH}$ values were determined with the Haasnoot-Altona equation. The RMSD of the 50 : 50 **2A**+**2B** mixture is also shown.

3.3.3 RDC analysis of the conformational exchange of salsolidine in solution

RDC of Salsolidine hydrochloride were determined on a sample oriented in the CPCL-LC lyotropic phase.

3.3.3.1 The Cetylpiridinium Chloride / NaCl / Hexanol Liquid Crystal (cpcl-lc)

The lyotropic liquid crystal *Cetylpiridinium Chloride / NaCl / Hexanol* (CPCL-LC) was introduced by Prosser *et al.* for the alignment of biomolecules,[157] in an attempt to overcome the protein binding problems showed by the dilute bicelles. Combination of *cetylpiridinium chloride* (CPCl) and hexanol in aqueous NaCl solutions can result in a wide range of phases, including an (*undesired*) isotropic micellar phase and a (*desired*) anisotropic lamellar phase. The lamellar phase results when equal weight fractions of CPCl and hexanol are combined in 200 mM NaCl to a total liquid crystal concentration between $1 - 5$ % (w/w). The CPCl lyotropic phase has been extensively used for the weak alignment of biomolecules and oligosaccharides.[158–161]

Materials for making this liquid crystal are inexpensive and the protocol to yield the lamellar phase is very easy to follow. The CPCL-LC liquid crystal is stable for long periods of time.

3.3.4 RDCs of 2 were extracted from a set of ^{1}H-coupled 1D ^{13}C and 2D HSQC spectra.

3.3.4.1 Alignment of salsolidine

RDCs from a 200 mM solution of racemic **2** were extracted from the difference in C–H couplings in an isotropic ($^{1}J_{CH}$) and anisotropic ($^{1}T_{CH} = {}^{1}J_{CH} + {}^{1}D_{CH}$) sample (Table 3.6). Deuterated water was used as the isotropic medium and the CPCL-LC prepared in D_2O as the anisotropic one.

We showed previously (Section 3.3.2.1 and Figure 3.4) that the N atom of salsolidine is protonated in the isotropic sample. An equivalent demonstration can not be done in the liquid crystal as changes of pH caused the disruption of the phase. Nevertheless, the ^{1}H and ^{13}C spectra of the anisotropic sample closely resemble those of the isotropic state, indicating that the conformational and protonation states are not perturbed by the alignment medium. The similarity of ^{13}C chemical shifts in both media is easily appreciated in Figure 3.10.

3.3.4.2 RDC extraction

Couplings were extracted from a set of 1D and 2D spectra. Most of the couplings were extracted from the gated-decoupled ^{13}C-detected experiment. Figure 3.10 shows a comparison of the isotropic and anisotropic samples. The stereochemical assignment

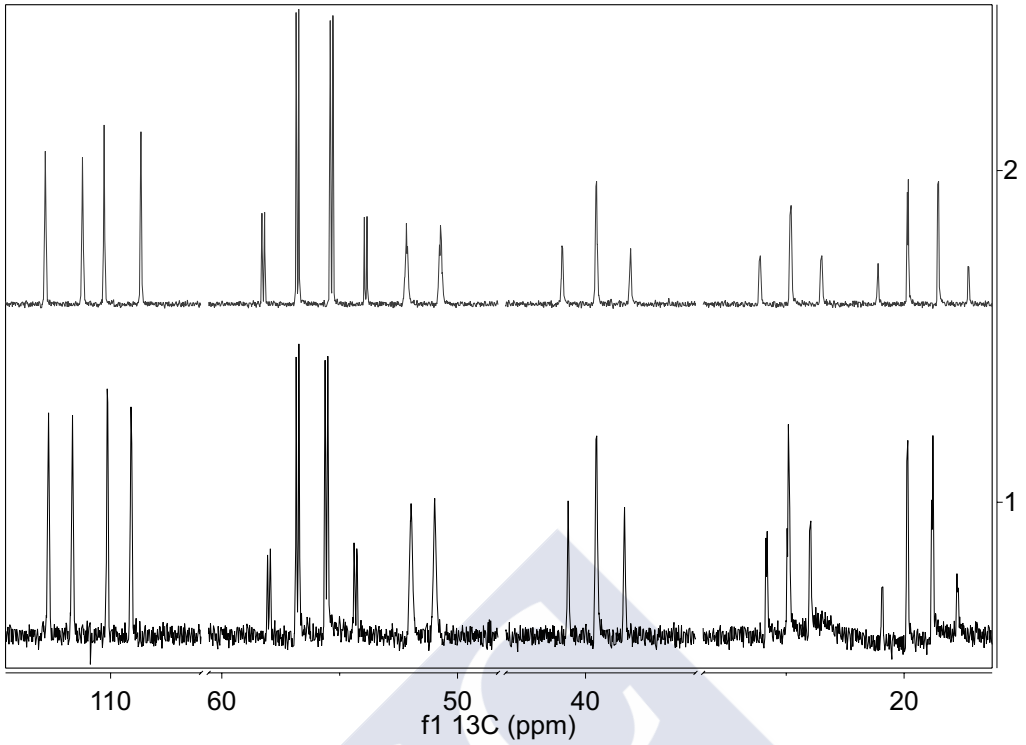


FIGURE 3.10: Gated-decoupled ^{13}C spectra of Salsolidine (**2**) at 125 MHz. (2) top, isotropic sample; (1) bottom, anisotropic sample.

of couplings of methylenic C–H vectors is not straightforward. The ^{13}C resonances of methylenes C3 and C4 appear as triplets in the 1D gated-decoupled ^{13}C spectrum of the isotropic sample, meaning that the α and β C–H vectors have equal values of the isotropic $^1J_{CH}$ (Figure 3.11). The situation is different in the spectrum of the anisotropic sample, where both methylenes appear as doublet of doublets, indicating that the $^1T_{CH}$ of the α and β C–H vectors are not equal (Figure 3.11). Unfortunately, their values can not be assigned directly from the ^{13}C gated-decoupled experiment to the α or β configuration.

Values of the C–H3α and C–H3β vectors were distinguished from the F_2-coupled CLIP-HSQC[53] spectrum, but no stereochemical assignment was possible for the overlapped benzylic H4 protons (Figure 3.12). Experimental $^1J_{CH}$ and $^1D_{CH}$ are given in Table 3.6. Consequently, we devised a protocol to average methylene CH couplings within the RDC fit procedure. The procedure is similar to that proposed by our group to average couplings of fast-rotating methyl and phenyl groups,[51] which has been discussed in Chapter 1 (Section 1.5) and will be explained in the next Section 3.3.4.3.

F2-coupled HSQC spectra are known to suffer from some limitations (see Section 1.7).

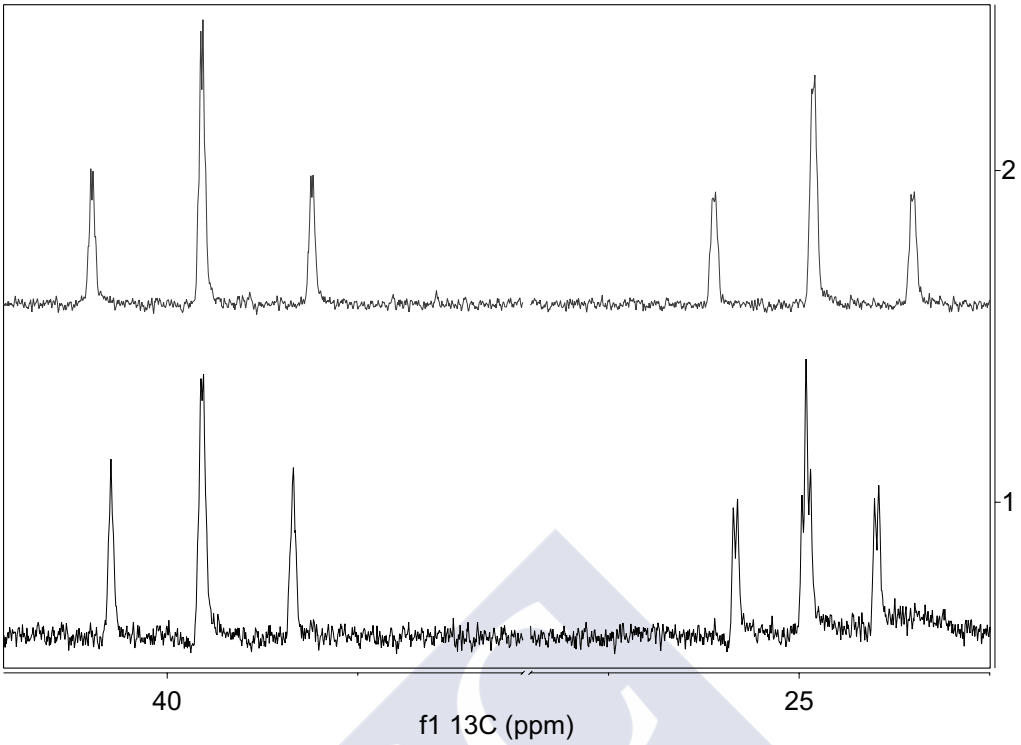


FIGURE 3.11: Methylene expansion of gated-decoupled ^{13}C spectra of Salsolidine (**2**) at 125 MHz. (2) Top, isotropic; (1) bottom, anisotropic.

One of such limitations is the difficulty to accurately determine the center of complex multiplets due to the superposition of numerous dipolar couplings. Moreover, linewidth may be affected by strong coupling and homonuclear dipole-dipole coupling that can be problematic in aligned samples. Therefore, we also acquired a F_1-coupled HSQC spectrum (Figure 3.13) of the anisotropic sample (Figure 3.12). There were no significant differences in the values of $^1T_{CH}$ obtained in the two versions of the HSQC and in the 1D gated-decoupled ^{13}C spectrum. Regarding the methylene groups, it must be noted that the *sum* of their two CH couplings is obtained from the F_1-coupled HSQC. All available one-bond C–H couplings were extracted (Table 3.6) except those of the methoxy groups, as their mobility make them useless for the determination of the alignment tensor.

3.3.4.3 Couplings of geminal protons are included in RDC fits as a half-sum when stereochemical assignment is not feasible

Usually, the RDCs of methylene groups can be measured, even though the stereospecific assignment of the geminal protons is often not known. In order to have as many experimental restraints as possible, it is desirable a protocol to introduce these data

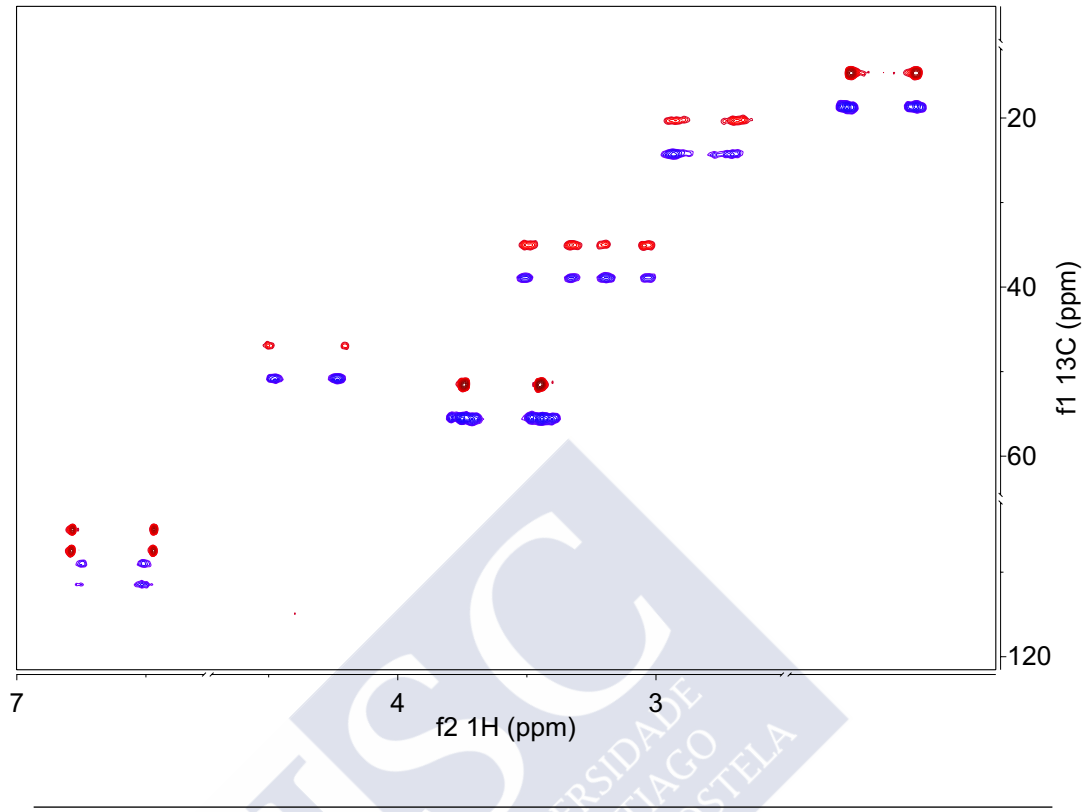


FIGURE 3.12: Overlay of CLIP-HSQC (CLean In Phase - HSQC) spectra of salsolidine (**2**) . Red, isotropic; blue, anisotropic.

TABLE 3.6: Experimental scalar and dipolar one-bond couplings of salsolidine, in Hz.

Vector	$^1J_{CH}$	$^1D_{CH}$
C1–H1	145.5	−19.1
C3–H3α	145.1	7.3
C3–H3β	145.1	3.4
C4–H4α/β[a]	130.9	−18.3
C4–H4α/β[a]	130.9	−11.4
C5–H5	158.8	−30.4
C8–H8	157.3	−30.4
C9–H9	129.4	−5.1

[a] The stereochemical assignment of the H4 α/β protons was unknown, hence their values were averaged in the calculations.

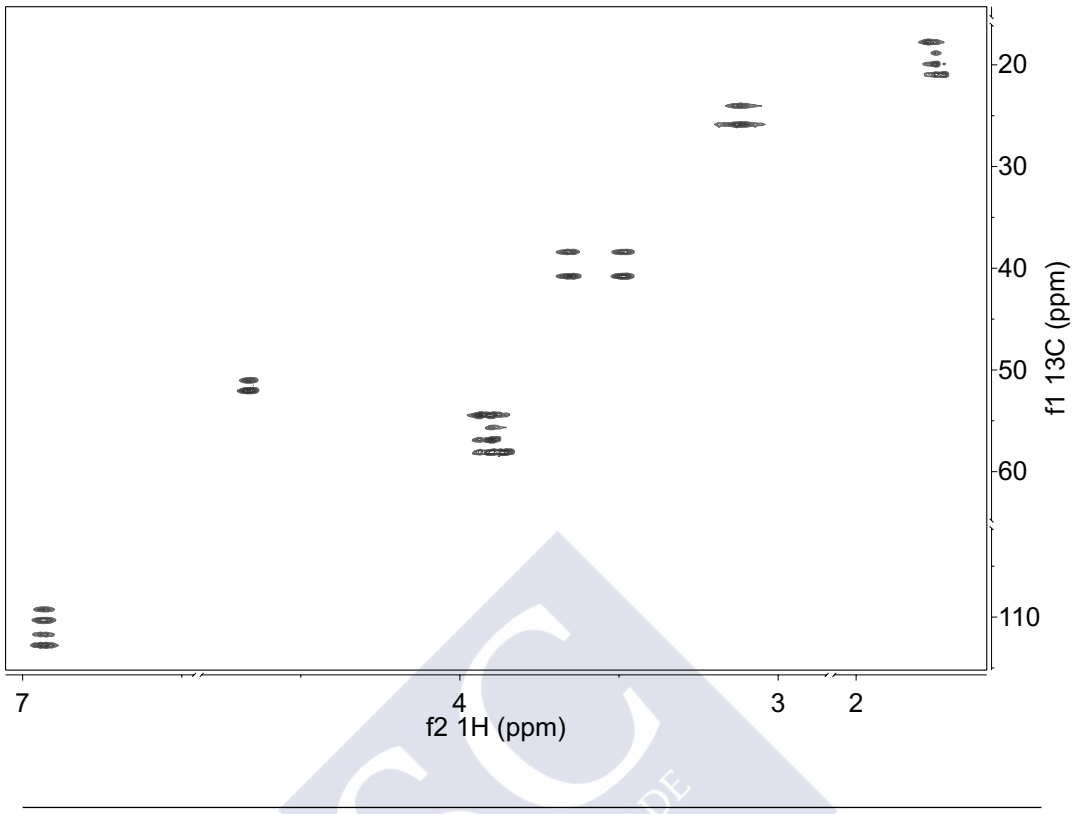


FIGURE 3.13: F_1-coupled HSQC spectra of isotropic Salsolidine (**2**). No *J*-evolution multiplication module was used.

in the calculations.This is especially critical when dealing with molecules with sparse C–H vectors and many degrees of freedom, as is usually the case of flexible small organic molecules. Averaging of the methylene C–H couplings is the solution. A protocol had been developed in our group to deal with the averaged RDC of fast-rotating groups (methyl and phenyl).[51] The mathematical basis has been discussed in Chapter 1 (Section 1.5). We will show in the next how methylene RDCs can be introduced in the fit when their stereospecific assignment is not available. This is even possible if the individual values are not accessible, provided the *sum* of their couplings can be measured. The F_1-coupled HSQC experiment is very useful for this goal, as the sum of methylene one-bond C–H couplings is directly extracted. The procedure devised here will be applied to salsolidine in this chapter and to other molecules in the next chapters.[54]

It was shown in Chapter 1 that the Residual Dipolar Coupling D between nuclei I and S is expressed in matrix form as

$$D = \frac{\kappa}{R^3} \vec{r}^T \hat{\mathbf{A}} \vec{r}, \tag{3.2}$$

where $\vec{r}$ is the internuclear unit vector, $\hat{\mathbf{A}}$ is the alignment tensor, and $\kappa =$

$-\frac{3}{8\pi^2}\mu_0\gamma_I\gamma_S\hbar$ (see Chapter 1 for further details). Since $\hat{\mathbf{A}}$ is a symmetric and traceless tensor, Equation (3.2) can be expanded leading to the following linear system

$$\vec{d} = \hat{\mathbf{M}}\hat{\mathbf{A}},$$

$$\begin{pmatrix} w_1 d_1 \\ \cdots \\ w_c d_c \\ w_d d_d \\ \cdots \\ w_i d_i \\ w_j d_j \\ w_k d_k \\ \cdots \\ w_p d_p \\ w_q d_q \\ \vdots \\ w_N d_N \end{pmatrix} = \hat{\mathbf{M}} \begin{pmatrix} A_{xy} \\ A_{xz} \\ A_{yy} \\ A_{yz} \\ A_{zz} \end{pmatrix}, \tag{3.3}$$

in which $d_i = D_i\frac{R^3}{\kappa}$ is the *reduced* RDC and $\hat{\mathbf{A}}$ is the alignment tensor expressed as a column vector containing its five independent components. This equation is equivalent to Equation (1.82), except that here additional elements with indexes c and d are shown explicitly to account for the averaging of methylene couplings.

The matrix $\hat{\mathbf{M}}$ in Equation (3.3) represents the molecular geometry and is composed by director cosine products. Explicitly

$$\hat{\mathbf{M}} = \begin{pmatrix} 2w_1x_1y_1 & 2w_1x_1z_1 & w_1(y_1^2-x_1^2) & 2w_1y_1z_1 & w_1(z_1^2-x_1^2) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 2w_cx_cy_c & 2w_cx_cz_c & w_c(y_c^2-x_c^2) & 2w_cy_cz_c & w_c(z_c^2-x_c^2) \\ 2w_dx_dy_d & 2w_dx_dz_d & w_d(y_d^2-x_d^2) & 2w_dy_dz_d & w_d(z_d^2-x_d^2) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 2w_ix_iy_i & 2w_ix_iz_i & w_i(y_i^2-x_i^2) & 2w_iy_iz_i & w_i(z_i^2-x_i^2) \\ 2w_jx_jy_j & 2w_jx_jz_j & w_j(y_j^2-x_j^2) & 2w_jy_jz_j & w_j(z_j^2-x_j^2) \\ 2w_kx_ky_k & 2w_kx_kz_k & w_k(y_k^2-x_k^2) & 2w_ky_kz_k & w_k(z_k^2-x_k^2) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 2w_px_py_p & 2w_px_pz_p & w_p(y_p^2-x_p^2) & 2w_py_pz_p & w_p(z_p^2-x_p^2) \\ 2w_qx_qy_q & 2w_qx_qz_q & w_q(y_q^2-x_q^2) & 2w_qy_qz_q & w_q(z_q^2-x_q^2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 2w_Nx_Ny_N & 2w_Nx_Nz_N & w_N(y_N^2-x_N^2) & 2w_Ny_Nz_N & w_N(z_N^2-x_N^2) \end{pmatrix}, \tag{3.4}$$

where x, y, and z are the director cosines of the unit vector $\vec{r}$. Each element of $\hat{\mathbf{M}}$ and $\vec{d}$ is preceded by a weighting factor w_i that is used to average certain elements. Typically, non-averaged couplings have the default value $w_i = 1$.

The inclusion of $^1D_{CH}$ of fast-rotating methyl and phenyl groups is addressed (Chapter 1) by the individual entries indexed with i, j , k (methyl), and p, q (phenyl). In an analogous way, additional terms with indexes c and d are set up in the matrix expansion to average methylene geminal C–H vectors. Methylene $\hat{\mathbf{M}}$ elements are transformed into their corresponding averaged values and weighted with factors $w_c = w_d = \frac{1}{2}$ as follows

$$M_c = M_d = w_c(2x_c y_c + 2x_d y_d). \tag{3.5}$$

Accordingly, the corresponding averaged reduced dipolar couplings are defined as

$$\langle d_c \rangle = \langle d_d \rangle = w_c(d_c + d_d). \tag{3.6}$$

Following this procedure, unassigned C–H4 RDCs can be included in the fit.

3.3.4.4 Fit of RDCs to Single Structures

A total of 8 experimental $^1D_{CH}$ of salsolidine were available (Table 3.6). C–H couplings of positions 9 (methyl) and 4 (methylene α/β) were averaged as indicated in the previous section (see Section 3.3.4.3 and Chapter 1).[51] By using this approach, only the average $^1D_{CH}$ coupling of the two C4 protons was taken into account in the fit, thus leaving a total of seven experimental data.

Initially, each of the conformers **2A-C** was fit separately to the experimental RDCs using the singular value decomposition (SVD) as implemented in a developer version of MSPIN.[41–43] The quality of fit (Table 3.7) is expressed in terms of the Cornilescu quality factor Q_C (Chapter 1).[44]

TABLE 3.7: Quality of fit of RDCs to salsolidine sole structures.

Structure	Q_C	*c.n.*[a]
2A	0.205	11.08
2B	0.176	15.49
2C	0.112	24.82

[a] *c.n.*: condition number (see Chapter 1).

The robustness of the SVD procedure is expressed by the condition number (*c.n.*, Table 3.7),[162] i.e. the ratio between the largest and smallest singular values in the SVD.

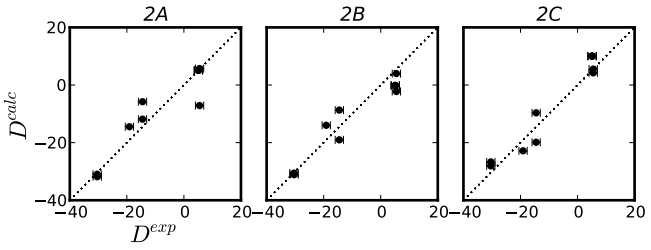


FIGURE 3.14: Plots of experimental (D^{exp}) *vs.* back-calculated (D^{calc}) $^1D_{CH}$ RDCs resulting from the fit of the sole salsolidine candidate conformers **2A**-**C**. The error bars are set to 1.5 Hz.

The condition number expresses the sensitivity of the system to variations in the experimental RDCs. Consequently, the larger the condition number, the more sensitive will be the alignment tensor to the uncertainty of experimental RDCs. Conformer **2C** has the lowest value of Q_C (i.e. the best fit) but the largest *c.n.*, indicating that this result should be interpreted with caution. The relatively large values of Q_C reflect that none of the conformers fits well to the experimental data. The poor fit between experimental and back-calculated RDC can be appreciated in Figure 3.14. Furthermore, the geometry that fits best (**2C**) has the worst condition number and the highest computed energy (+1.9 kcal/mol). Considering that conformational averaging was in fact expected from the previous analysis of chemical shifts, $^3J_{HH}$ couplings and DFT computations, we set to fit the RDC data to mixtures of conformers.

3.3.4.5 Fit of RDCs to Multiple Conformers. Single-tensor approximation

Fit of RDC data to a mixture of two or more conformers in equilibrium is challenging. In principle, the orientational probability of each conformer is independent from the others, meaning that there is one alignment tensor per conformer. In practice, this means a large number of unknowns, as each tensor has 5 independent components. For two conformers in equilibrium, this means a total of 11 unknowns, namely the 10 tensor components and one population (molar fraction). Due to the paucity of the RDC data available (7 experimental RDC values), multitensor fit of salsolidine would not be possible.

Nevertheless, since the conformational change in salsolidine results in a small perturbation of the overall molecular shape (see Figure 3.3), the single tensor approximation seems reasonable in this case. This reduces the number of unknowns to six: the five tensor components and one conformer population. Given that the C5–H5 and C8–H8 vectors on the benzene ring are nearly parallel, a total of six experimental $^1D_{CH}$ are linearly independent, hence equaling the number of unknowns.

Additionally, the **2C** conformer could be discarded due to its high DFT computed energy and previous NMR work, $^{3}J_{HH}$ analysis. Nevertheless, we intended to rely only on NMR structural restraints, therefore all candidate conformations were taken into account in terms of two-membered ensembles.

Previous NMR studies, as well as $^{3}J_{HH}$ analysis and ^{1}H and ^{13}C chemical shifts, point to an equilibrium of two conformations in solution. Based on this, populations of the two relevant conformers of salsolidine hydrochloride can be determined with the proposed single-tensor approximation. In order to apply the single tensor approximation, a common frame of reference is needed for both conformers. A reasonable assumption consists in superimposing all heavy atoms of **2A-C** conformers using a least-squares minimization procedure. This ensures a minimal movement of the structures with respect to the common external frame. The least squares superimposition was done with a previously described SVD-based algorithm using Cartesian-weighted coordinates of all heavy atoms.[125]

TABLE 3.8: RDC fit of all possible salsolidine pair conformers ensembles.

Ensemble	Q_c	*c.n.*	$p_1 : p_2$
2A+**2B**	0.047	6.36	51:49
2A+**2C**	0.045	6.39	61:39
2B+**2C**	0.121	18.24	8:92

All the solutions considered had the same assignment of all diasterotopic protons.

RDC data were fit with MSPIN. The protocol simultaneously optimizes conformer populations and tensor components in an iterative manner. Fit of the two-membered ensembles resulted in a better Q_C than that obtained with either of the separate conformers (Table 3.8 and Figure 3.16). Fits of **2A**+**2B** and **2A**+**2C** fits were indistinguishable in terms of Q_C and *c.n.*, whilst **2B**+**2C** fits clearly worse (Figure 3.15).

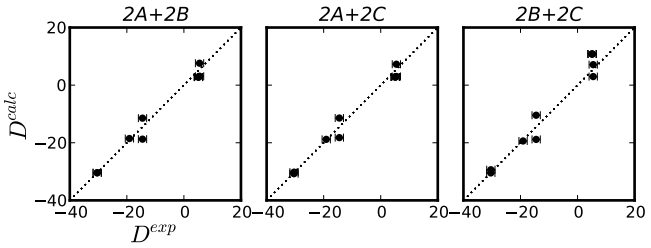


FIGURE 3.15: Plots of experimental (D^{exp}) *vs.* back-calculated (D^{calc}) $^{1}D_{CH}$ RDCs resulting from the fit of salsolidine 2-membered ensembles. The error bars are set to 1.5 Hz.

Further discrimination of the ensembles can be done based on $^{3}J_{HH}$ analysis. Following the previously presented methodology, ensemble averaged $^{3}J_{HH}$ couplings were predicted with MSPIN[42,43] and fitted to the experimental ones. Ensemble **2A**+**2B** fits

better to $^3J_{HH}$ couplings (Figure 3.16, $RMSD$ = 1.0 Hz), than **2A**+**2C** and **2B**+**2C** ($RMSD$ = 1.3 and 4.1 Hz, respectively). Populations of **2A**+**2B** converged to 51 : 49, very similar to the predicted values from DFT-computed energies (Table 3.1).

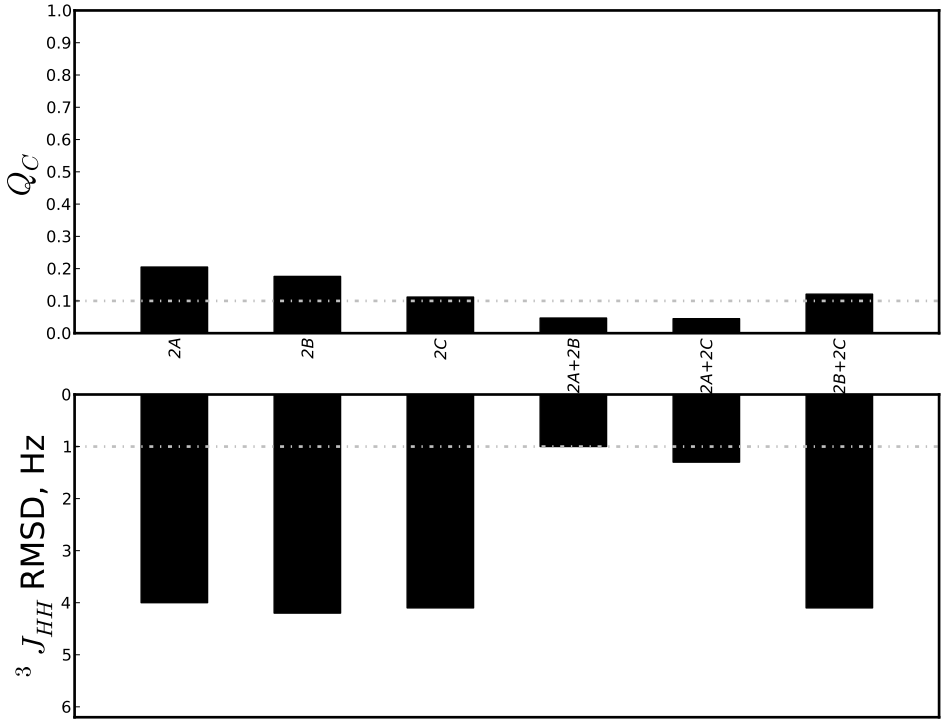


FIGURE 3.16: Quality of RDC and $^3J_{HH}$ fit to salsolidine sole conformers and 2-membered ensembles.

To test the impact of experimental errors on the computed populations and Q_C, a bootstrapping estimate was done, similarly to previous work by Sánchez–Pedregal *et al.*[51] Bootstrapping is a useful tool to estimate the statistics of the RDC fit. This method relies on a Gaussian replacement resampling with the same size of the original distribution (same number of experimental and resampled RDCs), which is done n times. The resultant simulated data sets are then fitted to the structures and Q_C and p_i are compared with those determined with the experimental data set to obtain the statistics (see Chapter 1, Section 1.4.5).

Experimental RDC values were randomly changed following a Gaussian distribution over a resample of 250 datasets. A value of 1.5 Hz was considered as conservative estimate of the experimental error, and set as the standard deviation of the measured RDCs. Each one of these synthetic datasets was fitted to the **2A**+**2B** ensemble. Back-calculated RDCs, populations and alignment tensor elements were kept. These 250 simulations gave consistently similar values for conformer populations, alignment

tensors and quality factors Q_C, indicating that the result is not very sensitive to experimental uncertainty.

On average, population of conformer **2A** was 49.5% (standard deviation, $\sigma = 4.2\%$) and the Q_C factor was 0.056 ($\sigma = 0.022$).

3.4 Conclusion

Early work on salsolidine hydrochloride had identified averaging of chemical shifts and $^3J_{HH}$ couplings, which was ascribed to a fast-exchange equilibrium between two half-chair conformers.[147] Noteworthy, this analysis was not supported by computational simulations or conformational search procedures, and lineshape analysis was not done to analyze the individual $^3J_{HH}$ of C3 protons.

In our approach, we address the conformational equilibrium of salsolidine with the aid of RDC-enhanced NMR and computational tools. Molecular mechanics conformational search followed by DFT optimization furnished three low energy conformers (Table 3.1 and Figure 3.3). The two lowest-energy ones being half-chairs, which are expected to be in 50 : 50 ratio according to Maxwell-Boltzmann statistics (Table 3.1).

Lineshape analysis of the 500 MHz ^{1}H spectrum gave values in the range of $6.0 - 6.5$ Hz for all four $^3J_{HH}$ couplings, which indicates extensive conformational averaging.

Salsolidine was proven to be protonated in the experimental conditions (Figure 3.4) and conformational equilibrium is not expected to be affected by the alignment medium as demonstrated by the unaffected chemical shifts of all resonances (Figure 3.10), particularly the C3, as predicted by DFT-GIAO computations (Table 3.2).

Fit of RDC to an ensemble composed by **2A** and **2B** conformers resulted in better Q_C factors than fit to any of the sole conformers **2A**-**C** or other ensembles (Figure 3.16). The error in the calculated conformer populations was estimated with the bootstrapping method. Assuming an experimental error of 1.5 Hz in the RDC, average conformer populations are estimated as 49.5 : 50.5 with $s.d. = 4.2\%$. This is in very good agreement with the computed energies of conformers **2A**-**B**, which barely differ in 0.1 kcal/mol.

The method of averaging of unassigned geminal C–H vectors in methylenes is expected to be of general use to study other molecules.

In conclusion, all classical NMR restraints ($^3J_{HH}$ and chemical shifts) support the early findings about salsolidine conformational equilibrium in solution. RDC multi-conformer fit supports the previously proposed fast exchange between the two half-chair conformers **2A** and **2B**.

3.5 Materials and Methods

3.5.1 Salsolidine hydrochloride

Racemic salsolidine hydrochloride **2**·HCl was purchased from Sigma-Aldrich and used without further purification.

3.5.2 Conformational Search

The conformational space of **2** was explored with the MM3 force field and the stochastic search procedure as implemented in PCMODEL.[111,130,150,151] The so-obtained conformations were then minimized at the DFT level of theory using the M052X[132] meta-GGA-hybrid functional and the 6-31+G** basis set. Solvation was taken into account by using the Onsager model[163] with a solvation radium of 4.95 Å and a water relative dielectric constant $\epsilon = 78.39$. GIAO[164] chemical shifts were computed using the GGA OPBE[165,166] functional and the specialized pcS-2[154] basis set on M052X structures, and referenced to tetramethylsilane. All DFT computations were performed with the Gaussian03 package.[167] DFT derived energies are shown in Appendices (Appendix B, Section B.2.2).

Conformational search and DFT calculations were conducted by Dr. Armando Navarro-Vázquez.

3.5.3 Preparation of the CPCl Liquid Crystal

The CPCl liquid crystal CPCL-LC was prepared with 1 g of D_2O (99.9 %, Spectra), 1.16 % (*w*/*w*) NaCl, 2.63 % (*w*/*w*) cetylpyridinium chloride (CPCl, Sigma-Aldrich) and 2.5 % (*w*/*w*) *n*-hexanol. All reagents were used as purchased without further purification.

NaCl (11.6 mg) was dissolved in D_2O (1.0 g) in a magnetically stirred glass vial. When all the NaCl was dissolved, CPCl (26.3 mg) was added while stirring vigorously and the mixture was heated at 70 °C for six minutes. The mixture was then cooled down to r.t., then *n*-hexanol (25 mg) was added and the mixture was heated again at 70 °C

for several minutes and allowed to cool to r.t. slowly. Stirring was maintained all the time.

3.5.4 NMR

Assignment

NMR experiments were recorded on a Varian 500 MHz spectrometer equipped with a 5 mm ID / PFG probe (50-202 MHz). Salsolidine resonances were assigned from a set of standard ^{1}H, ^{13}C and HSQC spectra. Phase-sensitive NOESY in D_2O was recorded with a mixing time of 500 ms. The assignment spectra can be seen in Appendix A, Section A.2.1 and the assignment in Appendix A, Section A.2.2.

Residual dipolar couplings

Residual dipolar couplings (RDC) were determined from a set of ^{1}H-coupled 1D and 2D spectra recorded on the isotropic and anisotropic samples. Gated-decoupled ^{13}C spectra were acquired with the standard Varian sequence based on *s2pul*, with 32768 complex points and recovery delay d_1 of 1 s. The Bruker used to record the F_2-coupled and F_1-coupled HSQC, can be found in Appendix C

CLIP-HSQC were acquired as 200* (^{13}C) $\times$ 2048* (^{1}H) data matrices, where N* refers to N complex pairs, and spectral widths of 4006 $\times$ 15723 Hz, respectively, using 32 transients per FID and 1.4 s delay between scans.

F_1-coupled HSQC were acquired as 800* (^{13}C) $\times$ 700* (^{1}H) data matrices, where N* refers to N complex pairs, and spectral widths of 4006 $\times$ 15723 Hz, respectively, using 32 transients per FID and 1.4 s delay between scans.

The deuterium quadrupolar splitting ($|\Delta\nu_Q|$) of the solvent was checked before and after recording the ^{13}C gated-decoupled and F_2-coupled HSQC spectra to assess the integrity of the aligned sample. Aligned samples were stable over several days at r.t. (the deuterium $|\Delta\nu_Q|$ remained constant at 18.5 Hz).

NMR experiments for the determination of the protonation state were recorded on a Bruker DPX 250 MHz spectrometer equipped with a 5 mm inverse probe.

Spectra were processed and analyzed with the MestReNova software.[42]

All NMR experiments, save for those recorded at 250 MHz, were acquired by Dr. Víctor Sánchez-Pedregal and Ramón Gesto-Rey.

3.5.5 RDC Fits

RDC analysis was performed using the MSPIN program.[42,43] The procedure is outlined in Chapter 1. Input and output data are reported in Appendices. Molecular coordinates of conformers **2A-C** (Appendix B, Section B.2.1), RDC input tables in MSPIN-ready format (Appendix B, Section B.2.3), and program outputs (Appendix B, Section B.2.4).

Bootstrapping

Bootstrapping estimates were performed with an unreleased developer version of MSPIN software. 256 data-sets where resampled and fitted to candidate structures as described in Chapter 1, Section 1.4.5 and analyzed to calculate the descriptive statistics of Q_C and p_i distributions.

Chapter 4

Application of one-bond RDC to the determination of absolute configuration and conformation in flexible molecules

4.1 Eburnamine-vincamine alkaloids

The eburnamine-vincamine alkaloids can be found in the plant family *Apocynaceae*. The main types of these alkaloids have been isolated from several *Hunteria* and *Vinca* species, including the *Tabernaemontanoideae* tribe. Quite a number of these alkaloids exert different pharmacological activities, from cell multiplication to cardiovascular system and brain functions.

The eburnamine-vincamine alkaloids group can be classified into two major parent skeletons, namely (−)-eburnamine and (+)-vincamine (Figure 4.1), which originate from the tryptamine metabolism. The five ring skeleton system is characteristic to these alkaloids. In this Chapter, alkaloids from this family are numbered following the biogenic numbering proposed by le Men and Taylor.[168]

In general, compounds having a 20*R*, 21*R* (20β, 21β) configuration are classified as *eburnane* type ("eburna" skeleton), whereas those having 20*S*, 21*S* (20α, 21α) configuration are classified as the *vincane* type ("vinca" skeleton).

Vincamine and other members of this family of compounds are used as nootropic (*brain enhancing* drugs) dietary supplements. Vincamine can be found in Europe as

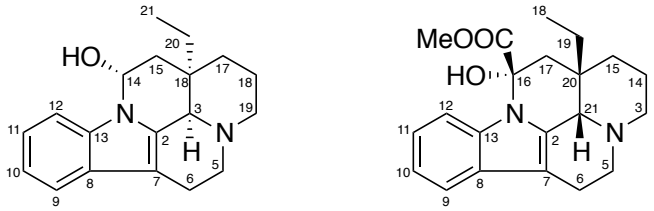


FIGURE 4.1: Eburnamine-vincamine alkaloids representative structures. Atoms numbered following the biogenic process, ring numbering are the same as for **3** and **4**, shown in Figure 4.2.

Oxybral SR, marketed by GlaxoSmithKline. Other important (semi-synthetic) supplement is vinpocetine, alkaloid that can be found under the brand names Cavinton and Intelectol.

4.2 19-OH-(−)-Eburnamonine natural alkaloid shows vasodilator properties

Eburnamonine was first isolated by Raymond-Hamet and co-workers in 1959 from *Bonafusia macrocalyx* as a mixture of (−)-eburnamonine and (+)-eburnamonine.[169] Care shall be taken in early references, as the genus *Bonafousia* (Family *Apocynaceae*, tribe *Tabernaemontanoideae*) was emended by Allorge in 1983.[170] Some reports have been published on the isolation of a number of indole alkaloids of the iboga type[171,172] and of the lignan glycoside bonafusioside type[173] from a collection of *B. macrocalyx* from French Guiana

A new indole alkaloid of the eburnan type, namely 19-OH-(−)-eburnamonine (**4**, Figure 4.2), was isolated from a collection of *Bonafousia macrocalyx* (Müll. Arg.) Boiteau & L. Allorge (Basionym: *Tabernaemontana macrocalyx* Müll. Arg.) from Loreto, Perú. Isolation was carried out by Roxana Sifuentes and Dr. Helena Maruenda, from Departamento de Ciencias — Sección Química, Pontificia Universidad del Perú.

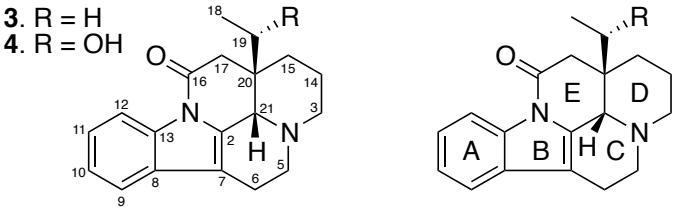


FIGURE 4.2: Eburnamonine, **3**, and its 19-OH derivative, **4**.

Initial studies on the biological effects of indole alkaloids reported increased glucose and cerebral oxygen consumption, as well as protection against cerebral edema.[174] Eburnamonine has been clinically used as a drug for improving cerebral circulation and metabolism. The mechanism of action *in vivo* possibly includes the stimulation of cholinergic neurotransmission. Particularly, (−)-eburnamonine has been shown to be a subtype specific allosteric effector of human recombinant muscarinic receptors.[174,175]

4.2.1 Evidences of C19 Hydroxylation

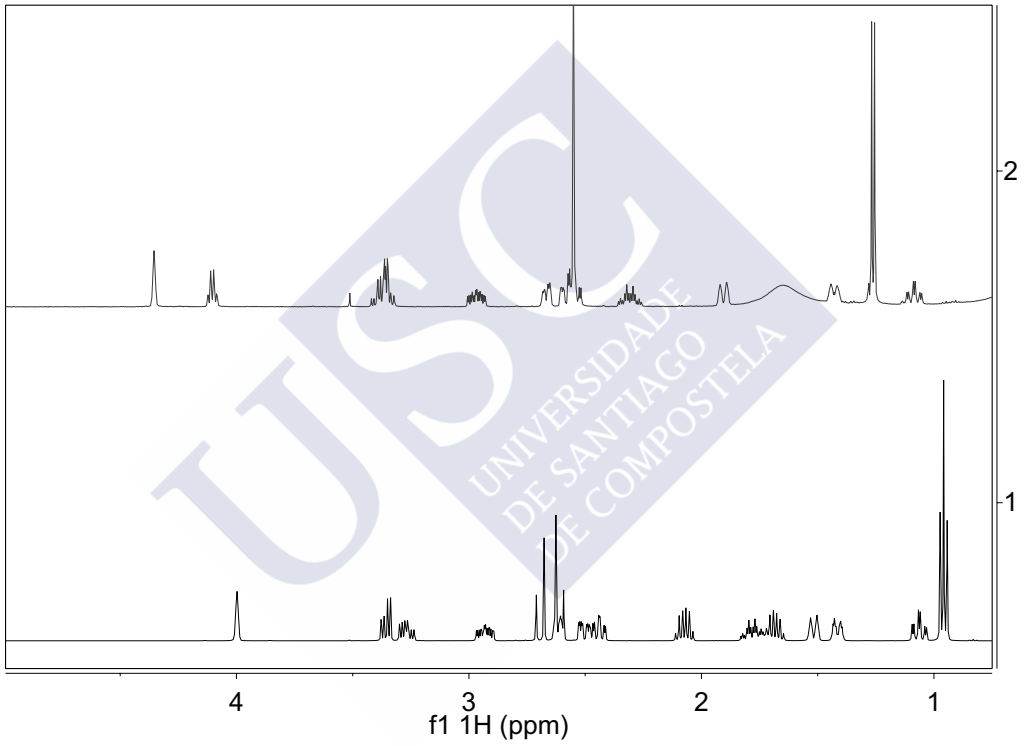


FIGURE 4.3: Eburnamonine, **3** (spectrum 1), and its 19-OH derivative, **4** (spectrum 2) ^{1}H 1D NMR spectra. Notice the disappearance of **3**'s methylenic C19 signals at 1.6 ppm (bottom, 1) and the new CH quartet at 4.1 ppm in **4** (top, 2), see text.

In Figure 4.3, 1D ^{1}H NMR spectra of the new indole alkaloid **4** and a commercial sample of **3**, acquired in $CDCl_3$, are shown. The ^{1}H NMR spectrum of **4** shows an evident similarity to the one of the commercial sample of (−)-eburnamonine (**3**, see Section 4.6). However, the ethyl group spin-system in **3** furnished a quarted-triplet spin system at 1.78 ppm, that is not present in the ^{1}H spectrum of **4**. New features in the **4** ^{1}H spectrum, as a quartet at 4.08 ppm (CH) and a doublet at 1.24 ppm (CH_3), clearly point to the hydroxylation of the parent compound **3** at the C19 position.

Due to the significant spectral similarities of **4** and **3**, the complete ^{1}H and ^{13}C NMR assignment was performed by comparing the NMR data **4** with the previous assignment of a commercial (−)-eburnamonine (**3**) sample previously done by Dr. Roberto R. Gil using a combination of 1D ^{1}H and ^{13}C, COSY, NOESY, multiplicity edited-HSQC and HMBC spectra. All the NMR data are fully consistent with the structure depicted in Figure 4.2.

High resolution mass spectroscopy (HRMS-ESI, Figure 4.4) of **4** natural sample resulted in an exact mass (m/z) of 311.1748 ± 0.0001 whilst eburnamonine (**3**) mass is 294.1732. The MW of **4** is 16 Da larger than that of **3**, suggesting the presence of an additional oxygen atom. The ^{1}H NMR differences also point to the presence of an additional oxygen atom on C19. It is safe to conclude that **4** is a 19–OH derivative of **3**.

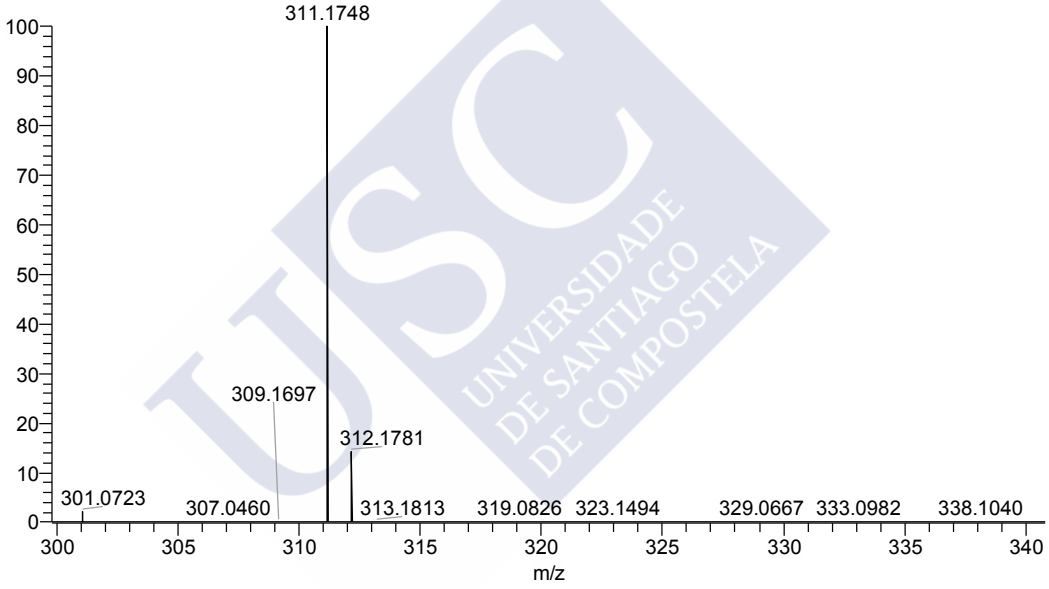


FIGURE 4.4: High resolution mass spectrum of 19-OH-(−)-eburnamonine (**3**).

The close spectral similarity between **4** and **3** and the presence of a NOE correlation in **4** between H19 and H21 (Figure 4.5) point to a *cis* D-âĂŞE ring fusion in the **4** skeleton (the same as in **3**). However, a new stereogenic center is created at carbon C19. The relative configuration of this new stereogenic center cannot be easily determined since NOE and $^{3}J_{HH}$ coupling constants analysis are hampered by the lack of protons at position C20 and by the presence of rotamers around the C19–C20 bond.

An additional NOE correlation between H19 / H17 was found in **4**, which suggests the presence of either a (−)-*synclinal* or a (+)-*synclinal* rotamer (Figure 4.7). Conformational exchange between rotamers, or presence of an *antiperiplanar* rotamer cannot

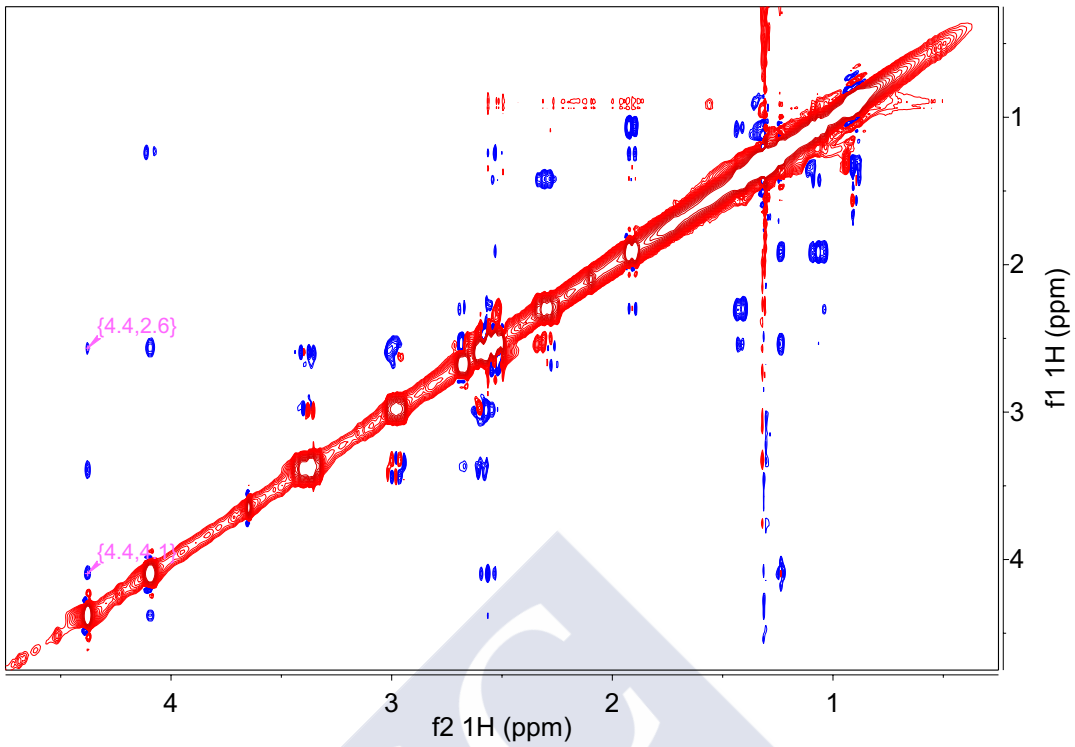


FIGURE 4.5: 19-OH-(−)-eburnamonine (**4**) geNOESY spectrum acquired at 500 MHz field in $CDCl_3$ with τ_{mix} = 600 ms. H19 / H21 (4.4, 4.1) and H19 / H17 (4.4, 2.6) correlations are labeled with magenta arrows, see text.

be however excluded.

Since the parent eburnamonine has been isolated from natural sources as either levorotatory or dextrorotatory form,[169] it is not only necessary to determine the relative but also the absolute configuration of the new hydroxylated compound (**4**), these will be accomplished by means of TD-DFT computations once the relative configuration has been determined on the basis of NMR experiments.

4.2.2 Configurational / conformational space of 19-OH-(−)-eburnamonine

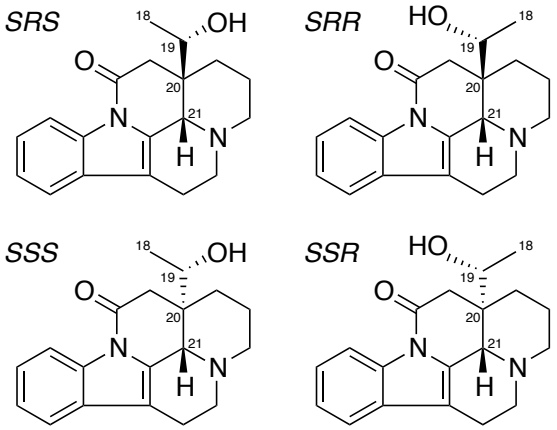


FIGURE 4.6: All possible configurations of **4** regarding C19 and C20 stereogenic centers.

As a first step for solving the relative configuration of **4**, all possible configurations were generated using the PCModel program,[111] by altering the corresponding bonds (C19–OH, C20–C19) whilst fixing the configuration of C21 as *S*, followed by an energy minimization step *in vacuo* (Figure 4.6).

Subsequently, the conformations of each of the candidate configurations were obtained by means of PCmodel[111] *in vacuo* stochastic Global-MMX[130] molecular mechanics computations, using an energy cutoff of 4.0 kcal/mol. The only conformations found are the different rotamers around the C19–C20 bond, which were classified as (−)-*synclinal*, (+)-*synclinal* and *antiperiplanar* after the C15–C20–C19–C18 dihedral angle, as shown in Figure 4.7.

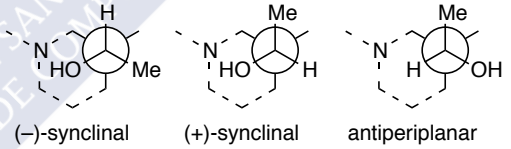


FIGURE 4.7: Schematic representation of all possible rotameric states of 19-OH-(−)-eburnamonine (**4**).

4.3 Relative configuration of 4 was determined by two independent methods

RDCs have proven to be a very useful tool for the determination of relative configuration in cases in which the classical NMR parameters, namely NOE correlations and

vicinal $^{3}J_{HH}$ couplings, provide ambiguous results to structural problems located in ^{1}H-deficient moieties.

Quantum mechanics (QM) calculations have been used for years to obtain reliable chemical shift predictions for given structures. *Ab initio* methodology for the calculation of chemical shifts has been recently reviewed.[176–178]

From the seminal paper of Ramsey in 1950[28] and the series of papers following, a number of equations for the calculation of NMR chemical shifts have been published. However, the routine calculation of isotropic shielding constants and therefore, NMR chemical shifts (see Chapter 1), was not feasible —at reasonable cost— until the mid 1970s, with the development of methods such as the Gauge Independent Atomic Orbital (GIAO),[179] which was further developed by Pulay and co-workers in the 1990s.[164] In the last decade, the computational prediction of ^{1}H and ^{13}C chemical shifts has became increasingly accurate and affordable. Such improvements derive from advances in computational techniques, mainly the development of the Density Functional Theory (DFT),[176–178] as well as in computers themselves, making the prediction of chemical shifts of routine use.

In the last years, *ab initio* predictions of chemical shifts have demonstrated a valuable tool for the determination of relative configuration of natural products.[180,181] In addition, the use of residual dipolar couplings (RDCs) have proven to be a powerful structural technique for the determination of the relative configuration.[33,45,182,183]

Based on these two recent developments and knowing that classical NMR approach can not solve the ambiguity of the configuration of C19, both alternative methods were used independently, furnishing the same configuration.

4.3.1 Relative configuration determination assisted by *ab initio* chemical shift calculations

The structures representing all the available configurational and conformational space of **4** were optimized at the DFT level (OPBE/6–31G(d))[153] in gas phase. Analytical frequencies were inspected for every calculation to ensure the nature of the stationary point.

Following Goodman[184] recommendations, *ab initio* NMR shielding constants were calculated by making use of the Gauge Independent Atomic Orbital (GIAO) method,[164] both in gas phase and in chloroform, taking solvation into account by single-point computations. Moreover, instead of the recommended and widely used

TABLE 4.1: 19-OH-(−)-eburnamonine DFT-computed $\Delta G_{298.15K}$ free energies and Boltzmann-averaged expected populations.

Structure	$\Delta G_{298.15K}$, kcal/mol	Populations[a], %
SSS		
anti	3.3	0
(+)-*sync*	0.1	44
(−)-*sync*	0.0	56
SSR		
anti	6.5	0
(+)-*sync*	10.3	0
(−)-*sync*	0.0	100
SRS		
anti	4.3	0
(+)-*sync*	6.1	0
(−)-*sync*	0.0	100
SRR		
anti	4.9	0
(+)-*sync*	1.1	13
(−)-*sync*	0.0	87

[a] Configuration-constrained Boltzmann-averaged ensembles of the candidate rotameric states.

B3LYP/6–31G(d,p)[185] level, we employed the OPBE[153] non-hybrid functional in combination with the specialized basis set pcS–1.[154]

To obtain the corresponding chemical shifts for every configuration, the *ab initio* obtained shieldings were averaged on the basis of the SCF energy differences following Maxwell-Boltzmann statistics. As a first step to obtain the chemical shifts (${}^{calc}\delta^k$) of every nucleus k, shieldings of the reference (σ^{TMS}) and **4** (σ^k) molecules are computed in the same way and transformed

$$ {}^{calc}\delta^k = \frac{\sigma^{TMS} - \sigma^k}{1 - \frac{\sigma^{TMS}}{10^6}} \approx \sigma^{TMS} - \sigma^k. \tag{4.1} $$

The configuration-constrained chemical shifts (δ^k_{aver}) are calculated by

$$ {}^{calc}\delta^k_{aver} = \frac{\sum_{i=1}^{n} {}^{calc}\delta^k_i \exp(-E_i/RT)}{\sum_{i=1}^{n} \exp(-E_i/RT)}, \tag{4.2} $$

where, δ^k_i is the calculated chemical shift of the k^{th} spin in the i^{th} conformer, taking

298.25K as T. The values of E_i were obtained from the single-point *ab initio* calculations.

The computed chemical shifts for each diastereomer were introduced in the JAVA web applet provided by Goodman's group,[186] for the calculation of DP4 probability. Chemical shifts calculated *in vacuo* and in chloroform are shown in Table A.6 and Table A.7 (Appendix A), respectively. In the following, only the chemical shifts computed in chloroform will be taken into account.

Correlation coefficients as the *mean absolute error* (MAE) can be employed for comparing the experimental to the calculated chemical shifts of the candidate stereochemistries

$$MAE = \frac{1}{n}\sum_{k=1}^{n} |{}^{calc}\delta^{k}_{aver} - {}^{exp}\delta^{k}| = \frac{1}{n}\sum_{k=1}^{n} |e^{k}|, \tag{4.3}$$

even though other fit estimators can be used, such as $RMSD$

$$RMSD = \sqrt{\frac{\sum_{k=1}^{k} ({}^{exp}\delta^{k}_{i} - ({}^{calc}\delta^{k}_{aver} + \delta_{offset}))^2}{N}}. \tag{4.4}$$

The novelty of the DP4 probability is the use of empirical corrections based on a database. Note that the DP4 method introduces a least-squares linear correction of the calculated ${}^{calc}\delta^{k}_{aver}$ shifts, which compensates for systematic deviations in the methodology. Then, by assuming an empirically derived Student t distribution of the error, every error probability is calculated. Finally, the probability of obtaining such errors is calculated for every structure and converted to a *correctness* probability using Bayes' theorem.

4.3.1.1 DP4 probability points to SRS (−)-*sync* as the correct configuration and conformation of 4

DP4 calculations (Table 4.2) were done including simultaneously both ^{1}H and ^{13}C chemical shifts and each of them separately. All combinations resulted in less than 0.2% probability for other diastereomer than *SRS*.

DP4 calculations clearly indicate *SRS* as the correct relative configuration, with (−)-*sync* as the only populated rotamer, as derived from DFT energies.

TABLE 4.2: DP4 probabilities for the four possible 19-OH-(−)-eburnamonine diastereomers.

	DP4 probability, %		
Diastereomer	$^{13}C+^{1}H$	^{13}C	^{1}H
SSS	0.0	0.0	0.0
SSR	0.0	0.0	0.0
SRS	100.0	99.8	100.0
SRR	0.0	0.2	0.0

4.3.2 RDC fit to the configuration-constrained ensembles of 4

19-OH-(−)-eburnamonine (**4**) was dissolved in $CDCl_3$ and pumped inside a compressible PMMA gel using the previously described, reversible compression / relaxation method.[96] Briefly, this method allows to wash out all the methylmethacrylate monomer by alternatively compressing and relaxing the gel and changing the chloroform several times. The diffusion of **4** was done in the same way. When having a swollen and monomer-clean gel, the diffusion process takes only about five minutes and the gel sample is ready for experiment acquisition. The gel was compressed to its maximum, giving a $|\Delta\nu_Q|$ of 27 Hz, which was maintained over the whole series of experiments.

TABLE 4.3: Isotropic $^{1}J_{CH}$ and corresponding $^{1}D_{CH}$ values of **4**. Experimental errors are shown. Values in Hz.

Coupled Pair	$^{1}J_{CH}$	$^{1}D_{CH}$	Error[a]
C9–H9	162.1	20.8	0.9
C6–H6[a]	129.2	−12.9	1.4
C5–H5[a]	138.3	−14.0	1.4
C3–H3[a]	134.0	6.0	0.6
C14–H14[a]	129.0	−7.9	1.1
C21–H21	138.1	−21.6	1.9
C18–H18	126.0	−5.4	0.3
C17–H17[b]	128.8	−8.2	0.6
C15–H15[b]	126.2	16.4	0.8
C19–H19	142.7	−38.6	1.5
C12–H12	167.8	20.4	1.2
C11–H11	160.6	10.6	1.4
C10–H10	161.7	19.8	1.1

[a] Individual errors were estimated as LW/SN, where LW is the linewidth (down-scaled by κ) and SN the signal-to-noise ratio in the weaker of the two 2D spectra, which is usually the one acquired from aligned sample.[48]
[b] There is only one entry per methylene group which is the half-sum of each individual $^{1}D_{CH}$ couplings, as described in[54].

A total of 13 $^1D_{CH}$ couplings (Table 4.3) were obtained using a set of *J*-scaled (JS) F_1-coupled HSQC experiments, recorded in isotropic ($^1J_{CH}$) and anisotropic samples ($^1T_{CH} = {}^1J_{CH} + {}^1D_{CH}$).

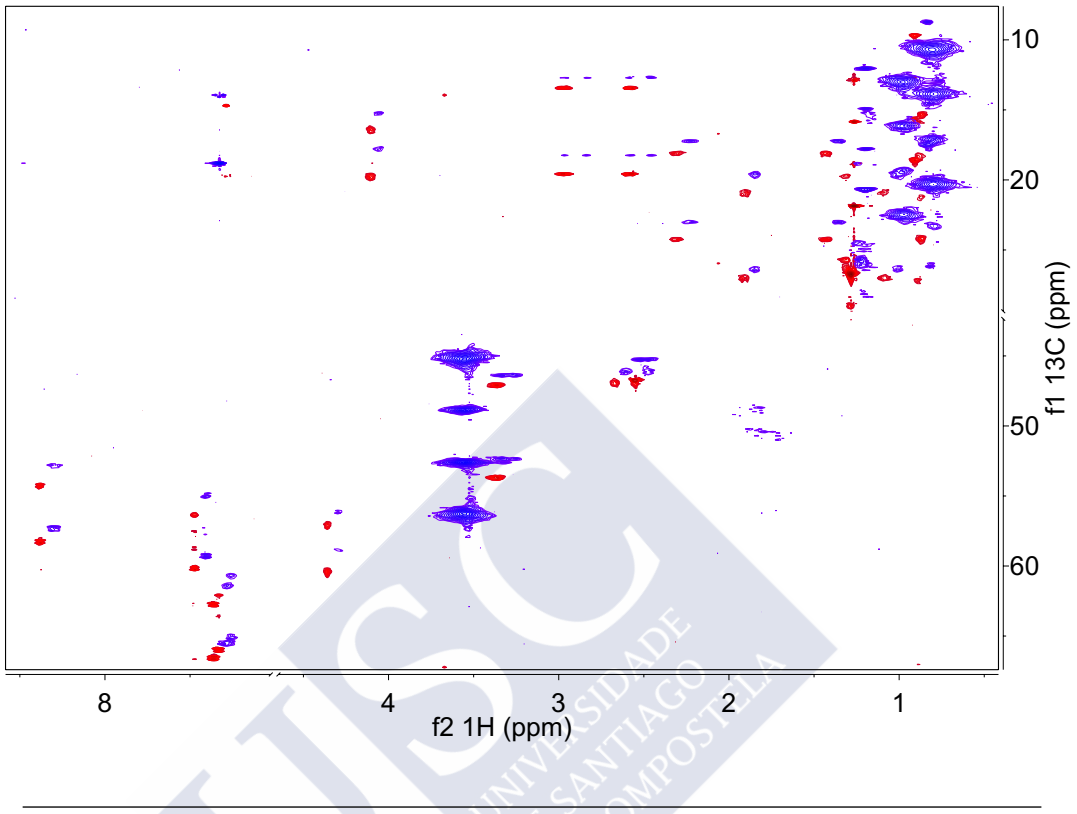


FIGURE 4.8: 19-OH-(−)-eburnamonine (**4**) F_1-coupled HSQC experiments containing a *J*-evolution multiplication module. red, isotropic; blue, anisotropic.

JS-HSQC (full spectrum in Figure 4.8) is actually the pulse sequence *hsqcetgpsp* from the Bruker library, modified for including a *J*-evolution multiplication factor (κ) for reducing the error in $^1J_{CH}$ determination without extending the experimental time.[52] The actual measured $^1J_{CH}$ evolves in the κt_1 period followed by a standard t_1 period for ^{13}C chemical shift evolution. This results in F_1 doublets of $\kappa \times {}^1J_{CH}$ splitting centered in the corresponding ^{13}C chemical shift. The use of this experimental approach permits better spectral resolution with shorter experimental times and allows better resolution for the fast-decaying FIDs of the molecules inside aligned media, in which usually methilenic protons magnetization can barely survive for $50 - 60ms$. In our hands, such an experiment produces high quality spectra, allowing reliable extraction of both $^1J_{CH}$ and $^1T_{CH}$ couplings, as can be appreciated in the spectrum expansion shown in Figure 4.9.

RDC data, obtained as $^1D_{CH} = {}^1T_{CH} - {}^1J_{CH}$ difference, was fitted to the previously described structures. The fitness between the experimental RDC data and the computed

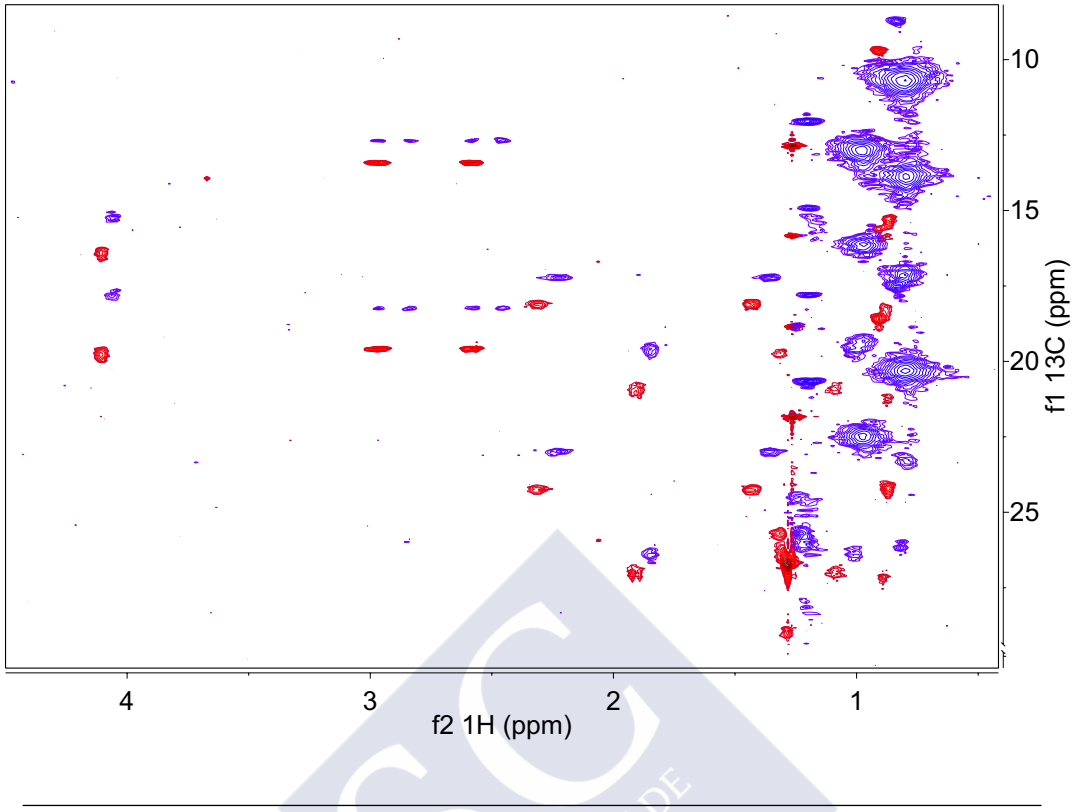


FIGURE 4.9: 19-OH-(−)-eburnamonine (**4**) F_1-coupled HSQC experiments containing a *J*-evolution multiplication module. Expansion of the aliphatic region. red, isotropic; blue, anisotropic.

trial structures was evaluated by making use of the singular value decomposition method (SVD),[41] as implemented in the MSPIN software.[42,43]

4.3.2.1 Experimental RDC fit to the single structures

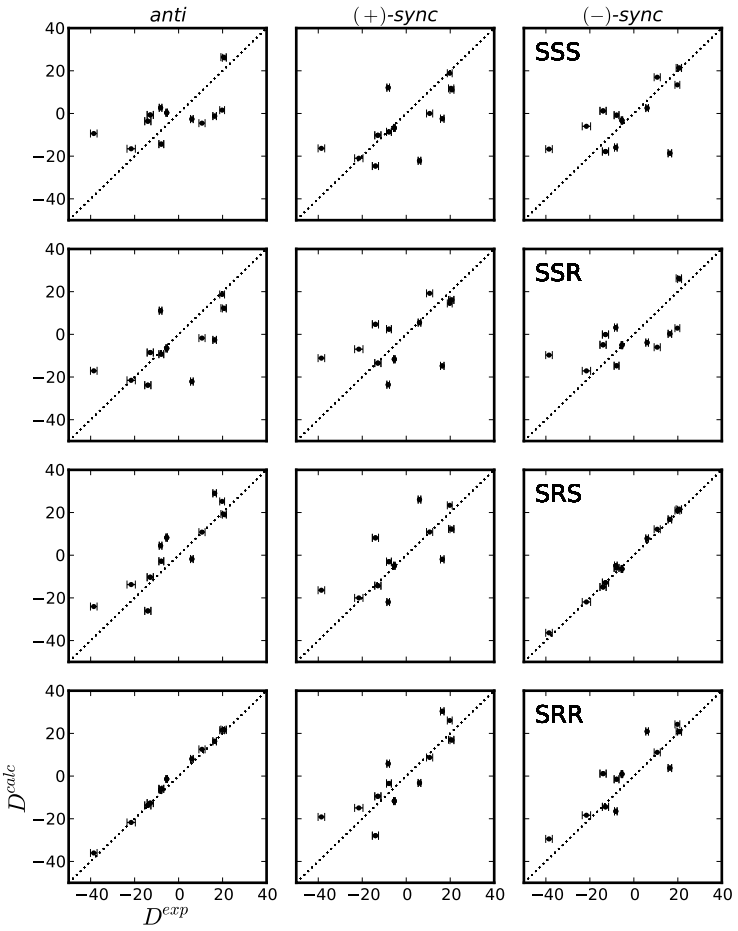


FIGURE 4.10: Plots of experimental (D^{exp}) *vs.* back-calculated (D^{calc}) $^1D_{CH}$ RDCs resulting from the fit of the sole 19-OH-(−)-eburnamonine geometries. The error bars are set to the experimental errors (Table 4.3).

As is normally done with RDC analysis in small molecules, an initial fit of all the individual structures is performed. Figure 4.10 shows the fit between the experimental (D^{exp}) and back-calculated (D^{calc}) RDCs when fitting the individual structures to experimental RDC data.

Initial SVD fit of all possible rotamers for every candidate configuration furnished more than one configuration with low Q_C, namely *SRS* and *SRR* (Figure 4.10).

For resolving the relative configuration and the conformation of **4** at the same time, we decided to employ two different fit approaches:

i) Boltzman population-constrained ensembles.

ii) Solving the populations at the same time as the alignment tensor using the single-tensor approximation.

4.3.2.2 RDC fit to Boltzmann-population constrained ensembles

The populations of the conformations were constrained to the Boltzmann values (Table 4.1) during the fit procedure. This was done taking into account the DFT energy of structures optimized taking solvent (chloroform) into account, as previously done with the DP4 probability calculation.

For fitting more that one structure at the same time (an ensemble), we made use of the single tensor approximation, in which a unique alignment tensor is computed for all structures of the ensemble. The different conformations were superimposed by minimizing the distance between the heavy atoms of the eburnan skeleton using a least-squares procedure.[51,54]

Table 4.4: Fit of RDC data to Boltzmann-averaged ensembles[a] of 19-OH-(−)-eburnamonine.

ensemble[a]	Q_C
SSS	0.868
SSR	0.739
SRS	0.085
SRR	0.509

[a] Boltzmann averaged ensembles with the DFT-derived energies shown in Table 4.1.

The quality of the fit was evaluated in terms of Cornilescu Q_C factor.[44] The best fit by far was observed for the configuration C21*S*, C20*R*, C19*S* configuration (Table 4.4), furnishing a $Q_C = 0.085$, that is lower that the other relative configurations ($\Delta Q_C > 0.4$). Thus, RDC fit clearly supports *SRS* as the correct relative configuration, with only one populated conformer, namely (−)-*sync*. This is consistent with DFT-computed energies.

Both DP4 and Boltzmann-populations constrained RDC fit supports the same result. Nevertheless both rely on the assumption of Maxwell-Boltzmann derived populations. Along this thesis, we try to avoid such a reliance, searching always for independent information which can support (or not) DFT-derived populations. In this case, we decided to repeat the RDC fit in an unconstrained fashion, which should provide the same result as the previous analysis.

4.3.2.3 RDC unconstrained fit to ensembles

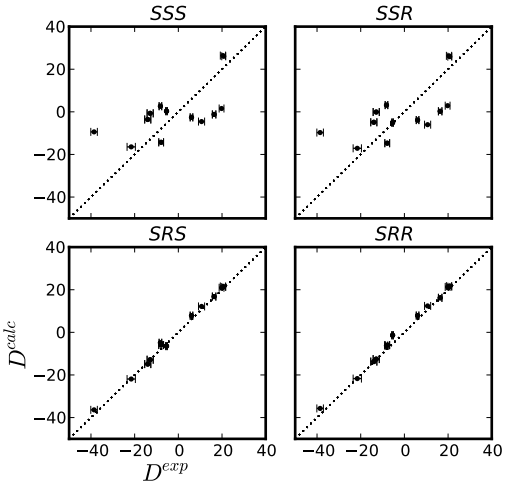


FIGURE 4.11: Plots of experimental (D^{exp}) *vs.* back-calculated (D^{calc}) $^1D_{CH}$ RDCs resulting from the fit of the sole 19-OH-(−)-eburnamonine configuration-constrained ensembles. The error bars are set to the experimental errors (Table 4.3).

To investigate the relative configuration of **4** without taking into account DFT-computed energies, we first performed a RDC fit to the sole structures (a different alignment tensor to every structure). Second, we computed the expected populations by means of RDC fits, using the single-tensor approximation, to configuration-constrained ensembles (one common tensor).

The unconstrained fit results shown in Table 4.5 are equivalent to the previous Boltzmann-constrained save for the *SRR* diastereomer. Both *SSS* and *SSR* configurations are safely discarded either in the single conformer (Figure 4.10) or in the diastereomer ensemble fit (Figure 4.11 and Table 4.5). The *SRS* diastereomer, particularlly (−)-*sync* rotamer, still furnished the lowest $Q_C = 0.085$ (Table 4.5), as in the previous Boltzmann-constrained approach (Table 4.4). Nevertheless, *SRR* configuration now furnished a lower $Q_C = 0.102$ with the *anti* rotamer being the only populated one (Table 4.5).

As the Q_C factor difference is not big enough for safely discard *SRR*, a cross-validated fit was performed. This procedure consists on repeated fits excluding the essential RDC values for differentiating these structures —namely C18–H18 and C19–H19, i.e. the vectors surrounding the new stereogenic center— one at a time, but back-calculating its expected value from the $\hat{\mathbf{A}}$ tensor calculated *without* that particular value. The fit was repeated only with the single structures *SRS*-(−)-*sync* and *SRS-anti* as they are the only well-fitting candidate structures.

TABLE 4.5: 19-OH-(−)-eburnamonine unconstrained fit to RDC data.

	Q_C factor[a]		
Structure	Single structures[b]	Ensembles[c]	Population
SSS[d]	–	0.752	
anti	0.752	0.752	100.0
(+)-*sync*	0.772	1.010	0.0
(−)-*sync*	0.764	1.010	0.0
SSR[d]	–	0.739	
anti	0.762	0.979	0.0
(+)-*sync*	0.825	1.138	0.0
(−)-*sync*	0.739	0.739	100.0
SRS[d]	–	0.085	
anti	0.505	0.661	0.0
(+)-*sync*	0.711	0.932	0.0
(−)-*sync*	0.085	0.085	100.0
SRR[d]	–	0.102	
anti	0.102	0.102	100.0
(+)-*sync*	0.546	0.713	0.0
(−)-*sync*	0.641	0.721	0.0

[a] Note that the Q_C factors for the individual conformers, listed in the *anti*, (+)-*sync* and (−)-*sync* rows correspond to the fit of these particular conformers back-calculated RDCs with the *global* alignment tensor determined for the ensemble.
[b] Fit to the single structures, every structure is fitted to a different alignment tensor, see text.
[c] Fit to the stereoisomer ensembles, containing the 3 possible conformers for each one, fitted making use of the single-tensor approximation, see text.
[d] Global ensemble Q_C factor resulting from the ensemble singe-tensor fit.

The results of the cross-validation of both C18–H18 and C19–H19 $^1D_{CH}$ are shown in Table 4.6 and Figure 4.12.

Both Q_C factors got reduced from the elimination of the two $^1D_{CH}$, as expected when removing structural restraints from the fit. The cross-validated value of C19–H19 $^1D_{CH}$ was very similar for both structures, as the spatial orientation of the C19–H19 bond is equivalent in both configurations. Interestingly, C18–H18 cross-validation furnished a completely different value that points clearly to *SRS*-(−)-*sync* as the correct structure.

TABLE 4.6: Cross-validated fit results. Only the two lowest Q_C structures are shown, see text.

Structure	Q_C factor	HC18–H18	C19–H19
SRS (−)-*sync*[a]	0.079	−6.28	−33.93
SRR anti[a]	0.068	−1.32	−33.09
Experimental		−5.40	−38.60

[a] Cross-validation: C*A*–H*A* RDC coupling is back-calculated from the alignment tensor calculated excluding that particular (${}^{exp}D_{CA\text{–}HA}$) value.

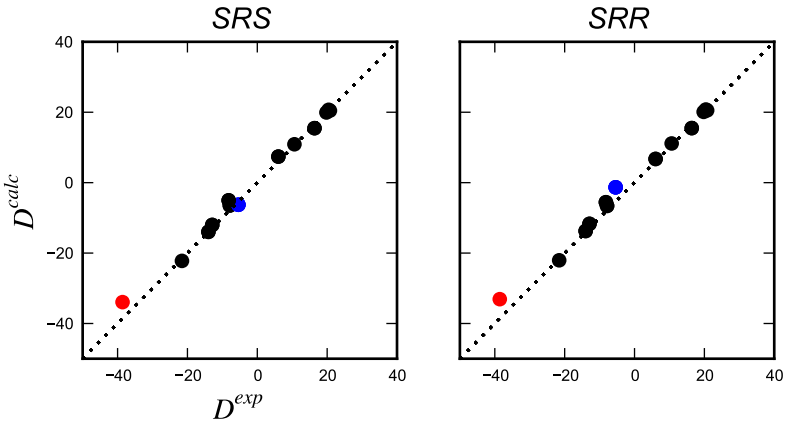


FIGURE 4.12: Plots of experimental (D^{exp}) *vs.* back-calculated (D^{calc}) ${}^1D_{CH}$ RDCs resulting from the cross-validation of 19-OH-(−)-eburnamonine candidate configurations. Back-calculated RDC for C19–H19 coupling in red and C18–H18 in blue.

Both *ab initio* calculations of chemical shifts, as well as RDC analysis pointed to the relative configuration *SRS* as the correct one. Furthermore, both methodologies permitted the determination of the (−)-*sync* rotameric state around C19–C20 bond as the only populated one.

These results are supported by DFT computations. DFT derived energies indicated a clear preference for the (−)-*synclinal* conformation in the *SRS* configuration, as this particular rotamer benefits from the possibility of an hydrogen bonding between the amine nitrogen and the hydroxyl group of the (C19) side chain (Figure 4.13).

Additional experimental information supports these findings. In this single conformation, the atoms H19–C20–C15–H15α(ax) adopt a *W* arrangement leading to a H19–H5α(ax) ${}^4J_{HH}$ of 1.4 Hz. In addition, the hydrogen bond between the amine N atom and the OH group is supported by the very broad singlet a 5.94 ppm, in full agreement with its DFT calculated value of 5.91 ppm.

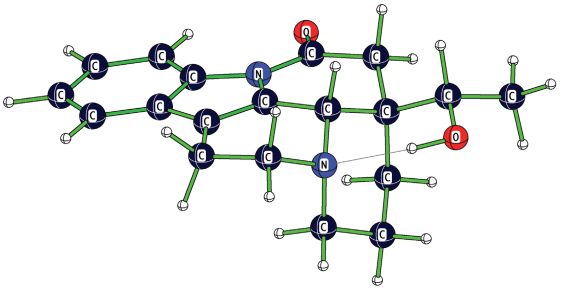


FIGURE 4.13: Relative configuration and conformation of **4** determined by RDC fits and DFT chemical shifts computations.

4.4 Absolute Configuration Determination

Both enantiomers of **3**, (+)-eburnamonine [CAS RN 47ebn–00–0] and (−)-eburnamonine [CAS RN 4880–88–0], have been isolated from natural source making necessary the determination of the absolute configuration of **4**.

The use of chiroptical methods for establishing the absolute configuration of one compound once its relative configuration and the conformational space has been determined by another technique (such as NMR) has been previously recognized.[176] Recently, this approach was combined with the efficient relative configuration determination by RDC-enhanced NMR by Griesinger and co-workers,[106,187] and Navarro-Vázquez and co-workers.[188]

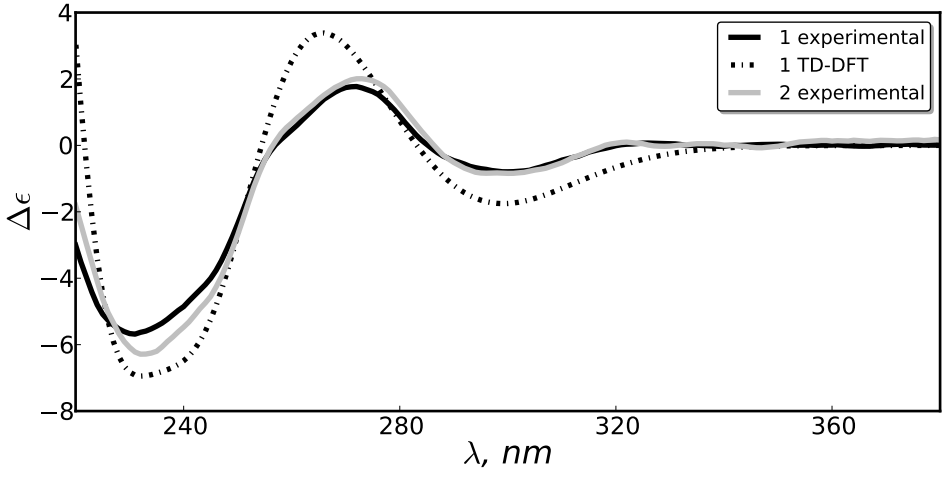


FIGURE 4.14: Circular dichroism spectra of **4** (solid line), **3** (shaded line) and TD-DFT computed of **4** (dash-dotted line).

The absolute configuration of **4** was determined using this procedure. The CD spectra of **4** and a commercial sample of (−)-eburnamonine (**3**) collected in acetonitrile were

nearly superimposable, suggesting that both compounds share the same absolute configuration of the eburnan-type skeleton (see Figure 4.14). In fact, PBE0/6-311+G(d,p) TD-DFT calculated CD spectra for *SRS* closely matches the experimental one, as shown in Figure 4.14. The experimental specific rotation at 589 nm ([α]D) of **4** ($-60.6°$) was in excellent agreement with the value calculated ($-60.5°$) using the same theory level as for the CD spectra.

4.5 Conclusion

A new eburnamonine derivative was isolated from its natural source and identified on the basis of 1D and 2D NMR experiments (Figure 4.3 and Figure 4.5) and high-resolution mass spectroscopy (Figure 4.4) as a hydroxylated derivative 19-OH-eburnamonine.

The candidate structures of the new hydroxylated derivative were generated by means of molecular mechanics calculations, and further optimized at DFT level (Figure 4.6 and Figure 4.7).

The relative configuration of 19-OH-eburnamonine was determined by two independent methods, namely DFT-computed chemical shifts comparison to experimental ones by the DP4 probability (Table 4.2) and unconstrained RDC fit (Table 4.5), treating every configuration as an ensemble of the three possible conformations. Both independent methods selected *SRS* as the correct relative configuration. Furthermore, RDC fits selected the rotameric state ($-$)-*sync* as the only populated one. A cross-validation of this result with the two lowest-Q_C structures (Table 4.6 and Figure 4.12) confirmed the results.

Once the relative configuration was known, the absolute configuration of 19-OH-($-$)-eburnamonine was determined from the comparison of the TD-DFT-computed ECD spectra with the experimental one (Figure 4.14). The absolute configuration was found to be *SRS*. Importantly both CD spectra were superimposable with the one of commercial ($-$)-eburnamonine.

4.6 Materials and Methods

4.6.1 Plant material and Extraction

Plant material collection, extraction and isolation of **4** were done by Roxana Sifuentes and Dr. Helena Maruenda, from Departamento de Ciencias — Sección Química, Pontificia Universidad del Perú.

Plant Material

Bonafousia macrocalyx Muell. Arg. was collected in Nauta — Rio Maranón, Long: 73°, 35′ W; Lat 04° 48′ S, Loreto, Perú, in November 2001 at an altitude of 150 meters.[1] The plant was identified by Prof. Juan Ruiz Caledonio (NC 5520) and the voucher specimen (34368) is deposited at the Herbarium AMAZ, Universidad Nacional de la Amazonia Peruana, Iquitos, Perú.

Extraction and Isolation

Dried ground leaves (120 g) were extracted at 25° C, first with petrol ether (3 × 1.2 L, 24h) and then with methanol (3 × 1.2 L, 24h). The methanol extract was evaporated to dryness to yield 14 g (116 g/Kg) of viscous material. The alkaloid fraction was obtained by partitioning the extract between 2% aqueous tartaric acid solution (30 mL) and ethyl acetate (60 mL). The aqueous layer was neutralized with $NaHCO_3$ and extracted with ethyl acetate (3 x 50 mL) to yield 200 mg (1.7 g/Kg) of an oily residue. Flash column chromatography of the later fraction, using CH_2Cl_2 / CH_3OH / Et_3N : 9.65/0.25/0.05 as the eluting solvent (R_f 0.3), yielded 45 mg of **4** as a white solid (0.4 g/Kg).

4.6.2 Computational Methodologies

Generation of all possible configurations, conformational search, and DFT calculations were carried out by Dr. Armando Navarro-Vázquez.

On each of the four possible configurations, geometries for the different rotamers around C19–C20 were generated by means of molecular mechanics MMFF94[123] computations. All these structures were refined at OPBE/6–31G(d)[153] level of theory

[1] The reader can view the location in Google Maps

in gas phase and vibrational frequencies computed to check that all obtained stationary points were true minima. Solvation was taken into account by single point OPBE/pcS-1[153,154] computations using the polarizable continuum model (PCM)[124] with chloroform gaussian09 parameters. DFT derived energies are shown in Appendices (Appendix B, Section B.3.2).

NMR shielding tensors were obtained at this same level. The computed shieldings were Boltzmann averaged (298.15 K) using the computed relative energies. In order to transform computed shieldings into chemical shifts the reference shielding was obtained by minimizing the difference between observed and computed data in a least-square sense. CD Rotational Strengths were computed at the TD-DFT level using the PBE0 functional (PBE1PBE keyword in gaussian09)[189] and the 6–311+G(d,p) basis set on OPBE/6-31G(d) previously optimized geometries. Solvation was included in the TD-DFT computations at the PCM level using acetonitrile parameters. CD spectra were simulated by convoluting rotational strengths intensities with Gaussian bands of 0.5 eV half-height full-width. All the DFT computations were done in Gausssian09.[131]

4.6.3 General

Melting point was measured on a Gallemkamp apparatus and is uncorrected. UV spectra were obtained in a Perkin Elmer UV/VIS Lambda 2-UVWinlab spectrometer, whereas the IR spectra on a Perkin Elmer 1600 Series FTIR âĂŞ Model 1620. CD spectra were collected in a JASCO J-815 spectrometer and optical specific rotation in a JASCO P-1020 polarimeter. Accurate mass spectra were recorded on a Thermo LTQ Orbitrap XL (Thermo-Fisher) HRMS. Purification of the compounds was monitored by thin-layer chromatography on aluminum-backed silica gel 60 F254 plates. Silica gel 60, particle size 40-63 μm (mesh 230-400) was used for flash chromatography. The solvents were HPLC and LC-MS reagent quality. Commercial (−)-eburnamonine **3** [CAS RN 4880–88–0] was purchased from Sigma-Aldrich.

Circular dichroism spectrum of a $3.9 \cdot 10^{-5}$ M solution of **4** in acetonitrile was recorded within a $190 - 400$ nm spectral window and 16 scans were accumulated.

4.6.4 NMR Spectroscopy

All NMR experiments were carried out in a Bruker Avance III 500 MHz spectrometer, operating at 500.13 MHz for ^{1}H and 125.76 MHz for ^{13}C. The assignment spectra can be seen in Appendix A, Section A.3.1 and the assignment in Appendix A, Section A.3.2.

One-bond ^{1}H–^{13}C Residual Dipolar Couplings (RDCs) were extracted as the difference in signal splitting between isotropic ($^{1}J_{CH}$) and anisotropic ($^{1}T_{CH} = {}^{1}J_{CH} + {}^{1}D_{CH}$) F_1-coupled JS-HSQC.[52,182] Both JS-HSQC experiments were acquired as a 399* × 512* real point matrix using a 4000 × 7545 Hz spectral window, 32 increments per FID, and a total (F_1) acquisition time of 67.8 milliseconds. Both experiments included a *J*-scaling factor (κ =3) for reducing the error in coupling measurement without extending acquisition time.

4.6.5 RDC fit

RDC analysis was performed using the MSPIN program[42,43] with the procedure outlined in Chapter 1.

Input and output data are reported in Appendices. Molecular coordinates of conformers **1A-C** (Appendix B, Section B.3.1), RDC input tables in MSPIN-ready format (Appendix B, Section B.3.3), and program outputs (Appendix B, Section B.3.4).

Chapter 5

Application of long-range RDCs to determination of the configuration

5.1 Long-range Residual Dipolar Couplings

Up to date for most published studies, all the application of RDCs to configurational, constitutional or conformational analysis of small molecules in solution made use of the one-bond ^{13}C–^{1}H dipolar couplings. The use of long-range $^{n}D_{CH}$ couplings to solve such structural problems is very infrequent.

Short-range ($^{1}D_{CH}$) couplings are relatively easy to measure in small molecules at natural ^{13}C abundance. In most of the cases, $^{1}D_{CH}$ couplings can be extracted from the direct (F_2) or indirect (F_1) dimensions through ^{1}H-coupled HSQC experiments. It is even possible to measure accurate $^{1}D_{CH}$ couplings with the very simple mono-dimensional ^{13}C gated-decoupled experiment (if enough signal dispersion exists).

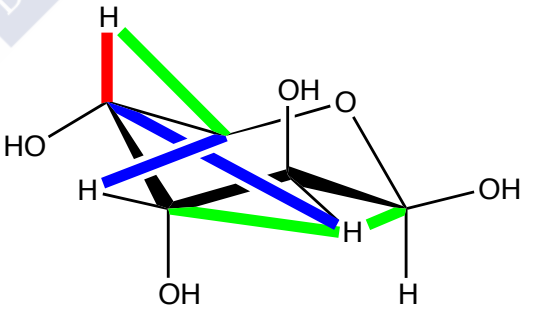


FIGURE 5.1: Schematic representation of $^{1}D_{CH}$, $^{2}D_{CH}$, and $^{3}D_{CH}$ vectors in a model pentose.

In general, $^{1}J_{CH}$ coupling constants are large and well resolved, if enough signal dispersion exists. When dissolved in alignment media, the $^{1}D_{CH}$ contribution is easily extracted as the splitting difference. All published studies making use of $^{1}D_{CH}$ have shown that short-range couplings are, in most cases, sufficient to solve a wide variety of structural problems in small molecules.

The scarce use of long-range couplings is mostly due to the intrinsic problems accompanying their measurement. Mainly those problems arose from the fact that the couplings are rather small ($1 - 10$ Hz, same magnitude as $^{3}J_{HH}$ couplings) and are correlated with low sensitivity nucleus, such as ^{13}C.

5.1.1 Long-range RDCs are valuable structural restraints

The key structural value of long-range ($^{n}J_{CH}$) couplings is well known and has been investigated since the 1970s. One of the first applied methods was developed by Matsumori *et al.* by the introduction of the "*J*-based configuration analysis".[190] This was extensively used[191,192] and still is being used for solving the relative configuration of complex natural products.[193–197]

This "*J*-based configuration analysis" relies on a combination of $^{2}J_{CH}$ and $^{3}J_{CH}$ along with $^{3}J_{HH}$ coupling constants to determine the relative stereochemistry between any two stereogenic centers, as long as any carbon between them contains at least one proton (methine or resolvable diasterotopic methylene protons). Due to the well known angle-magnitude dependence, described by *Karplus*-type relationships, the extracted value can be converted to a dihedral angle between the two atoms. This methodology has been successfully applied in many molecules, but it is unhelpful for solving the relative configuration of (*magnetically*) disconnected stereogenic centers. Magnetically inactive linkers, such as heteroatoms or carbons not attached to a proton atom, interrupt the local-sequential information that both $^{3}J_{CH}$ and $^{3}J_{HH}$ contain.

Despite the technical and methodological challenges associated with long-range couplings, there are many situations in which additional structural restraints are needed:

i) molecules with disconnected stereogenic centers containing several C–H bonds nearly parallel, or moieties not having enough protons

ii) (semi-)rigid molecules in which many candidate configurations and/or conformations can not be distinguished —in terms of fit quality factor— based only on $^{1}D_{CH}$ data fit.

This problems are quite common in oxidized natural products, among a wide range of natural and synthetic molecules. Accurate long-range couplings can provide more structural restraints in this challenging situations.

5.1.2 Available experiments for long-range couplings measurement

There has been so much development of experiments capable of accurately measuring long-range $^{n}J_{CH}$ couplings.

Contrary to $^{3}J_{HH}$ coupling measurement in isotropic samples, both *magnitude* and *sign* determination of the long-range couplings are of key importance in weakly aligned samples.

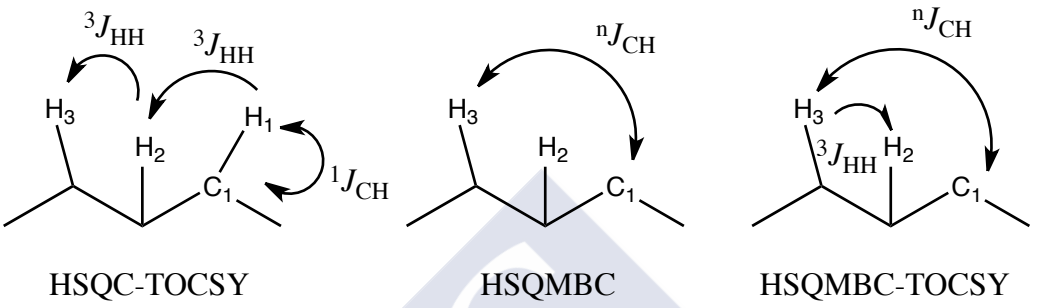


FIGURE 5.2: Magnetization transfer schemes of different types of Long-range $^{n}J_{CH}$ measurement experiments.

Available experiments for the measurement of long-range couplings have been recently reviewed by Kobzar and Luy[79] and, not so recently but extensively, by Márquez *et al.*[198] Briefly, experiments that allow the measurement of long-range couplings can be grouped depending on the design approach for obtaining the couplings:

i) **HMBC**: *J*-resolved HMBC, *J*-IMPEACH-MBC, psHMBC.

- HMBC coherence transfer(s) (Figure 5.2) allows determination of $^{n}J_{CH}$ involving quaternary carbons.

ii) **HSQC-TOCSY**: HETLOC, HSQC-HECADE, HSQC-TOCSY, HSQC-TOCSY-IPAP.

- TOCSY transfer steps (Figure 5.2) permit the evolution of $^{n}J_{CH}$ to continuous ^{1}H-containing spin-systems.[1]

iii) **HSQMBC**: GSQMBC, HSQMBC, G-BIRD-HSQMBC, P.E.HSQMBC.

iv) **CPMG**-based: LR-CAHSQC, CAGEBIRD-CPMG-HSQMBC.

[1] TOCSY transfer step causes a loss of about 75% of signal, when measuring in aligned media (L. Castañar, personal communication).

Even when there is a respectable number of experiments that can provide long-range *J* couplings, their use to extract RDCS is, in practice, very limited due to the inherent difficulties of aligned spectra such as linebroadening.

For very small molecules, such as menthol, long range ^{13}C–^{1}H RDCs were measured with a modified HMBC experiment proposed by Griesinger and co-workers[78] based on previous work by Keeler and Neuhaus.[199] Thiele *et al.* have used HETLOC experiments to measure the sign and magnitude of long-range CH RDCs of an α-methylene-γ-butyrolactone where a very limited amount of $^{1}D_{CH}$ couplings were available in order to solve the structure.[69]

5.1.3 Selective *J*-Scaled HSQC (SJS-HSQC)

Experiment Description

In this thesis disertation we will show the benefits of a new selective *J*-scaled HSQC experiment developed in Ad Bax group. The new pulse sequence, shown in Figure 5.3, was designed for incorporating the conceptual simplicity of the selective ^{1}H-flip experiment[200] into the ^{1}H-detected HSQC experiment. Additionally, the experiment is enhanced by gradient selection of coherence pathways resulting in sensitivity enhancement and suppression of artifacts.[201] Further suppression of artifacts and a minor sensitivity enhancement, resulting from positive ^{1}H–^{1}H NOE interactions, is accomplished by generating a mostly saturated state for ^{12}C-attached protons by inverting them through a bilinear rotation and subsequent recovery delay Δ,[202] prior to the start of the actual selective *J*-scaled (SJS) HSQC experiment.

The actual gradient-enhanced SJS-HSQC experiment starts at the end of the Δ period with an INEPT transfer of ^{1}H magnetization to ^{13}C, and uses τ and τ' optimized delays for simultaneous detection of methine, methylene, and methyl signals.[203] Prior to the evolution period, the experiment contains a selective *J*-dephasing period, of duration κt_1, which enhances by a factor κ the F_1 dimension $^{n}J_{CH}$ splitting to the proton selectively inverted by the 180° ϕ_2 pulse.[52]

Importantly, effects of static field inhomogeneity, which can be problematic in aligned samples, are refocused at the end of the κt_1 period and do not hamper resolution of the doublet splitting. For technical reasons, the experiment is coded with two 180° pulses during the time where ^{13}C magnetization evolves. This makes possible to use adiabatic pulses for refocusing purposes, as these compensate each other's phase imperfections.[204]

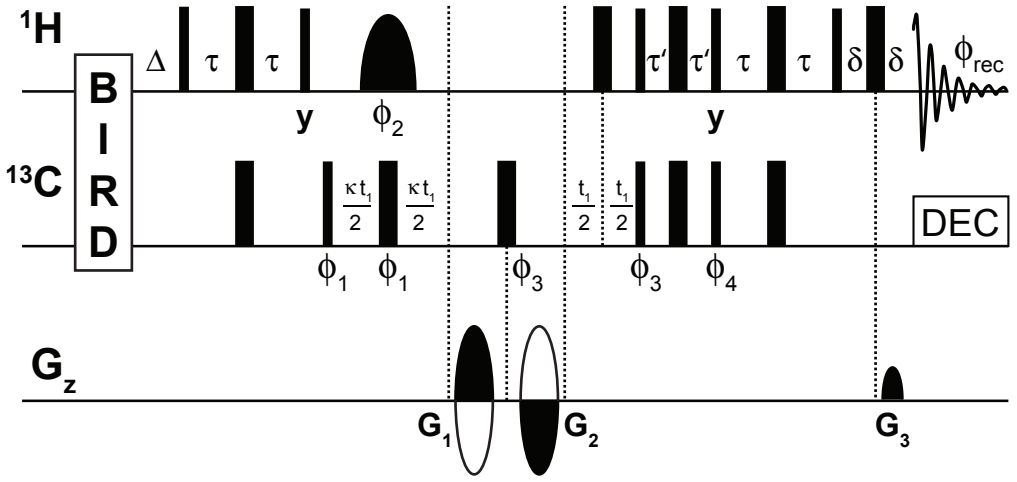


FIGURE 5.3: Selective *J*-Scaled HSQC experiment, see text. $\tau = 1/(4J)$, INEPT delay adjusted usually to 145 Hz; κ, J evolution multiplication factor; BIRD^{d,X} delay adjusted to $1/(2J)$; $\delta = 1/(8J)$, multiplicity selection. Phase cycling: $\phi_1 = x, -x$; $\phi_2 = x, x, x, x, -x, -x, -x, -x$; $\phi_3 = x, x, -x, -x$; $\phi_4 = y, y, -y, -y$; $\phi_{rec} = x, -x, -x, x$. Gradients power, as % of maximum power, $G_1 = 80$, $G_2 = -80$, $G_3 = 80$.

Use of the gradient-enhanced pulse scheme to transfer magnetization from ^{13}C back to ^{1}H, to first order, leaves unchanged the spin state of protons not attached to a ^{13}C atom (neglecting pulse imperfections and ^{1}H–^{1}H dephasing and relaxation during the short $2\tau + 2\tau' + 2\delta$). Therefore, when a proton $^{1}H_A$ is selected by the 180° ϕ_2 ^{1}H pulse between ^{13}C evolution and ^{1}H detection, this results in an E.COSY type multiplet.[205]

Coupling Extraction

In the same fashion as in the original selective ^{1}H-flip experiment,[200] a single proton is inverted per experiment, originating plain doublets for every ^{13}C atom having a coupling to this proton. It is evident that if a methylene proton is selectively inverted, this will result in a triplet splitting of the coupled resonances, and in quartets if a methyl group is selected.

This method results in simple and clean spectra. Experiment sensitivity is comparable with that of the optimized gradient-enhanced HSQC experiment, meaning that for a doublet splitting —resulting from a methine inversion— the sensitivity of SJS-HSQC experiment is decreased only by about two-fold relative to the reference ge-HSQC.

The use of a large κ scaling factor —$\kappa = 20$, in all the experiments shown in this thesis—, furnishes high resolution of the F_1 doublets while using a limited number of t_1 increments, consequently shortening the experimental time and permitting the rapid exploration of all resonances of interest.

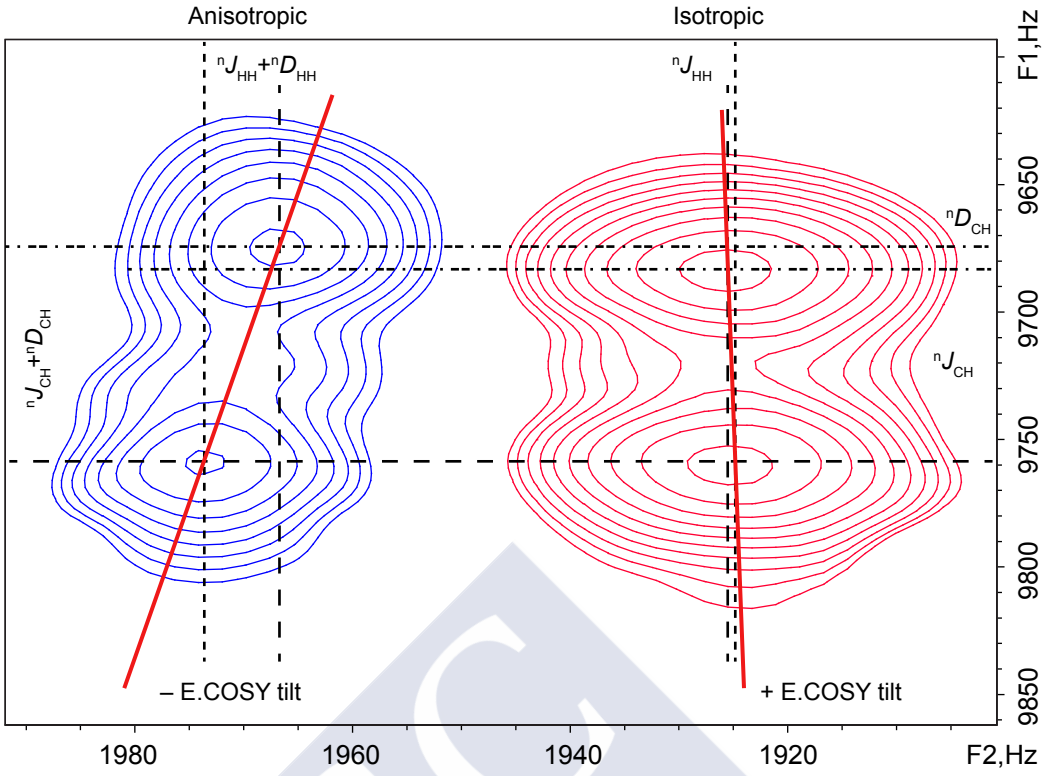


FIGURE 5.4: C7 resonance of 10-Epi (**5**, see text) splitting resulting from the selective inversion of H5. Superimposed SJS-HSQC spectra of isotropic (red) and anisotropic (blue) samples are shown. The E.COSY patterns shown indicate that $^{n-1}J_{HH}$ (with $n = 3$, $^3J_{CH}$) $^nT_{CH}$ has the same sign as the corresponding ^{1}H–^{1}H coupling (negative E.COSY tilt), whereas the isotropic $^nJ_{CH}$ has the opposite sign as the $^{n-1}J_{HH}$ coupling.

In the figure above (Figure 5.4), the resulting SJS-HSQC from the selective inversion of *H5* is shown (see next section). In the general case, the ^{13}C-dimension splitting of the *B* resonance for a given ^1H_B–$^{13}C_B$ peak corresponds to the (long-range) $J_{HA–CB}$ coupling and the F_2 dimension displacement of the α and β doublet components corresponds to the homonuclear coupling, $J_{HA–HB}$.

The E.COSY encoded pattern contains information about the relative sign of the long-range ($J_{HA–CB}$) heteronuclear coupling with respect to the corresponding ($J_{HA–HB}$) homonuclear coupling. As shown in Figure 5.4, the E.COSY pattern can be easily recognized from the relative displacement direction of the two doublet components.

5.2 10-Epi-8-deoxicumambrin B is an aromatase inhibitor targeted to breast cancer treatment

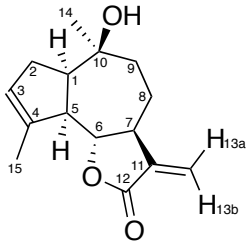


FIGURE 5.5: 10-Epi-8-deoxycumambrin B (**5**) structure showing the known absolute configuration.

10-epi-8-deoxycumambrin B (**5**, Figure 5.5), which will be abbreviated along the text as 10-Epi, is a biologically relevant natural sesquiterpene lactone isolated from *Stevia yaconensis* var. *subeglandulosa*.[206] This compound has shown significant activity against aromatase, an enzyme involved in hormone-dependent postmenopausal breast cancer.[207] The ^{1}H and the ^{13}C NMR spectra of **5** were assigned by Gil and co-workers.[206,208]

5.2.1 Biosynthetic restrictions leave 16 candidate 10-epi configurations

The absolute configuration of 10-Epi is well known. Additionally, the natural absolute configuration at C7 for sesquiterpene lactones isolated from species of the *Compositae* family, such as 10-Epi, is *S* due to the bio-synthetic pathway (H7 in α-orientation when the structure is drawn as shown in Figure 5.5).[209] However, the configuration at the stereocenters C1, C5, C6 and C10 can be either *S* or *R*, giving origin to a total of 16 diastereomers in which the configuration at C7 is maintained as *S*.

All possible configurations of 10-Epi were generated in Maestro software by manually altering the corresponding bonds, followed by an energy minimization step *in vacuo*.[150,151] The available conformational space for every generated conformation was examined by means of molecular mechanics MM3 force field calculations in vacuum.[150,151] Sampling of the available potential surface was done with the mixed torsional/low-mode sampling algorithm,[130] within an energy gap of 8 kcal/mol, as implemented in MacroModel.[110] The molecular mechanics optimized structures were further optimized at DFT level of theory using the OPBE[165,166] functional combined with the 6–31G(d) basis set *in vacuo*. Analytical frequencies were computed to ensure the nature of the stationary point.

Each structure was named after its configuration at carbons C1, C5, C6, C7 and C10. Meaning that compound **5** was named *RRSSS*.

DFT derived energies were analyzed in order to determine the expected conformer populations —for every isolated configuration— following Maxwell-Boltzmann statistics, assuming $T = 298.15\ K$

$$\text{Population} = \frac{\exp(-\Delta G_{298.15}/RT)}{\sum_{i=1}^{n} \exp(-\Delta G_{298.15}/RT)}. \tag{5.1}$$

Most configurations showed a single dominant conformation. Only four out of sixteen configurations were expected to present more than one conformer in solution based on relative energies difference. The configurations *RSSSR*, *RSSSS* and *SRRSS* were predicted to be in a two-state exchange with a nearly 1 : 1 ratio of the two low-energy conformers —with relative energy difference $\Delta G_{298.15} \leq 0.2\ kcal/mol$—. The remaining *SRSSR* configuration was the only one presenting more than two possible conformers, according to DFT energies —with an approximate population ratio of 6.0 : 2.5 : 1.5—.

TABLE 5.1: 10-epi-8-deoxycumambrin B diastereomers with more than one populated conformer energies and Maxwell-Boltzmann computed populations.

Diastereomer	Conformer	$\Delta G_{298.15}$, kcal/mol	p_i, %[a]
RSSSR	1	0.0	58.4
	2	0.2	41.6
RSSSS	1	0.0	59.4
	2	0.2	40.6
SRRSS	1	0.0	61.7
	2	0.3	38.3
SRSSR	1	0.0	60.5
	2	0.5	24.5
	3	0.8	15.1

[a] Conformers with expected populations lower than 10% have been excluded from the final calculation shown in this table.

All diastereomers presenting more than one populated conformer will be constrained to Boltzmann-averaged ensembles in the RDC fits. As in the previously discussed 19-(OH)-(−)-Eburnamonine case (see Chapter 4), the single-tensor approximation will be used. The atomic coordinates were conveniently superimposed for maximizing the decoupling between internal conformational movements and external molecular tumbling.

5.3 RDC-based analysis of the configuration of 10-epi

5.3.1 RDC extraction from F_1 HSQC-based experiments

The sample of 10-epi-8-deoxycumambrin B (**5**) used for these experiments was extracted and reisolated from *Stevia yaconensis* var. *subeglandulosa* by Dr. Viviana E. Nicotra from Universidad Nacional de Córdoba, Córdoba, Argentina.

A 25 mm length PMMA gel was swollen in $CDCl_3$. Residual monomer was washed out following the protocol developed by Gil and co-workers. Briefly, with this method it is possible to wash out all the methylmethacrylate monomer by alternatively compressing and relaxing the gel, changing the chloroform several times. Once the gel was clean of monomer, 3 mg of compound **5** were dissolved in $CDCl_3$ (200 μL) and added on top of the PMMA gel. The "forced dialysis" property of these gels is also employed for accelerating the diffusion of the sample.[45,96] When the gel was equilibrated with the compound, it was compressed to the maximum, and the tube was sealed with Teflon tape for holding the position of the plunger and avoiding solvent evaporation. The gel recovered the latter maximum $|\Delta\nu_Q|$ of 27 Hz and maintained it during the acquisition of all the spectra. An equivalent sample of 3 mg of **5** in 500 μL of neat $CDCl_3$ was prepared to collect the NMR data in isotropic conditions.

5.3.1.1 Measurement of $^1D_{CH}$

The one-bond $^1D_{CH}$ couplings shown in Table 5.2 were measured as the $^nD_{CH} = {}^nT_{CH} - {}^nJ_{CH}$ difference from C1–H1, C3–H3, C5–H5, C6–H6, and C7–H7 using (proton) F_1-coupled *J*-scaled HSQC experiments (JS-HSQC).[52] The *J* amplification factor (κ) was set to 3 (Figure 5.6). The two one-bond $^1D_{CH}$ and the geminal $^2D_{HH}$ couplings of vinylidene CH13 were measured with the *J*-modulated HMQC-type experiment (*J*-HMQC-ge/se-HSQC) proposed by Kövér and co-workers (Figure 5.7).[210] The measurement of accurate $^1D_{CH}$ and $^2D_{HH}$ couplings for the methylenic protons corresponding to C2, C8, and C9 aliphatic carbons, was prevented by extensive long-range $^nD_{HH}$ couplings to other (neighboring) protons. As an alternative, based on our previous work, the individual $^1D_{CH}$ of each of the methylenes was extracted from the JS-HSQC experiment. Likewise, couplings of the C14 and C15 methyl groups were measured from the same experiment.[54]

TABLE 5.2: 10-epi-8-deoxycumambrin B one bond J couplings and their corresponding RDCs and experimental errors, in Hz.

Coupled Pair	n[a]	$^{n}J_{CH}$	$^{n}D_{CH}$	Error
C1–H1	1	127.4	−7.5	1.0
C2–H2[b]	1	129.3	−7.6	0.8
C3–H3	1	159.3	12.0	0.6
C5–H5	1	129.5	−26.3	1.1
C6–H6	1	153.9	−32.8	0.8
C7–H7	1	127.6	−26.9	1.5
C8–H8[b]	1	127.3	−12.9	1.2
C9–H9[b]	1	124.2	−26.2	2.1
C14–H14	1	125.7	−5.5	0.3
C15–H15	1	125.8	1.1	0.4
C13–H13a[c]	1	163.1	23.4	0.3
C13–H13b[c]	1	160.2	4.4	0.3
H13a–H13b[c, d]	2	$\leq 2; \geq -2$	11.00	2.00

[a] All couplings are expressed as $^{n}J_{A–B}$, being A–B the coupled atoms (^{13}C and ^{1}H, respectively) and n the number of bonds in between.
[b] There is only one entry per methylene group which is the half-sum of each individual $^{1}D_{CH}$ coupling.
[c] Values obtained from the J-HMQC-ge/se-HSQC experiment, see text.
[d] $^{2}D_{HH}$ coupling.

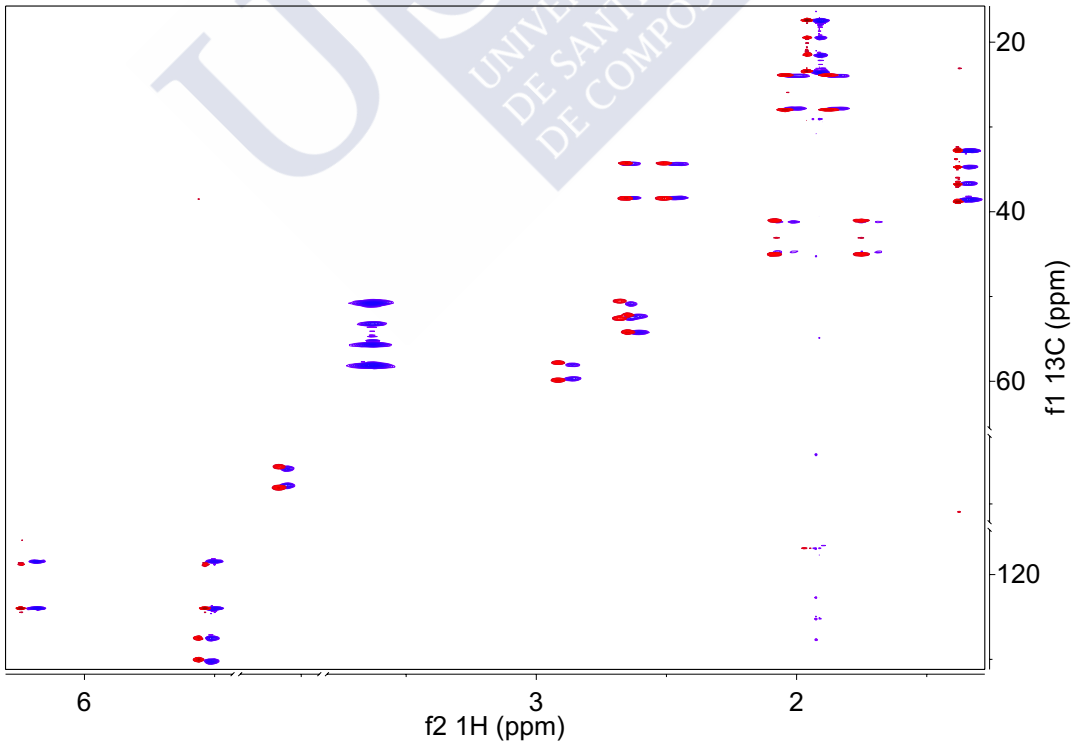


FIGURE 5.6: 10-epi (**5**) F_1-coupled HSQC experiments containing a J-evolution multiplication module. Red, isotropic; blue, anisotropic.

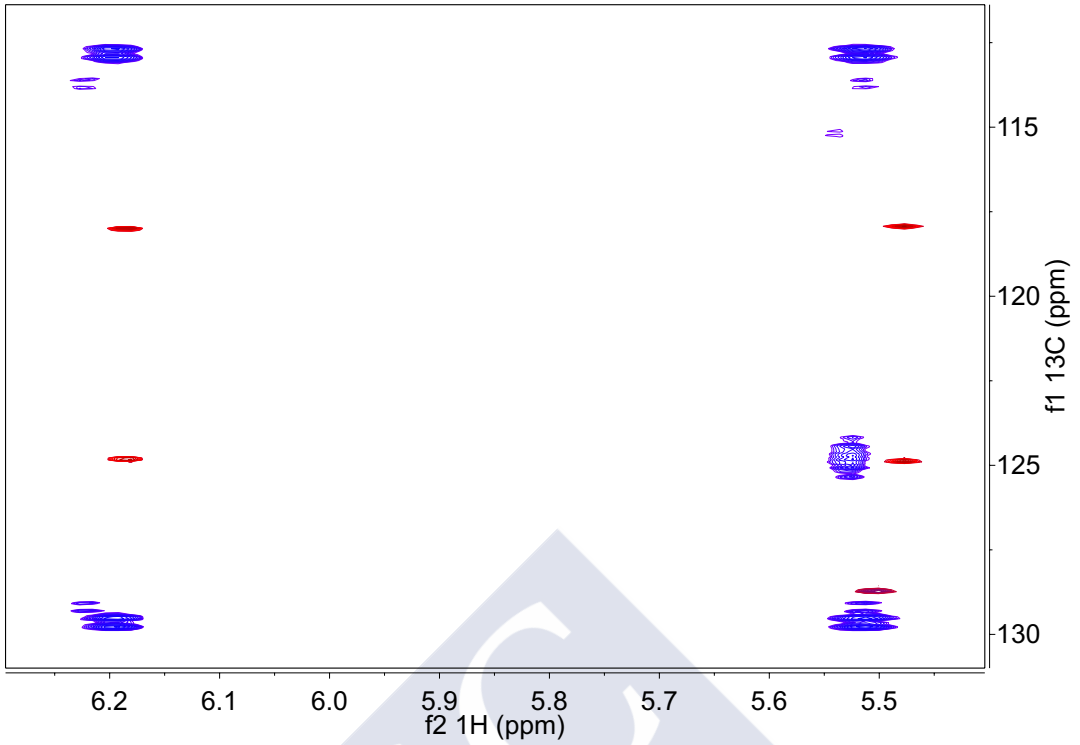


FIGURE 5.7: 10-epi (**5**) *J*-HMQC-ge/se-HSQC experiments containing a *J*-evolution multiplication module. Expansion of the C13–13a/b region. Notice the severe dipole-dipole interactions affecting H3 resonance at 129 / 5.5 ppm (^{13}C, ^{1}H; respectively). Red, isotropic; blue, anisotropic.

5.3.1.2 Measurement of 10-epi long-range ($^{n}D_{CH}$) couplings

A set of four SJS-HSQC experiments were recorded both in isotropic and anisotropic conditions, each inverting selectively one ^{1}H resonance. For selective inversion, well resolved proton resonances were chosen, namely H2α, H5 and H6. Additionally, the overlapping H3/H13b resonances were selectively inverted in the same experiment. In total, 15 more long-range RDCs were obtained from the set of SJS-HSQC experiments (Table 5.3). Importantly, the selective inversion of H5 allowed the measurement of its couplings to C1, C2, C3, C6, C7 and C15, hence furnishing enough couplings for, in principle, fit the alignment tensor.

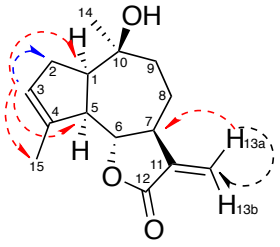


FIGURE 5.8: 10-epi structure showing the measured $^{n}D_{CH}$ couplings from H3 and H13b. n is color coded: n= 1, n= 2, n= 3.

Although, as explained previously and based on the ^{1}H-flip experiment requirements, it seems to be a prerequisite for the selective inversion of a proton resonance (H_A) that is well resolved and isolated from other resonances in the spectrum, in such a way it can be selectively inverted while the other resonances still unaffected. In practice, this requirement is not relevant as long as the other overlapped resonances ($H_{A'}$) are not coupled to the C_B whose coupling (H_A–C_B) is to be measured. This was demonstrated for the resonances H3 and H13b, whose spectral proximity allowed the inversion with the same selective pulse and are located in opposite parts of the 10-epi molecule (Figure 5.8). This single experiment furnished the long-range couplings between H13b and C7 and between H3 and C1, C2, C5, and C15. All these couplings are depicted in the wrapped Figure 5.8. The corresponding peaks of the SJS-HSQC are shown in Figure 5.9. In the same way as for JS-HSQC, SJS-HSQC experiments needs to be acquired both on isotropic and aligned states of the sample, with the isotropic $^{2}J_{CH}$ and $^{3}J_{CH}$ couplings being configurationally and conformationally informative by themselves as discussed above.[190,211]

Importantly, the simple and clean multiplet pattern originated by the selective inversion in the SJS-HSQC experiment allows the direct extraction of the couplings from the 2D spectrum. Particularly in this case, coupling extraction was performed in a pseudo-automatic way by using the automatic peak-picking module included in TopSpin 3.2 software, fitting the center of the peak with the parabolic interpolation algorithm.[212]

Sign determination of the coupling from the $^{2}J_{CH}$ is straightforward as $^{3}J_{HH}$ is known to be positive. In this case, the sign of a $^{2}J_{CB–HA}$ coupling can be determined by the E.COSY tilt from the known sign of the corresponding $^{3}J_{HB–HA}$. It is important to remind the sign relationships derived from the E.COSY effect, refering to lowest-^{13}C frequency peak as the α component of the doublet, if β component is at higher-^{1}H field strength the couplings have opposite sign; whilst if the β peak is at lower-^{1}H field strength, the couplings have the same sign (Figure 5.4). In the case of $^{3}J_{CH}$, it is known to be positive or very close to zero. The determination of the sign of the $^{n}T_{CH}$ coupling can be challenging as $^{n}D_{CH}$ can be either possitive or negative. It is often the case that, when compared to $^{n}J_{CH}$, an oppositely signed $^{n}J_{CH}$+$^{n}D_{CH}$ would require a $^{n}J_{CH}$ value outside of the achievable range. Care should be taken when $^{n}J_{CH}$ is small, as both possible solutions may furnish couplings in the feasible range. In this case, both possibilities must be explored when fitting the RDCs to the structure. Additionally, the measurements can be repeated and higher or smaller alignment strength, in the particular case of compressible gels or, as shown in Chapter 2, ionic additives can be added to the gel to tune the alignment degree, in the case of strained and not compressible gels.

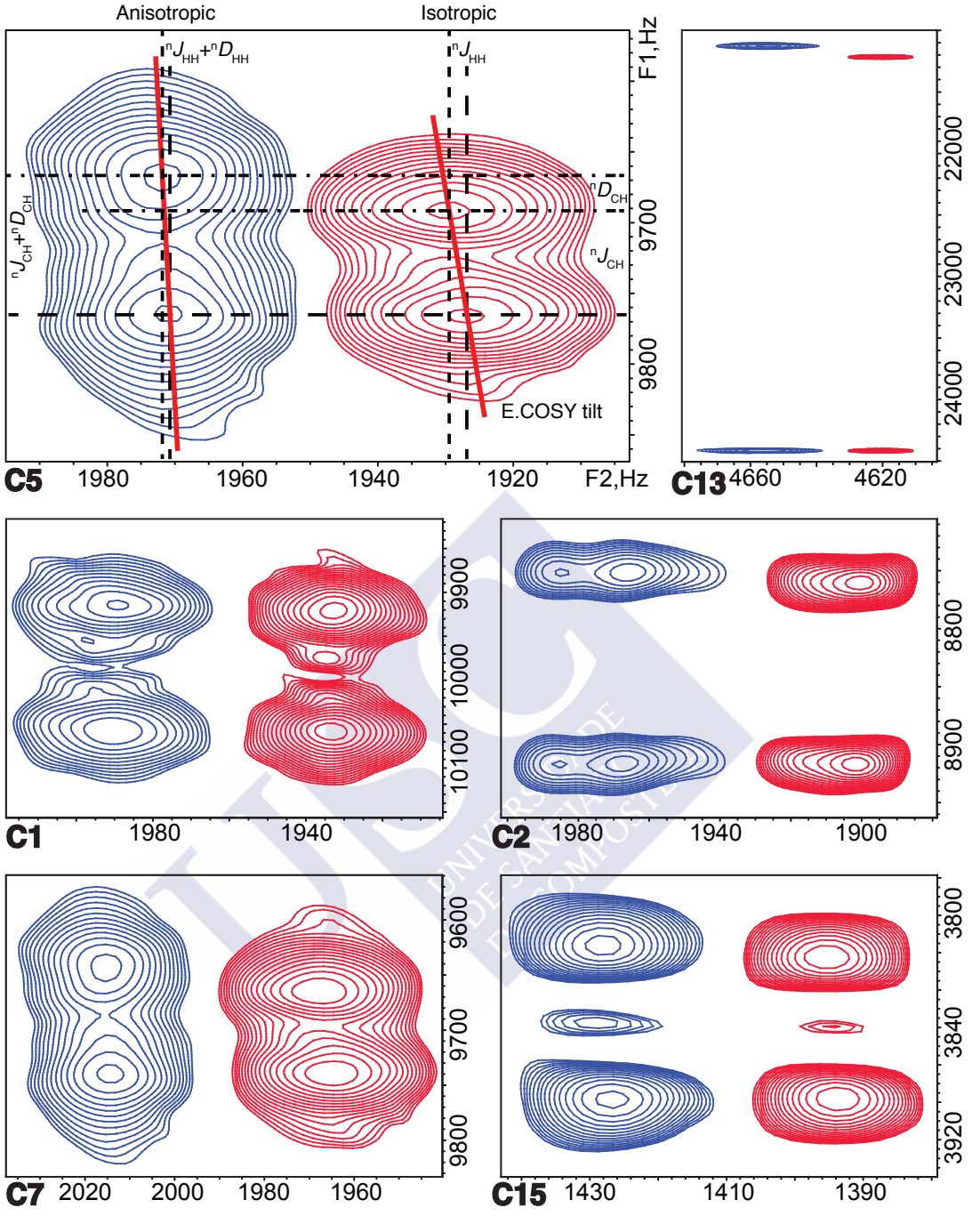


FIGURE 5.9: Overlayed SJS-HSQC spectra of isotropic (red) and anisotropic (blue) experiments, from the simultaneous H13b and H3 selective inversion.

The error on the determination of the center of the peak can be estimated with the procedure proposed by Kontaxis *et al.*[48] Briefly, for well-resolved doublets, the lower limit for the determination of the center of the peak can be given by LW/SN, where LW is the linewidth (down-scaled by κ) and SN the signal-to-noise ratio in the weaker of the two 2D spectra, which is usually the one acquired from the aligned sample.

TABLE 5.3: 10-epi-8-deoxycumambrin B long-range *J* couplings and their corresponding RDCs and experimental errors, in Hz.

Coupled Pair	n[a]	${}^{n}J_{CH}$	${}^{n}D_{CH}$	Error
H13a–H13b[b][c]	2	$\leq 2; \geq -2$	11.0	2.00
C1–H2a	2	$\leq 0.5; \geq -0.5$	1.90	0.50
C3–H2a	2	−5.99	1.53	0.05
C1–H3	3	6.88	0.28	0.09
C2–H3	2	7.82	0.42	0.05
C5–H3	3	8.96	0.33	0.03
C15–H3	3	3.80	0.31	0.01
C1–H5	2	−2.86	−1.55	0.12
C2–H5	3	–	0.10	0.10
C3–H5	3	4.47	−0.06	0.03
C6–H5	2	−7.77	−1.22	0.03
C7–H5	3	3.80	0.42	0.14
C15–H5	3	–	0.90	0.50
C5–H6	2	−0.60	−1.00	0.20
C8–H6	3	2.74	0.62	0.07
C7–H13b	3	3.71	1.14	0.08

[a] All couplings are expressed as ${}^{n}J_{A–B}$, being A–B the coupled atoms (^{13}C and ^{1}H, respectively) and n the number of bonds in between.
[b] Values obtained from the HSQC–HMQC experiment, see text.
[c] ${}^{2}J_{HH}$ coupling.

For 10-epi, κ-scaled ^{13}C linewidths in the SJS-HSQC experiments were measured in the $1 - 1.5$ Hz range, with SN ranging between 20 and 100. Therefore, for well-resolved doublets, estimated errors in the measurement of splittings fall in the $0.01 - 0.1$ Hz range (Table 5.3).

In the case of unresolved doublets in the ^{13}C dimension, the E.COSY pattern can allow the measurement of the sign and the magnitude of the coupling, undoubtedly with higher errors (${}^{2}J_{H6–C5}$ in Figure 5.10).

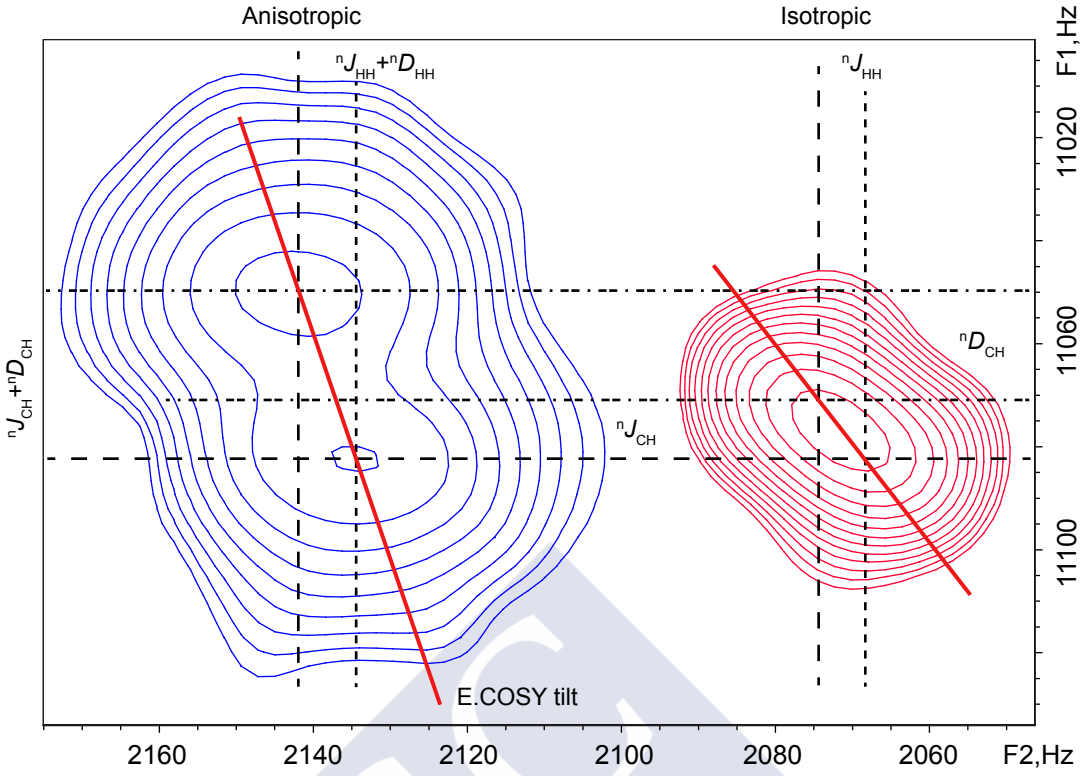


FIGURE 5.10: Superimposed SJS-HSQC spectra of isotropic (red) and anisotropic (blue) C5 resonance showing the coupling with the inverted H6.

For several ^{13}C nuclei, for example C1 (in the wrapped Figure 5.11), and C5 (Figure 5.10) an increased linewidth can be appreciated in the aligned sample. This line broadening can be attributed to enhanced proton–proton dipolar coupling interaction and shorter relaxation times in the aligned state. Therefore, there is an increased uncertainty regarding the exact value and sign of the coupling (see Table 5.3).

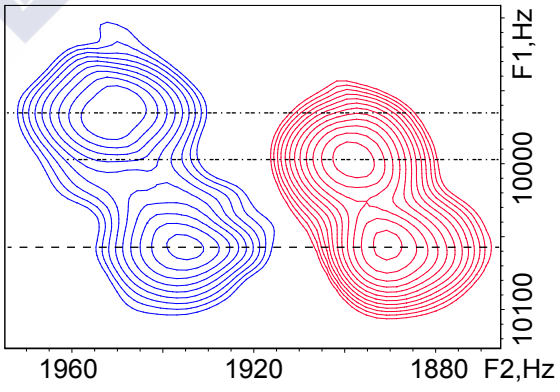


FIGURE 5.11: Superimposed SJS-HSQC spectra of isotropic (red) and anisotropic (blue) C1 resonance arising from H5 inversion.

5.3.2 Evaluation of the structural discrimination of $^1D_{CH}$ and $^nD_{CH}$

When determining the configuration of small molecules with rigid or semirigid

skeletons based on RDC analysis, usually the correct configuration is chosen as the one that best fits the RDC data among a complete set of structures exploring all possible configurations.[58,68,92,100] RDC fits were done with the MSPIN program.[41–43]

In order to evaluate the impact of long-range RDCs on the structural discrimination of **5**, two different scenarios were compared. The scenario **A** contained all $^{1}D_{CH}$ save for the averaged couplings from methylenic C2, C8, and C9, and from the methyl groups C14 and C15 and including the additional $^{2}D_{HH}$ H13a–H13b coupling. The scenario **B** contained all available $^{1}D_{CH}$, including the averaged RDCs (from methyl and methylene groups) and the homonuclear $^{2}D_{HH}$ H13a–H13b RDC. In both scenarios, an initial fit with only the short-range data was performed. Afterwards, the available long-range couplings were introduced in the fit and the differences in structural discrimination were evaluated.

In the particular case of configurations with more than one predicted conformer in solution (*RSSSR*, *RSSSS*, *SRRSS*, *SRSSR*) multiconformer fit must be performed. The conformers were superimposed by minimizing the distance of the heavy atoms, and their populations were constrained to the Maxwell–Boltzmann predicted population (Table 5.1) in the single-tensor SVD fit.[47,70] Ensembles in which a single structure represented more than the 90% population according to the computed DFT energies the SVD[41] fit was done using only the basal conformation. Fit quality was scored using the Cornilescu quality factor, Q_C.[44]

5.3.2.1 Scenario A (excluding averaged protons): fit of short-range RDCs

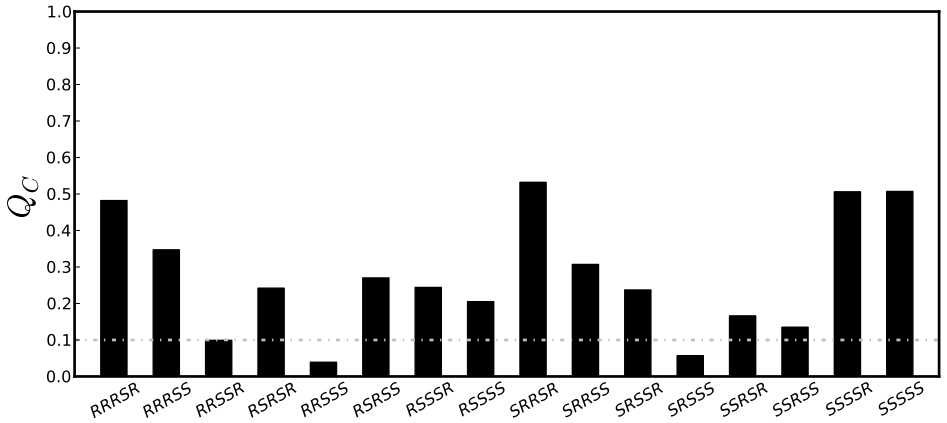


FIGURE 5.12: Scenario **A**, SVD fit of short-range RDCs of 10-epi geometries. Quality of the fit is expressed in function of Q_C factor. A gray, dash-dotted line indicates the usual "acceptance" level.

Scenario **A** includes the available $^1D_{CH}$ couplings save for the averaged ones (from methylene and methyl groups). As a result of the initial fit for the scenario **A**, three isomers were found to fit well with Q_C factors below 0.1: 10-epi *RRSSS* ($Q_C = 0.039$), *SRSSS* ($Q_C = 0.057$), and *RRSSR* ($Q_C = 0.099$, Figure 5.12 and Table 5.4).

The correct structure (*RRSSS*) furnished the best Q_C factor, but the fit score difference with the closest *SRSSS* diastereomer ($\Delta Q_C = 0.018$) is not as high as desirable for the success of structural determination based on RDC fits. Additionally, the fit of experimental (D^{exp}) and back-calculated (D^{calc}) RDCs is comparably good in several diastereomers, hampering the structural discrimination (Figure 5.13).

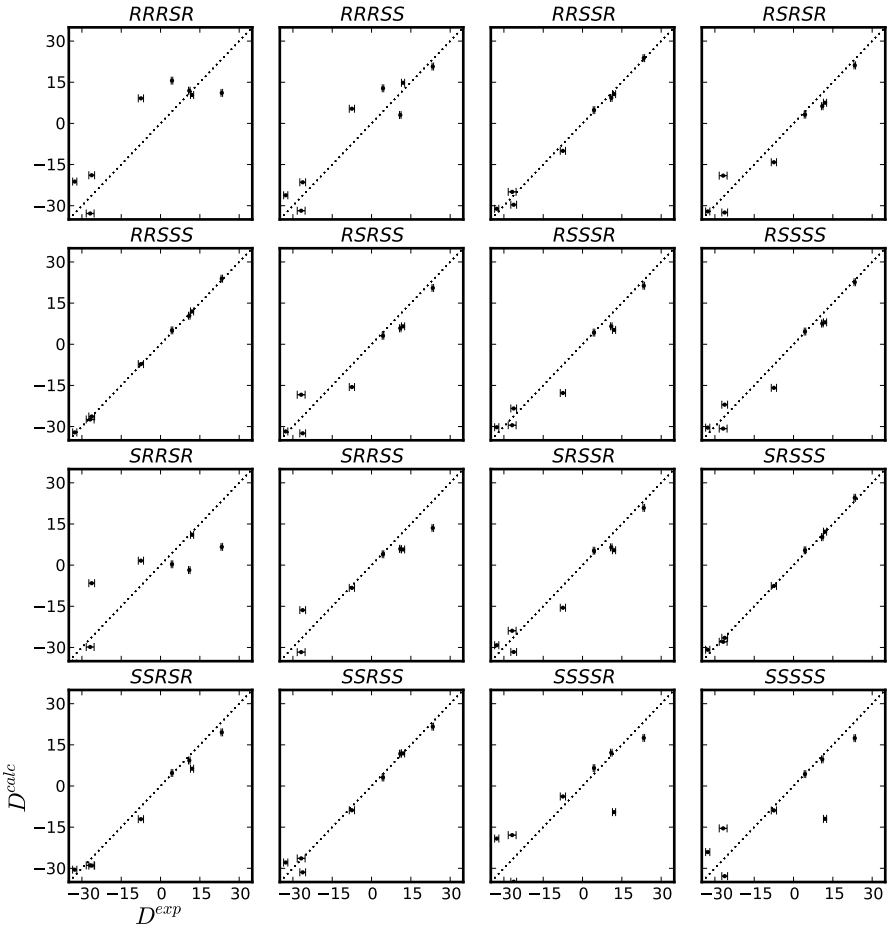


FIGURE 5.13: Plots of experimental (D^{exp}) *vs.* back-calculated (D^{calc}) Scenario **A** short-range RDCs resulting from the fit of 10-epi geometries. The error bars are set to the experimental error (Table 5.2).

5.3.2.2 Scenario B (including averaged protons): fit of short-range RDCs

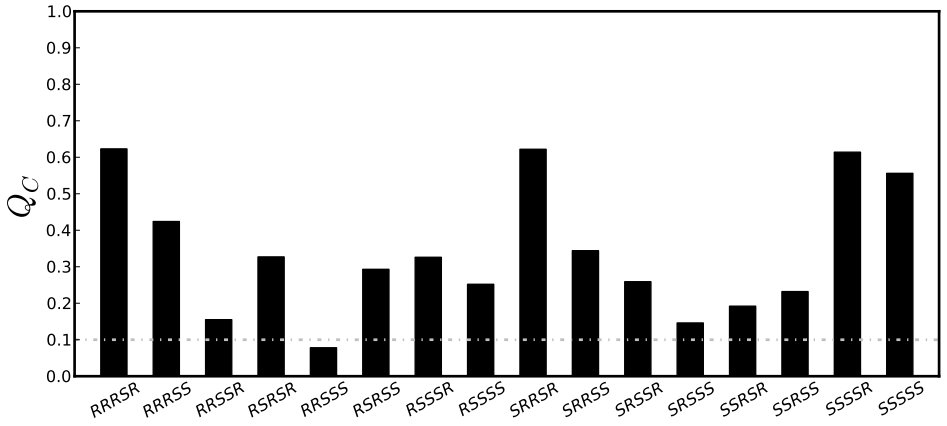


FIGURE 5.14: Scenario **B**, SVD fit of short-range RDCs of 10-epi geometries. Quality of the fit is expressed in function of Q_C factor. A gray, dash-dotted line indicates the usual "acceptance" level.

In the second scenario, when the CH_2 half-sum splitting for C2, C8, and C9, as well as the methyl groups RDCs were included in the fit, but no long-range RDCs were used, the Q_C factor of 10-epi increased to 0.078, whilst the Q_C factors for *SRSSS* and *RRSSR* increased to 0.146 and 0.155, respectively (Table 5.4 and Figure 5.14). The inclusion of more RDCs pushed forward the difference between the *RRSSS* and the *SRSSS* diastereomers ($\Delta Q_C = 0.068$, Table 5.4).

The improvement in the discrimination between the mentioned candidate conformations can be noticed looking at Figure 5.15, showing the correlation of experimental and computed RDCs for every candidate structure.

Up to this point, the structural discrimination is almost complete. However, the inclusion of long-range couplings will push Q_C values to safer levels of confidence, making the discrimination error-proof.

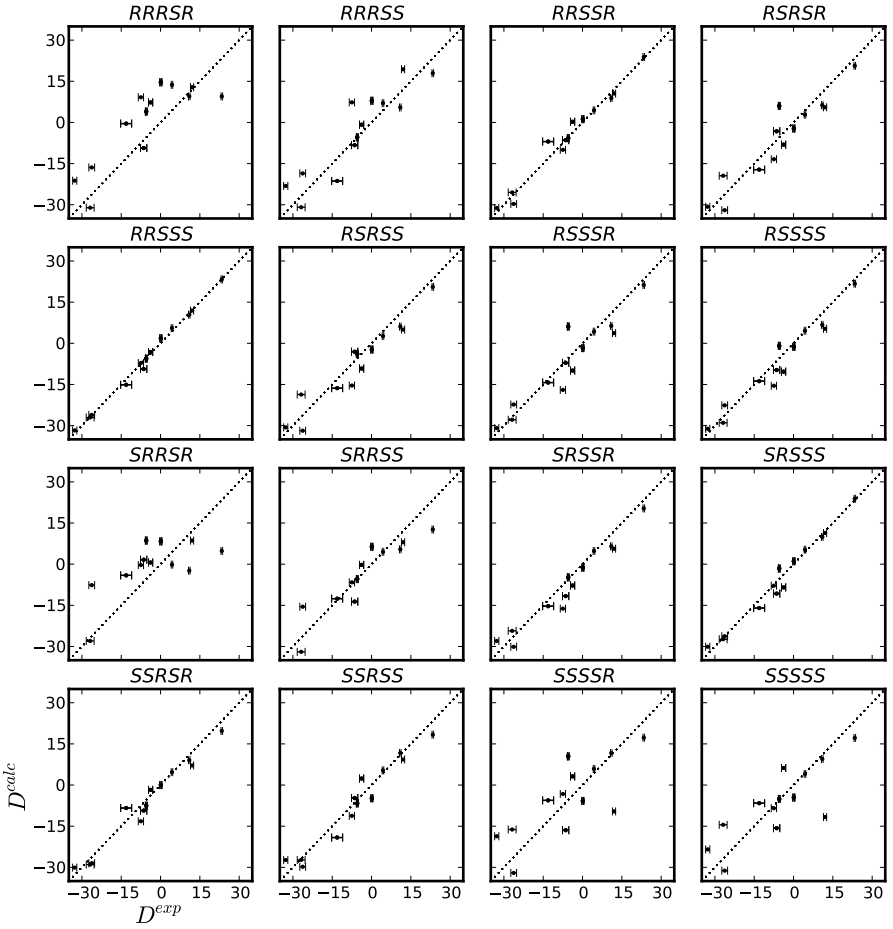


FIGURE 5.15: Plots of experimental (D^{exp}) *vs.* back-calculated (D^{calc}) Scenario **B** short-range RDCs resulting from the fit of 10-epi geometries. The error bars are set to the experimental error (Table 5.2).

5.3.2.3 Scenario A (excluding averaged protons): fit of short- and long-range RDCs

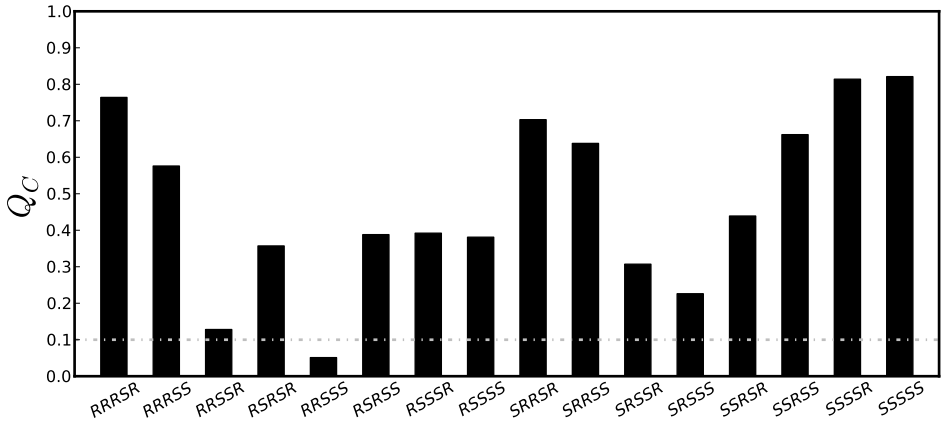


FIGURE 5.16: Scenario **A**, SVD fit of long-range RDCs of 10-epi geometries. Quality of the fit is expressed in function of Q_C factor. A gray, dash-dotted line indicates the usual "acceptance" level.

When the 15 long-range couplings are including in the scenario **A** fit, the discrimination between both *RRSSS* and *SRSSS* structures, which were the two closest diastereomers in the initial fit, became complete ($\Delta Q_C = 0.168$, Table 5.4). Visual inspection of the Figure 5.17, in which experimental and computed RDCs for every candidate configuration are compared, shows the resemblance of SVD-computed and experimental RDCs of the *RRSSS* configuration.

When the 15 additional ${}^{n}D_{CH}$ were introduced in the SVD procedure, fit of 10-epi (*RRSSS*, the correct structure) worsened marginally giving a $Q_C = 0.060$. This rising of the Q_C factor was expected as more experimental constraints were introduced in a fit with the same number of unknowns. Of key importance was the behavior of the diastereomers previously indistinguishable (short-range fit) *SRSSS* and *RRSSR*, which showed a larger deterioration in the fitness when long-range data was introduced in the fit ($Q_C = 0.228$ and $Q_C = 0.128$, respectively; Table 5.4 and Figure 5.16).

To have comparable weights in the fitting process when using both one-bond and long-range couplings, experimental and computed long range RDCs were scaled by r_{CH}^3, where r_{CH} is the distance between the coupled nuclei, furnishing scaled ${}^{n}D_{CH}$ couplings with a magnitude comparable to ${}^{1}D_{CH}$ couplings.

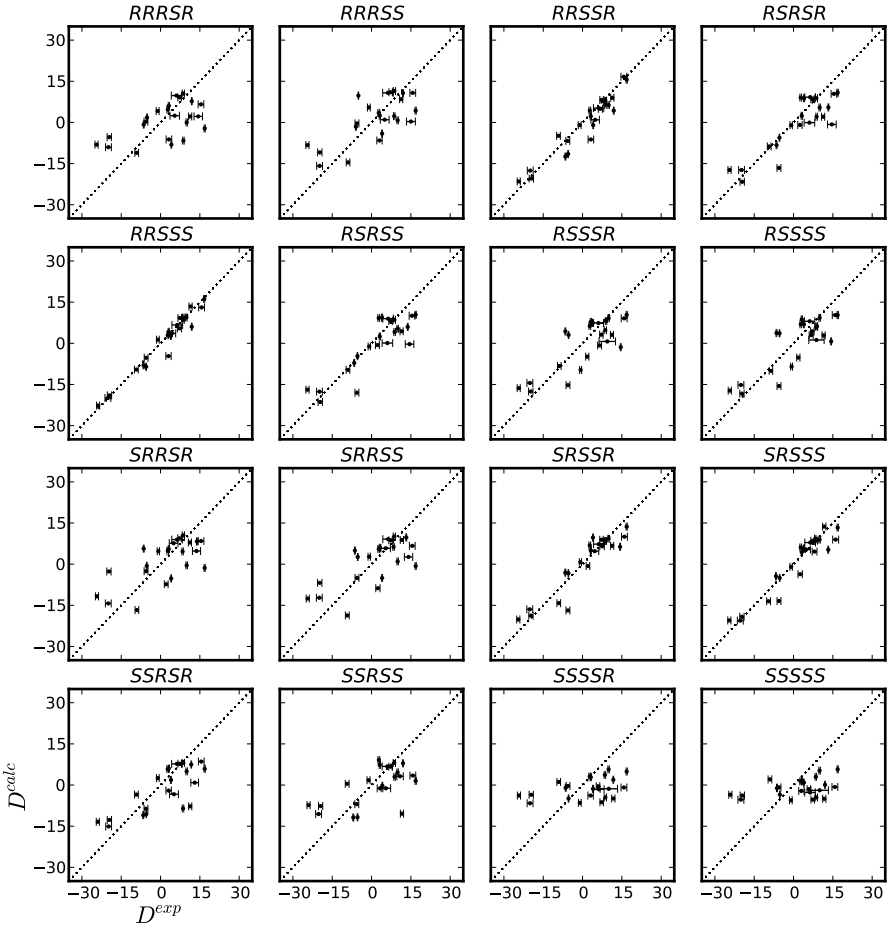


FIGURE 5.17: Plots of experimental (D^{exp}) *vs.* back-calculated (D^{calc}) Scenario **A** long-range RDCs resulting from the fit of 10-epi geometries. The error bars are set to the experimental error (Table 5.2 and Table 5.3).

5.3.2.4 Scenario B (including averaged protons): fit of short- and long-range RDCs

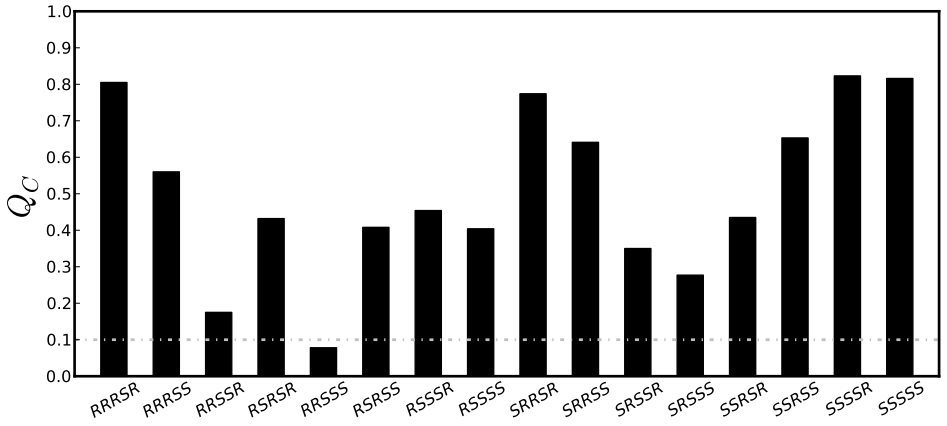


FIGURE 5.18: Scenario **B**, SVD fit of long-range RDCs of 10-epi geometries. Quality of the fit is expressed in function of Q_C factor. A gray, dash-dotted line indicates the usual "acceptance" level.

When long-range couplings were included in the RDC set of scenario **B**, 10-epi structure fit quality remained almost the same (from $Q_C = 0.078$ to $Q_C = 0.085$, Figure 5.18). As already explained, some degree of fit worsening is expected from the introduction of more restraints, without increasing the number of adjustable parameters (unknowns). Importantly, the Q_C difference between the structures that were indistinguishable by short-range RDC fit (*RRSSS* and *SRSSS*) is, in this scenario, $\Delta Q_C = 0.193$.

Additionally, in an equivalent behavior as scenario **A**, the configurations that furnished the lowest Q_C factors in the short-range fit (*SRRSS*, *SSRSR*, and *SSRSS*), increased significantly their Q_C factors. This is illustrated in Figure 5.19, in which experimental and computed RDCs for every candidate configuration are compared taking into account the experimental errors in the determination of the coupling.

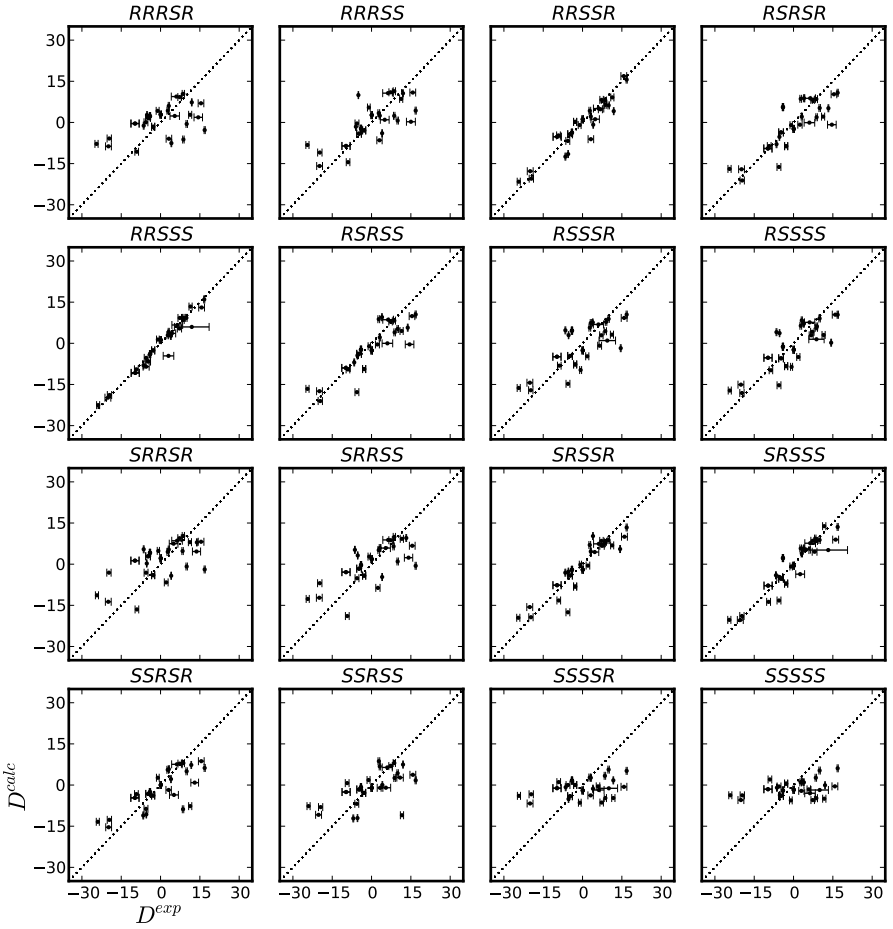


FIGURE 5.19: Plots of experimental (D^{exp}) *vs.* back-calculated (D^{calc}) Scenario **B** long-range RDCs resulting from the fit of 10-epi geometries. The error bars are set to the experimental error (Table 5.2 and Table 5.3).

As stated in the introduction to long-range couplings, the use of such anisotropic parameter is quite sparse. This novelty, along with the possibility of more than one reasonable value for a long-range coupling, made necessary the back-prediction of all measured ${}^{n}D_{\mathrm{CH}}$ from the tensor obtained in the scenario **A** with only short-range ${}^{1}D_{\mathrm{CH}}$ and in the equivalent scenario **B**. As depicted in Figure 5.21, experimental long-range couplings of **5** were correctly predicted from the short-range RDC fit determined tensor.

This comparison probes the alignment tensor similarity between short- and long-range fits. Furthermore, this demonstrates the correct assignment of the ${}^{n}J_{\mathrm{CH}}$ and ${}^{n}T_{\mathrm{CH}}$ signs, based on the E.COSY pattern enconded by the SJS-HSQC experiment.

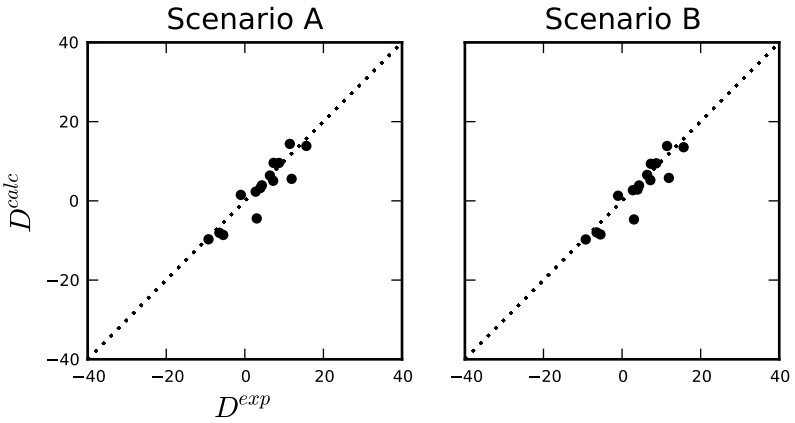


FIGURE 5.20: Plots of experimental (D^{exp}) *vs.* back-calculated (D^{calc}) Scenario **A** and Scenario **B** back-calculated long-range RDCs resulting from the short-range RDC fit of 10-epi geometries.

FIGURE 5.21: Scenario **A** and **B**, SVD Experimental *vs.* back-calculated long-range RDCs.

The comparison between the *structural discrimination factor* between the two fit scenarios including long-range couplings (Table 5.4) illustrates the high structural value of long-range couplings, which —in this particular case— made unnecessary the inclusion of the averaged values ($\Delta\Delta Q_C = 0.025$), which were of key importance for the discrimination in the short-range-only scenarios.

5.4 Conclusion

Molecular mechanics modeling furnished all possible diastereomers of 10-epi (**5**, Figure 5.5). Further DFT refination and population computations following Maxwell-Boltzmann statistics resulted in four out of sixteen diasteromers presenting more than one populated conformer.

The use of JS- and SJS-HSQC experiments along with the *J*-HMQC-ge/se-HSQC permitted the extraction of 12 $^{1}D_{\mathrm{CH}}$, 1 $^{2}D_{\mathrm{HH}}$ and 15 long-range (6 $^{2}D_{\mathrm{CH}}$ and 9 $^{3}D_{\mathrm{CH}}$) (Table 5.2 and Table 5.3).

RDC fits separated in two different scenarios, concerning the inclusion of averaged one-bond RDCs (scenario **B**) or not (scenario **A**) clearly pointed to *RRSSS* as the correct 10-epi configuration.

Table 5.4: 10-epi-8-deoxycumambrin B RDC fit results.

	Scenario **A**		Scenario **B**		
Isomer	$^1D_{CH}$	$^1D_{CH}$+$^nD_{CH}$	$^1D_{CH}$	$^1D_{CH}$+$^nD_{CH}$	$^nD_{CH}$
RRRSR	0.482	0.763	0.623	0.805	0.259
RRRSS	0.347	0.576	0.424	0.560	0.270
RRSSR	0.099	0.128	0.155	0.178	0.149
RSRSR	0.242	0.360	0.327	0.435	0.249
RRSSS	0.039	0.060	0.078	0.085	0.084
RSRSS	0.270	0.391	0.293	0.412	0.219
RSSSR	0.246	0.392	0.330	0.453	0.306
RSSSS	0.205	0.381	0.252	0.405	0.250
SRRSR	0.532	0.703	0.622	0.774	0.239
SRRSS	0.307	0.636	0.344	0.640	0.234
SRSSR	0.237	0.335	0.259	0.352	0.156
SRSSS	0.057	0.228	0.146	0.278	0.124
SSRSR	0.166	0.439	0.192	0.437	0.303
SSRSS	0.135	0.661	0.232	0.654	0.291
SSSSR	0.506	0.813	0.614	0.824	0.409
SSSSS	0.507	0.822	0.556	0.819	0.495

Interestingly, the accuracy of measured long-range couplings pushed forward the structural discrimination defined as Q_C difference between the correct (*RRSSS*) and the structurally closest diastereomer (*SRSSS*).

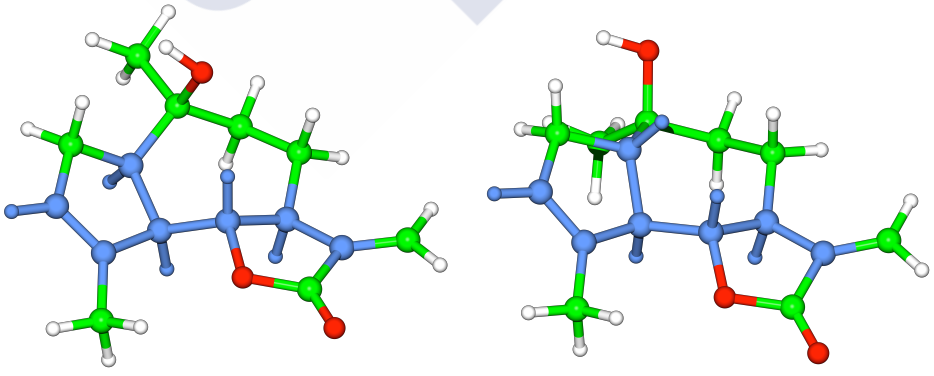


Figure 5.22: 10-Epi *RRSSS* and *SRSSS* diastereomers shown in the same relative orientation. Relevant carbon squeleton is shown in blue. See text for details.

In Figure 5.22 the superimposition of both structures is shown. The carbon skeleton comprising C3–C4–C5–C6–C7–C11 and C1–C5–C6–C7–C11 is almost superimposable

in both structures. The structural differences between these two lowest-Q_C-structures is located in the carbon skeleton involving C8–C9–C10–C14, and this is the structural reason of the lack of discrimination of the short-range scenario **A**.

In conclusion, the SJS-HSQC experiment acquired in isotropic and anisotropic conditions furnished up to 15 long-range RDCs, that permitted the determination of the correct configuration of the 10-epi molecule, with 5 stereogenic centers, based only in RDC analysis.

5.5 Materials and Methods

5.5.1 Materials

The sample of 10-epi-8-deoxycumambrin B (**5**) was reisolated from *Stevia yaconensis* var. *subeglandulosa*[208] by Dr. Viviana E. Nicotra from Córdoba National University, Córdoba, Argentina.

5.5.2 NMR experiments

All NMR experiments were carried out at 298 K on a Bruker *Avance III* NMR spectrometer operating at 747.33 MHz ^{1}H frequency and equipped with a triple resonance room temperature TXI probehead, containing a three-axis pulsed field gradient accessory. The assignment spectra can be seen in Appendix A, Section A.4.1 and the assignment in Appendix A, Section A.4.2.

JS-HSQC spectra were acquired as a 512* (^{13}C) $\times$ 800* (^{1}H) data matrices, where N* refers to N complex pairs, with acquisition time of 22.5 ms (^{13}C) and 89.2 ms (^{1}H), and a spectral window (SW) of 22727 Hz for ^{13}C dimension and 8971 Hz for ^{1}H dimension, using 4 transients per FID and a 1.5 s delay between scans. J amplification factor κ was set to 3. The total measurement time was approximately 1.8 h per spectrum.

SJS-HSQC spectra were acquired as a 512* (^{13}C) $\times$ 800* (^{1}H) data matrices, where N* refers to N complex pairs, with acquisition time of 22.5 ms (^{13}C) and 89.2 ms (^{1}H) and a spectral window (SW) of 22727 Hz for ^{13}C dimension and 8971 Hz for ^{1}H dimension, using 8 transients per FID and a 1.5 s delay between scans. J amplification factor κ was set to 20. The total measurement time was approximately 3.8 h per spectrum.

5.5.3 Alignment of 5 Using Reversible Compression / Relaxation of PMMA Gels

A 25 mm length PMMA gel was swollen and washed in $CDCl_3$ following the protocol of Gil and co-workers.[96] The maximum ^2H $|\Delta\nu_Q|$ achievable was about 27 Hz. Compound **5** (3 mg) was dissolved in $CDCl_3$ (200 μL) and added to the NMR tube containing the clean and fully relaxed swollen PMMA gel. The Shigemi plunger was subsequently inserted and the gel was compressed and relaxed several times by gently pumping it with the Shigemi plunger.

Once the gel was equilibrated with the compound **5** solution, the gel was compressed to its maximum, and the tube was sealed with Teflon tape for holding the position of the plunger and avoid solvent evaporation.[96] In the current work, all anisotropic NMR measurements were conducted with the gel fully compressed, resulting in a $CDCl_3$ ^2H $|\Delta\nu_Q|=$ 27 Hz. A separate sample of 3 mg of **5** in 500 μL of $CDCl_3$ was prepared to collect the NMR data in isotropic conditions.

5.5.4 Computational details

The conformational space of the different configurations was explored by means of molecular mechanics using the MM3 force field[150,151] and the mixed torsional/low-mode sampling algorithm[130] as implemented in MacroModel.[110]

All conformers in the resulting ensembles were then minimized at the DFT level using the OPBE functional[165,166] and the 6-31G* basis set using the Gaussian09[131] program, retaining only conformers with a relative energy below 2 kcal/mol. Twelve configurations yielded a single dominant conformation. The configurations *RSSSR*, *RSSSS* and *SRRSS* were found to have a nearly 1 : 1 ratio mixture of two conformers ($<$ 0.2kcal/mol relative energy difference) and configuration *SRSSR* showed three conformers with a population ratio of ca. 6.0 : 2.5 : 1.5 based on the DFT energy differences. DFT derived energies are shown in Appendices (Appendix B, Section B.4.2).

5.5.5 RDC fit to the candidate structures

Alignment tensor determination, back computation of RDC values and calculation of quality factors (Q_C) were conducted using the Mspin software suite.[42,43] The goodness of fit between experimental and back-computed RDCs was expressed in terms of the Cornilescu quality factor Q_C.[44] To have comparable weights in the fitting process when using both one-bond and long-range couplings, experimental and computed

long range RDCs were scaled by r_{CH}^3, where r_{CH} is the distance between the coupled nuclei.

Input and output data are reported in Appendices. Molecular coordinates of conformers **1A-C** (Appendix B, Section B.4.1), RDC input tables in MSPIN-ready format (Appendix B, Section B.4.3) and program outputs (Appendix B, Section B.4.4).

Chapter 6

Application of long-range RDC to the analysis of conformational equilibria

6.1 Introduction

6.1.1 The anti-obesity drug Lorcaserin has a flexible 7-membered ring

The 3-benzazepine skeleton (Figure 6.1) is found in many natural compounds and pharmaceuticals. Many 3-benzazepine derivatives have been found as strong effectors —either agonist or antagonist— over dopamine receptors —both D_1 and D_2 families—, along with the important adrenergic and serotonin receptors, all of them belonging to the family of transmembranal G-Protein Coupled Receptors (GPCRs).[213]

There has been considerable effort in finding derivatives with favorable pharmacological activity.[214] Noteworthy, very recently the crystal structure of the complex between the 5-HT_{1B} receptor (GPCR) with ergotamine has been resolved.[215]

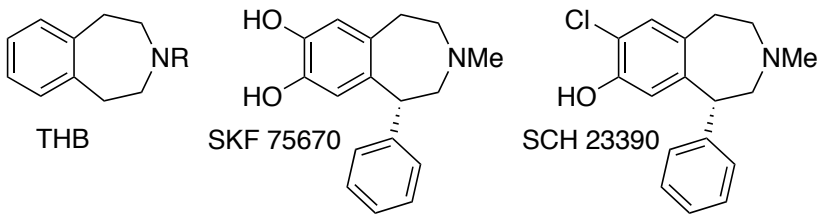


FIGURE 6.1: Relevant tetrahydro-3-benzazepines. Left: the tetrahydrobenzazepine (THB) skeleton, and right: bioactive THB D_1 ligands

Therefore, the bioactive conformation of this kind of derivatives is unknown, at least, from direct methods. Instead, there has been considerable effort in the computational analysis of the conformational properties of these compounds and Structure-Activity relationship studies,[216,218] which can give clues on the bioactive conformation.

In the absence of double or triple bonds or steric restrictions, seven-membered rings are conformationally flexible and can access a number of conformations. Fusion with the benzene ring poses some restriction on the flexibility of the seven-membered ring in tetrahydro-3-benzazepines (THB, Figure 6.1) but still they conserve a considerable degree of conformational freedom in the rest of the bonds.

Several computational studies done on compounds of this class, with diverse substituents, have identified a number of accesible conformations for the benzazepine ring, comprising crown-chair, chair, boat and twist conformations.

TABLE 6.1: Relative MM2 energies calculated for 1Ph-3-THB (NH^+) by Petterson *et al.*[218] as function of the 1-Ph substituent disposition, in kcal/mol.

conformer	equatorial Ph	axial Ph
crown-chair	0.0	0.0
twist	1.8	2.1
boat	4.2	2.0

In the 90s, Petterson *et al.*[218] studied the conformational space accessible to a series of substituted 3-THB by means of MM2 calculations, finding two possible low energy structures for the 1-phenyl free-base derivative, namely axial and equatorial crown-chairs, with the same relative energy already found in previous investigations.[217,219] Additionally, twist and boat conformers, both axial and equatorial, were also found with higher relative energies (Table 6.1).[218] Equivalent results were found by Snyder *et al.*[220] In the available crystal structures of 1-phenyl-THB derivatives, they adopt a crown-chair conformation with an equatorial disposition of the phenyl ring.[217,219] Alkorta *et al.*[221] achieved comparable results by means of AM1, MM3 and DFT calculations (Table 6.2). In the absence of any substituent on the 3-benzazepine skeleton, the crown-chair is the most stable conformation.[221]

In this chapter, we will take advantage of NMR, prominently RDC, to analyze the conformational state of lorcaserin hydrochloride, a new drug of the THB family recently approved for the treatment of obesity (Figure 6.2). Firstly, the computational and NMR-based evidences of conformational exchange will be presented. Secondly, one-bond RDC will be used to address its conformational equilibrium, in conjunction with

TABLE 6.2: Relative MM3 ($\Delta\Delta G_{298}$) energies calculated for THB (Figure 6.1, NR = NH^+), mTHB (Figure 6.1, NR = NMe^+) and SKF 75670 (Figure 6.1) by Alkorta *et al.*[221] SKF 75670 energies are shown as function of the 1-Ph substituent disposition. Energies in kcal/mol.

conformer	THB[a]	mTHB[a]	SKF 75670	
			equatorial Ph	axial Ph
crown-chair	0.0	0.0	0.0	0.3
twist	3.7	4.2	3.2	4.2
boat	3.3	3.5	3.6	5.4

[a] The energies shown correspond to the lowest energy inversion isomer.

$^3J_{HH}$ and NOE data. Finally, the procedure and benefits of introducing long-range RDC will be discussed.

6.1.2 Lorcaserin is a promising anti-obesity drug of the 3-THB class

The prevalence of overweight has become epidemic in the past decades, as recently discussed by Bays.[222] Obesity is estimated to affect at least 300 million human beings in the whole world.[222,223] Approximately 65% of United States adults are classified as overweight and more than 30% present obesity. Enforced by the wide prevalence of overweight in USA and the lack of a selective treatment, the NIH *Heart, Lung and Blood Institute*, as well as other authors, have pointed to the importance of effective prevention and development of selective obesity treatments.[224,225] Overweight has been found to be associated with conditions such as Type II *diabetes mellitus*, cardiovascular diseases —mainly, high blood pressure and atherosclerosis— and metabolic diseases.[222]

Effective treatments targeting overweight are scarce.[222] Recently, lorcaserin hydrochloride (**6**) (Figure 6.2) was developed by Arena Pharmaceuticals as the first 5-HT_{2C} highly-selective agonist.[226–228] Lorcaserin has demonstrated to help obese patients to lose weight in several clinical trials, including double blind trials.[229,230] Recenlty, the USA Federal Drug Administration (FDA) has accepted the new drug application submitted by Arena Pharmaceuticals for the chronic weight management on obese adults using lorcaserin.[231,232]

6.1.2.1 5-HT_{2C} specificity is of key importance to prevent severe side-effects

Serotonin (5-hydroxytryptamine, 5-HT) is a tryptophan-derived neurotransmitter. Among other major effects, it is known to regulate mood, appetite, and sleep. It is popularly acknowledged as a "well-being and happiness molecule". Serotonin binds

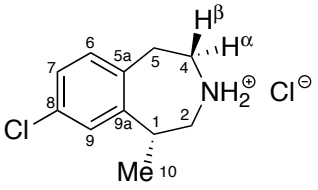


FIGURE 6.2: Structure of lorcaserin, **6**.

to at least 14 receptor subtypes —including 5-HT_{2C} receptors— that appear to be restricted to the central nervous systems in mammals. This receptor family is an important (agonist) drug target for diseases including sexual dysfunction, schizophrenia, and obesity.[233–235] Noteworthy, a lack of selectivity over the 5-HT_{2C} receptor can lead to important to severe side-effects *via* the highly homologous 5-HT_{2A} (causing hallucinations and cardiovascular conditions) and 5-HT_{2B} (associated to valvulopathy and pulmonary hypertension), among others.[216,230]

Given the difficulty to attain selectivity, it is critical to design compounds that precisely match the target binding sites. Therefore, structural studies are critical to design new generations of bioactive, selective compounds. As explained in Chapter 1, the bound conformation is very likely to be one among the dominant ones in the free-state in solution. Therefore, if the 3D structure of the receptor/drug complex is not available, solution NMR is the tool of choice for the free-state conformational analysis of drugs.[236]

In the particular case of lorcaserin hydrochloride, crystallographic analysis indicated that the molecule adopts a crown-chair conformation, but only the one with the equatorial orientation of the 1-methyl substituent (**6B**, Figure 6.3).[227] This however, does not discard the presence of other low-energy conformers in solution. In the following we will show how the conformational space of lorcaserin in quasi-physiological conditions can be determined through the use of RDC-enhanced NMR.

6.2 Computational and experimental (NMR) evidences of conformational averaging of lorcaserin in solution

6.2.1 Computational search of the conformational space of lorcaserin

The conformational space of **6** was explored by Molecular Mechanics (MM3 force field) calculations using the GMMX stochastic conformational search procedure as implemented in PCMODEL.[111,150,151] Six different conformations were found within

an energy cutoff of 5 kcal/mol. Geometries were further optimized at the DFT level using the M052X functional[132] along with the 6-31+G(d,p) basis set. Solvation effects were taken into account by the Polarizable Continuum Model in its Integral Equation Formalism (IEFPCM)[124] using Gaussian03 water parameters (Figure 6.3).[167]

Computed analytical frequencies were inspected to verify that the stationay points were true minima. It is important to note that the chloride counterion was not taken into account either in the force field or in the DFT level calculations due to technical difficulties. All relative energies are reported as Gibbs ($\Delta G_{298.15K}$) free energies differences.

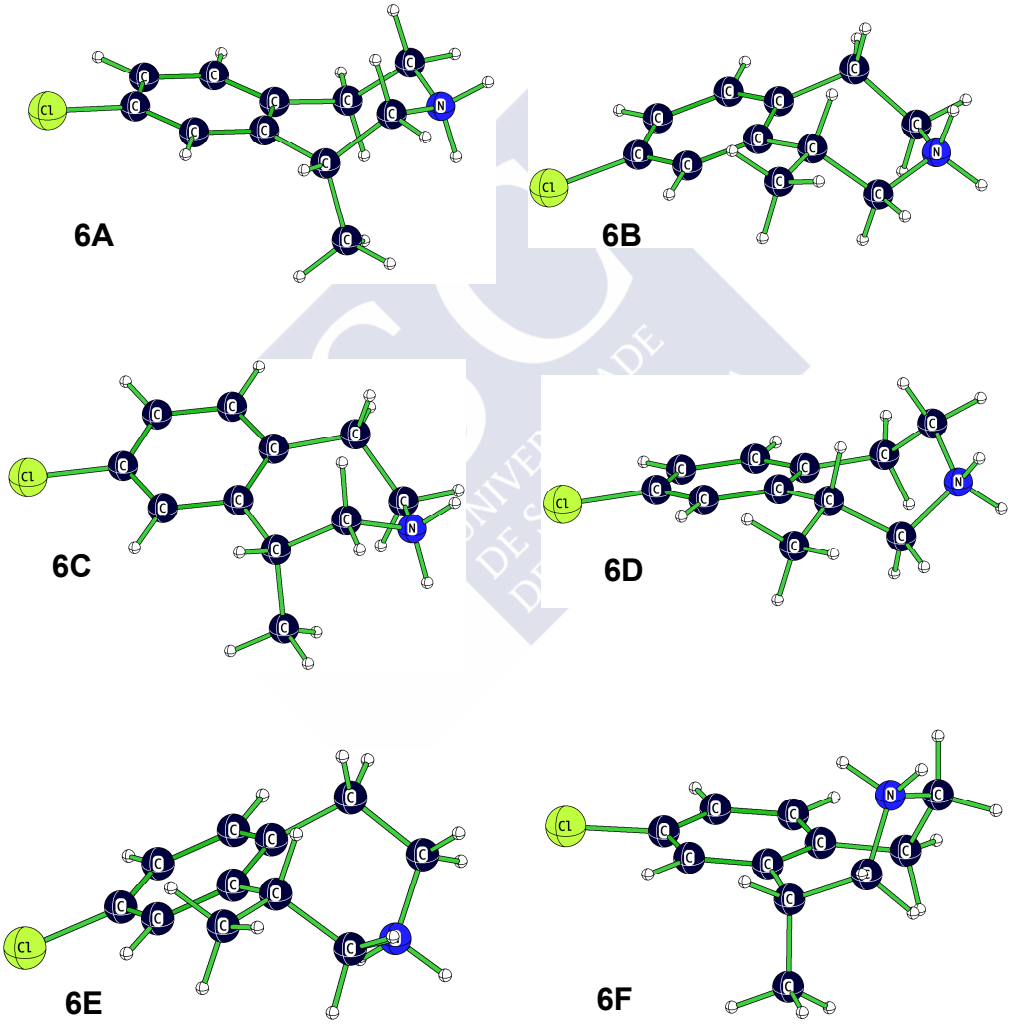


FIGURE 6.3: Lorcaserin conformational space.

The following conformations were found: crown-chair (**6A** and **6B**), twist (**6C** and **6D**), boat (**6E**) and twist-boat (**6F**). DFT-computed energies are presented in Table 6.3. The axial (**6A**) or equatorial (**6B**) disposition of the 1-methyl substituent has little influence on the relative energy of the crown-chair forms (less than 0.1 kcal/mol). The other

conformations have higher energies. The twist conformer with the axial methyl (**6C**) has a free energy 2.2 kcal/mol higher than **6A**, while that with the equatorial methyl (**6D**) is 1.1 kcal further above (to a total of 3.2 kcal/mol over **6A**). The least stable conformers are the boat (**6E**) and twist-boat (**6F**) with the 1-methyl in equatorial and axial orientation, respectively. It seems that the equatorial or axial orientation of the 1-methyl substituent has only a minor influence on the stability of these conformers.

Their expected populations were estimated from their DFT-computed energies following Maxwell-Boltzmann statistics (Equation (6.1)) assuming $T = 298.15$ K

$$\text{Population} = \frac{\exp(-\Delta G_{298.15K}/RT)}{\sum_{i=1}^{n} \exp(-\Delta G_{298.15K}/RT)}, \tag{6.1}$$

where $\Delta G_{298.15K}$ is the relative (DFT) free energy and n is the total number of conformers.

TABLE 6.3: Gibbs free energies and Boltzmann populations of lorcaserin conformations.

Structure	$\Delta G_{298.15K}$, kcal/mol	p_i, %
6A	0.0	51.9
6B	0.1	46.2
6C	2.2	1.4
6D	3.2	0.2
6E	3.3	0.2
6F	3.8	0.1

The **6B** crown-chair corresponds to the crystal structure. However, DFT-computed energies suggests a equilibrium in solution between **6A** and **6B**, with populations close to 50 %. We set to address this question experimentally by NMR.

6.2.2 NOESY correlations support the simultaneous axial and equatorial disposition of the methyl group

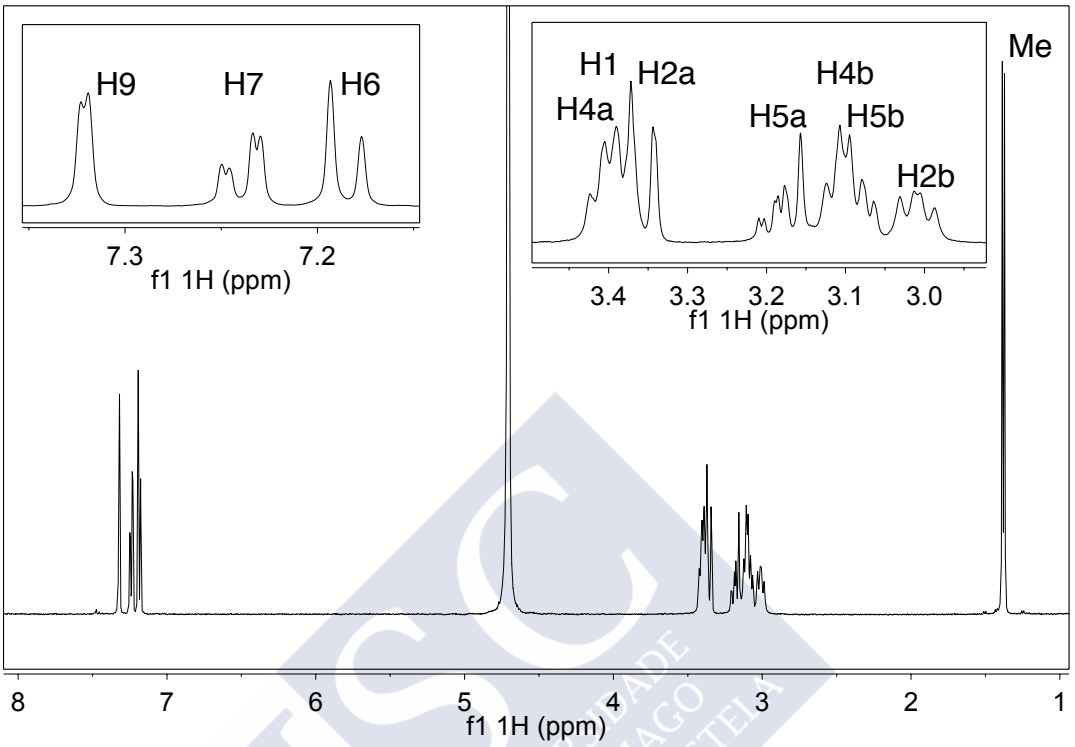


FIGURE 6.4: 1D ^{1}H NMR of lorcaserin, 500 MHz. Labels *a* and *b* correspond to the low-field and high-field resonances, respectively.

Lorcaserin was assigned from a combination of COSY, NOESY, edited HSQC and HMBC experiments. The final assignment is labeled in the following ^{1}H (Figure 6.4) and ^{13}C-HSQC (Figure 6.6) for clarity.

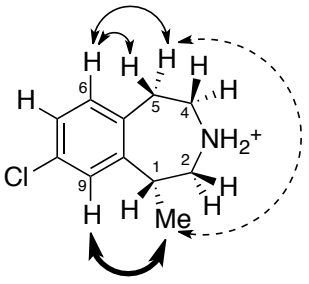


FIGURE 6.5: Representative NOE correlations of lorcaserin. *Intense*, bold line; *medium*, thin line; and *weak*, dashed line.

NOE correlations (τ_{mix} = 400 ms) of lorcaserin are depicted in Figure 6.5. Figure 6.7 shows the selective inversion of the H10 methyl signal in a 1D NOESY experiment. Methyl H10 shows the strongest NOE correlation with H9, suggesting an equatorial disposition of the methyl group (average H9/H10 distance = 2.5 and 3.5 Å in **6B** and **6A**, respectively).

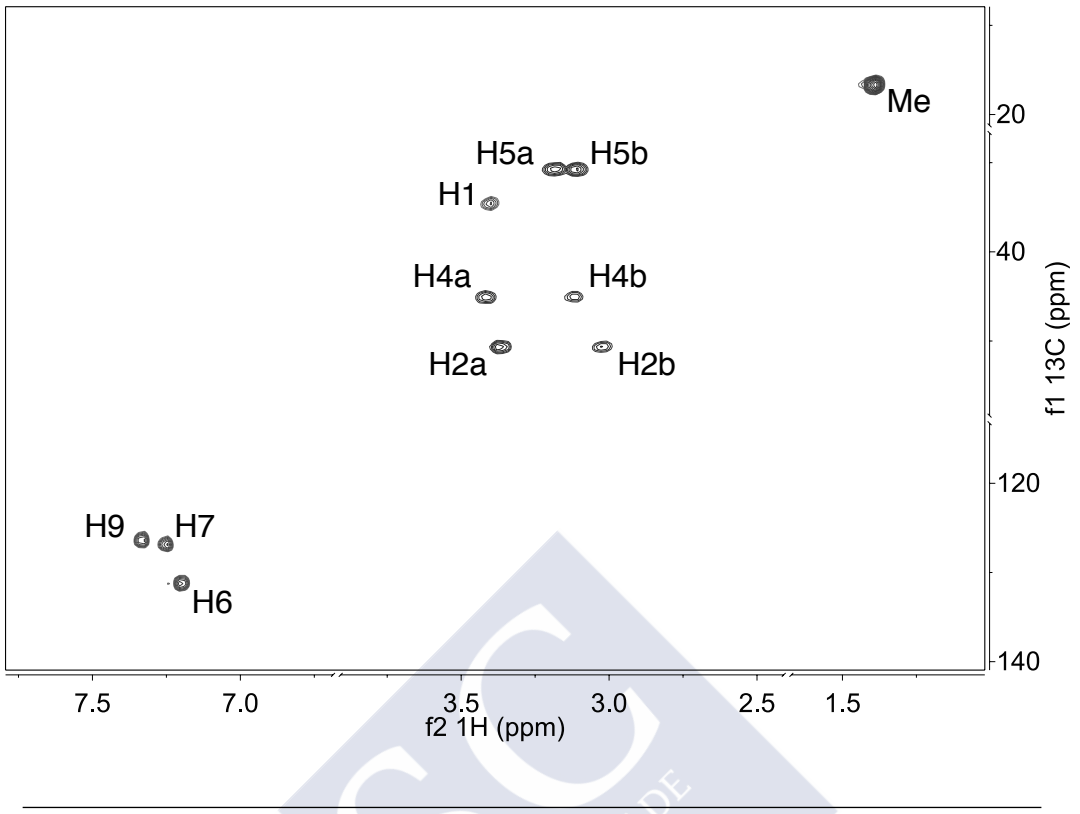


FIGURE 6.6: 2D ^{13}C-^{1}H HSQC of lorcaserin, 500 MHz. Labels *a* and *b* correspond to the low-field and high-field resonances, respectively.

Importantly, there is a weak correlation between methyl H10 and H5b (Figure 6.7, high field H5 resonance).[1] Due to distance constraints, this is possible only in conformers having an axial methyl (e.g. average distance H10/H5α is 3.20 and 3.45 Å in **6A** and **6F**, respectively, as the equatorial methyl is too far away (e.g. distance H10 / H5α is larger than 5 Å in **6B**, **6D** and **6E**). Therefore, this H10/H5b NOE correlation permits the assignment of resonance H5b to H5α and indicates the presence of a conformer with an axial disposition of the methyl group.

Additionally, selective inversion of H6 gave strong NOE correlations with both H5 protons (Figure 6.8). This can be attributed to the conformational mixing of any pair of conformers where the H5α/β protons exchange their axial and equatorial disposition (e.g. **6A** and **6B**, see Figure 6.3).

The methyl group H10 showed more intense correlations with the low-field resonance H2a than with the high-field resonance H2b. Given that the stereoassignment of H2 is yet unkown, interpretation of these intensities is unclear at this point. It will be discussed later in this chapter.

[1] See lorcaserin asignment in Appendix A, Section A.5.2.

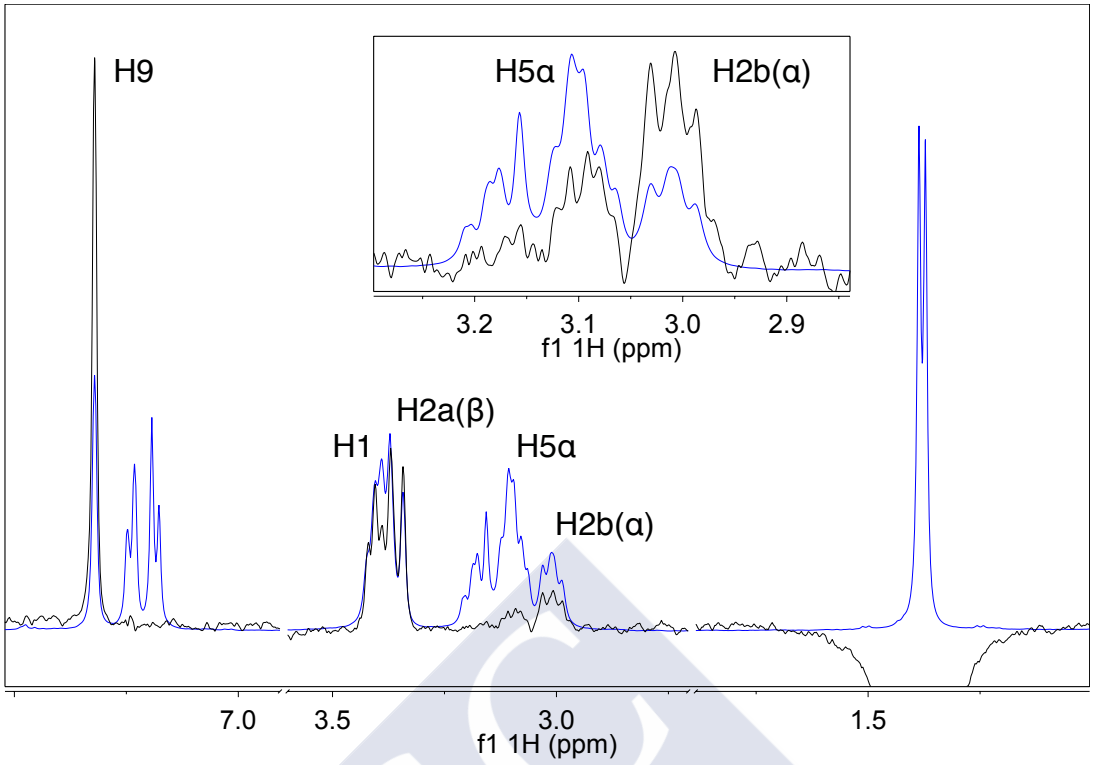


FIGURE 6.7: Lorcaserin 1D-NOE correlations from the selective inversion of the 1-Me resonance (H10, 1.35 ppm) . Blue, proton spectrum; black, NOE spectrum. $\tau_{mix} = 400$ ms, 500 MHz.

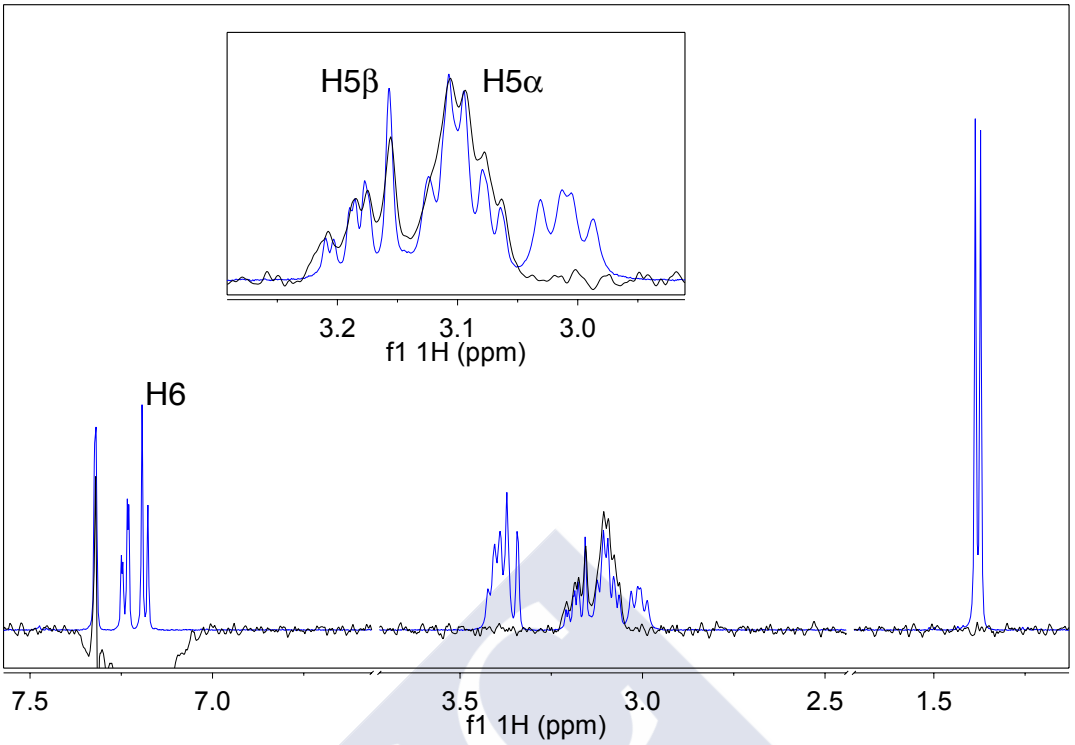


FIGURE 6.8: Lorcaserin 1D-NOE correlations from the selective inversion of the H6 resonance (H6, 7.18 ppm) . Blue, proton spectrum; black, NOE spectrum. τ_{mix} = 400 ms, 500 MHz.

In summary, these NOE correlations support the conformational exchange of lorcaserin between —at least— two conformations having alternate equatorial/axial disposition of the methyl group. According to the DFT energies, it is likely that these conformers in fast exchange are the crown-chairs **6A** and **6B**.

6.3 Analysis of one-bond RDC with the single-tensor approximation furnishes conformer populations

6.3.1 One-bond RDC are determined from ^{1}H-coupled HSQC spectra

All samples were prepared from a stock of 25 mM lorcaserin hydrochloride dissolved in deuterated water. This stock was used as isotropic sample and as swelling solution for the aligned samples in AMPS-acrylamide gels. Three aligned samples were prepared using gels neutralized with 200 mM solutions of KOH, NaOH or HCl.[94,108]

We will refer to them in the next as *K*-gel, *Na*-gel and *H*-gel, respectively. A fast F_1-coupled spectra of a small slice of the aromatic region of these gels show a remarkable difference in observed couplings (Figure 6.9 and Figure 6.10).

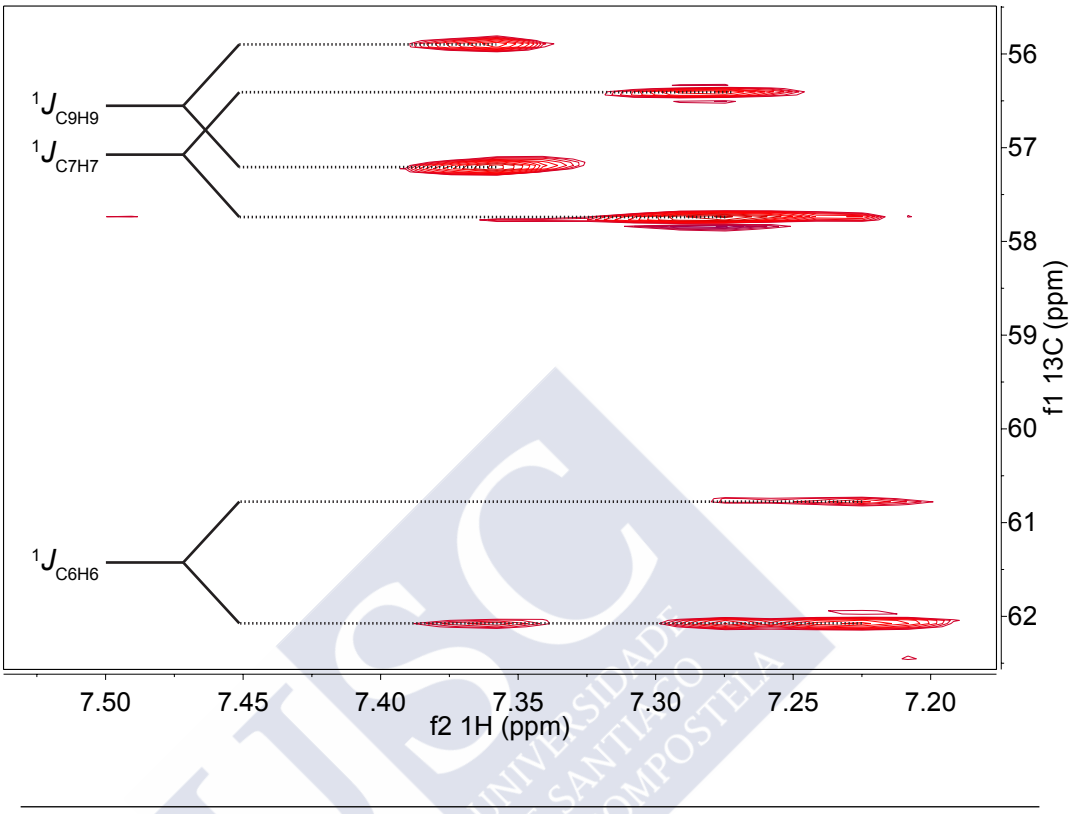


FIGURE 6.9: Lorcaserin isotropic F_1-coupled HSQC spectrum (aromatic expansion), indicating the $^1J_{CH}$ splittings. No *J*-evolution multiplication module was used in this series of HSQC.

One-bond RDCs ($^1D_{CH}$) were measured from a set of F_1-coupled *J*-Scaled (JS) HSQC spectra,[52] using a *J*-scaling factor $\kappa = 3$ (Figure 6.11). $^1D_{CH}$ were determined as the difference $^1D_{CH} = {}^1T_{CH} - {}^1J_{CH}$ between anisotropic and isotropic couplings (Table 6.4). Couplings of the methylene C–H were taken as half the splitting in the HSQC, as the observed splitting of methylene groups in F_1 is the sum of their two one-bond C–H couplings.[54]

A total of seven independent one-bond RDC were measured in each of the alignment media Table 6.4. Inspection of the RDC values indicates that the degree of alignment of lorcaserin in AMPS-acrylamide gels depends on the neutralization conditions of the gel. Gels neutralized in HCl and KOH solutions furnished one-bond RDCs in the same range (between -30 and $+9$ Hz). Instead, the *Na*-gel sample furnished smaller RDCs. This contrasts with the case of methylcodeinium ion (**1**, Chapter 2), which aligned more strongly with the *Na*-gel than with the *H*-gel. All in all, the influence of the counterion is not surprising, as was discussed in Chapter 2.

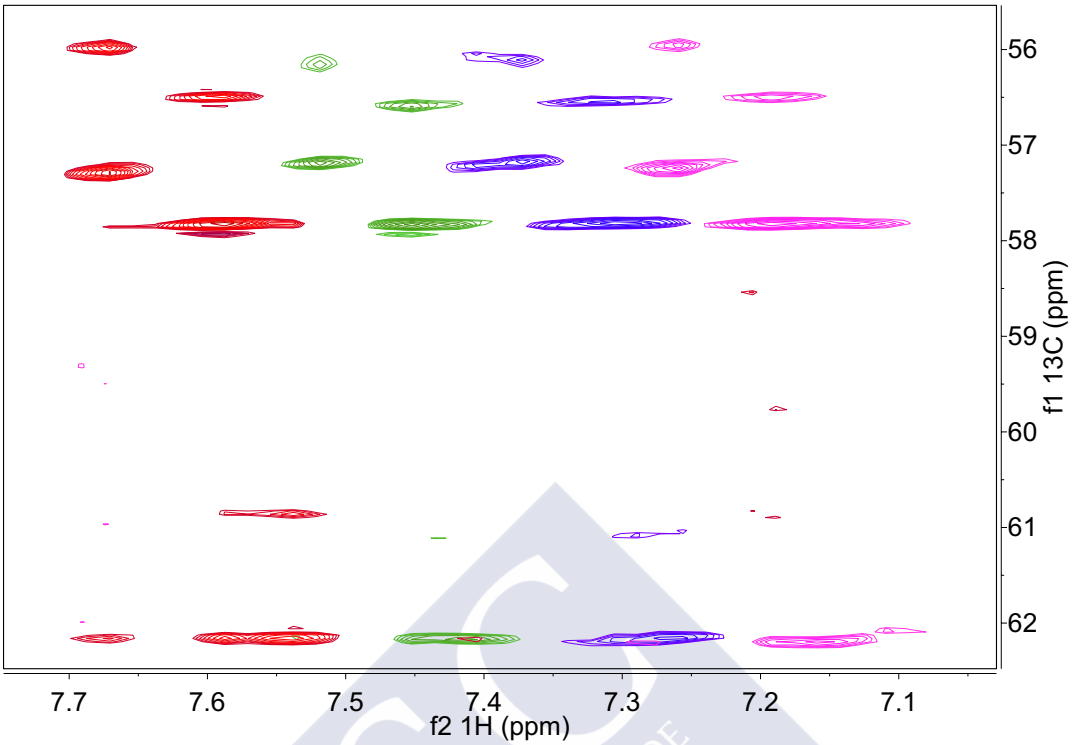


FIGURE 6.10: Overlay of a series of F_1-coupled HSQC spectra of lorcaserin in different aligning conditions. Only the expansion of the aromatic region is shown. Red, isotropic; blue, *K*-gel; pink, *Na*-gel; green, *H*-gel. No *J*-evolution multiplication module was used in this series of HSQC.

TABLE 6.4: Scalar ($^1J_{CH}$) and dipolar ($^1D_{CH}$) one-bond C–H couplings of lorcaserin, in Hz. RDCs measured in each of the three anisotropic samples are reported.

		$^1D_{CH}$			
Vector	$^1J_{CH}$	*H*-gel	*K*-gel	*Na*-gel	Error[a]
C10–H10	128.1	3.0	5.1	0.5	0.4
C1–H1	127.6	8.6	3.3	2.9	1.9
C2–H2[b]	144.9	2.1	2.4	−0.4	1.4
C4–H4[b]	145.5	2.9	4.4	1.0	0.3
C5–H5[b]	129.2	−1.3	−4.4	−1.1	0.4
C7–H7	168.7	−13.9	−8.1	−1.9	0.2
C9–H9	164.5	−30.1	−30.0	−1.9	0.7

[a] Error was estimated as $\kappa \times LW/SN$ in the H-gel spectrum, where κ is the *J*-evolution scaling factor, *LW* the linewidth in Hz, and *SN* the signal-to-noise ratio.
[b] There is only one entry per methylene group, which is the half-sum of of their two individual $^1D_{CH}$ couplings as described in the text.

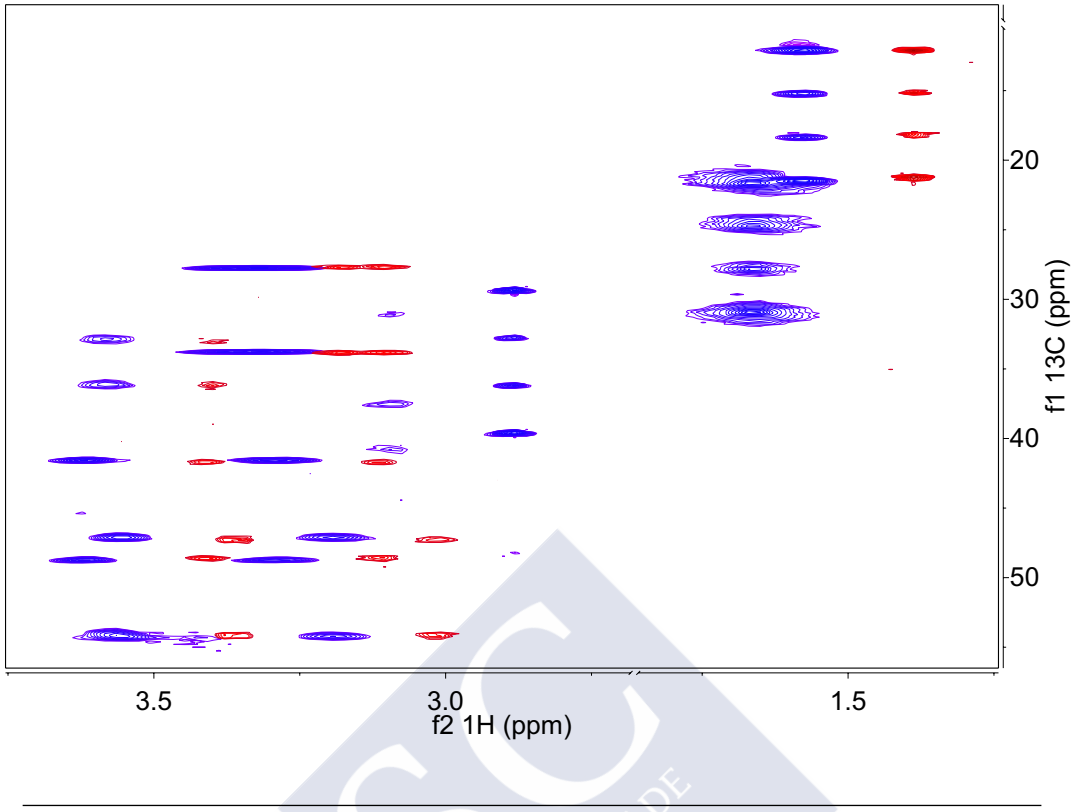


FIGURE 6.11: Overlay of two F_1-coupled HSQC spectra of lorcaserin in isotropic (red) and anisotropic (blue, *H*-gel) conditions. Only the expansion of the aliphatic region is shown. A *J*-evolution multiplication module was used with scaling factor $\kappa = 3$.

6.3.2 RDC Analysis. One-bond RDC fit to the individual conformers

In a similar way as in Chapter 5, it is safer to find first an approximate solution with the well-established methodology based on $^1D_{CH}$ and, afterwards, introduce the long-range couplings in the fit.

Initially, we followed the same procedure as in previous chapters. One-bond RDCs were fitted to each of the sole conformers of **6** using the RDCFIT scripts coded by Dr. Armando Navarro-Vázquez. The averaged RDCs from the methyl group, as well as methylenic protons half-sums were included in the calculations as previously described (see Chapter 1 and Chapter 3).[51,54] This circumvents the assignment of the individual diastereotopic methylene protons and reduces the effect of strong coupling in the accuracy of the values. The fit quality was assessed with the Cornilescu quality factor, Q_C (Table 6.5).[44]

The best *individual* fit (lowest Q_C) was obtained with conformer **6B**, which corresponds to the X-ray structure. Conformers **6E** (boat, samples *H*-gel and *K*-gel) and **6D** (twist, *Na*-gel) also gave Q_C values below 0.2 (Table 6.5, and Figure 6.12). It is common

knowledge that there is not a value of Q_C that guarantees that a fit is good (i.e. that it agrees well with the experimental data) and fits are interpreted on a case by case basis. It may happen that even the best-fit conformer (that with lowest Q_C) gives in fact a poor fit. Certainly, Q_C is a good merit function for optimization and for selecting the best structure among a set of alternative choices but further tests are usually done, like comparing experimental (D^{exp}) *vs.* back-calculated (D^{calc}) $^1D_{CH}$. Figure 6.12 shows that none of the fits is particularly good, as back-calculated $^1D_{CH}$ clearly deviate from the experimental ones.

TABLE 6.5: Fit of the experimental one-bond RDCs to the sole conformations **6A-F**. Quality of fit is expressed in terms of Q_C. Data from the three samples (*H-*, *K-* and *Na*-gel) are shown.

	Q_C		
Structure	*H*-gel	*K*-gel	*Na*-gel
6A	0.381	0.382	0.219
6B	0.132	0.099	0.051
6C	0.297	0.365	0.319
6D	0.355	0.340	0.069
6E	0.179	0.153	0.304
6F	0.342	0.314	0.371

In the fit to the single conformers, both *H-* and *K*-gels furnished comparable quality factors. *Na*-gel RDCs fit resulted in lower Q_C factor for **6B** ($Q_C = 0.052$) and **6D** ($Q_C = 0.069$), which are values in the usual acceptance region. The small quality factor of conformer **6B** is not surprising but the **6D** small quality factor is completely different than the result in the other gels.

It is not surprising that fit of the *Na*-gel data deviates from fits with the *H*-gel and *K*-gel data, as experimental RDC of the former sample are of very small size (Figure 6.12). Data of the *Na*-gel sample was analyzed with the others only for completion and was not used any further.

The bad fit of $^1D_{CH}$ RDCs (Table 6.5 and Figure 6.12) to the candidate structures **6A-F**, along with NOE evidences (Figure 6.7) of conformational averaging, point to the presence of more than one populated conformer in solution. This is also compatible with the DFT-calculated energies of the conformers under study. We will address in the next the problem of conformational averaging in solution by means of RDC fits taking advantage of the single-tensor approximation.

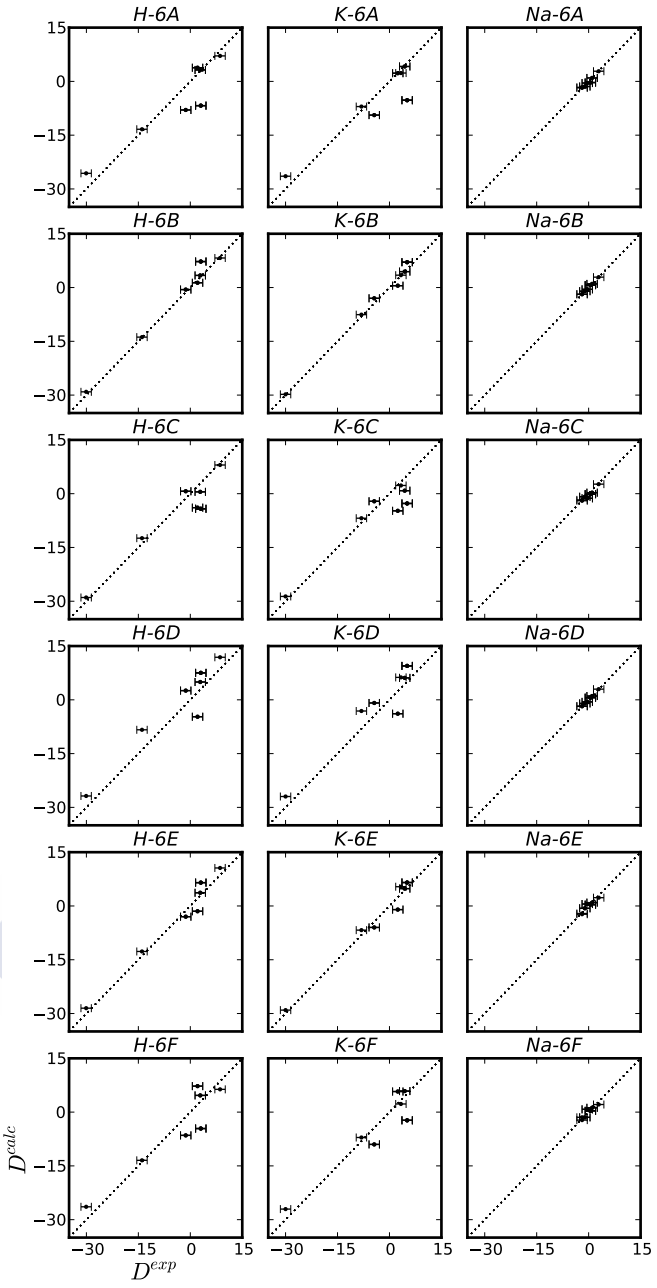


FIGURE 6.12: Plots of (D^{exp}) *vs.* back-calculated (D^{calc}) $^1D_{CH}$ RDCs resulting from the fit of all possible 2-membered **6** ensembles; experimental data from *H*-gel, *K*-gel and *Na*-gel. The error bars are set to 1.5 Hz.

6.3.3 RDC Analysis. One-bond RDC fit to all possible pair-ensembles

As was already introduced in Chapter 1 and in Chapter 3, the analysis of flexible small molecules by RDC fit is not straightforward given that, in principle, the orientational probability depends on the conformation. If this is the case, an independent

alignment tensor must be fitted to every structure, meaning that 5 independent RDCs are needed for every member of the ensemble. Easily, this represents an unfeasible number of experimental RDCs to fit. In the case of short-amplitude conformational changes —the conformational change does not dramatically change the global shape of the molecule— the single-tensor approximation can be used.[47]

For the determination of a common reference frame for all conformations, DFT structures were superimposed by minimizing the RMSD between the cartesian coordinates of all heavy atoms.[125] Besides, other superimposition schemes were tested (all heavy atoms, exclusion of the methyl group; only the aryl ring atoms), leading to nearly equivalent results.

The use of the single-tensor approximation permits the simultaneous determination of conformer populations *and* alignment tensor components. These values are used to back-calculate the RDCs, which can be compared with the experimental ones, e.g. graphically or with an appropriate merit function.[51] As seven independent one-bond RDCs are available, it is possible to fit two-membered ensembles ($5 + (2 - 1) = 6$ unknowns) with confidence. Furthermore, two are the lowest energy conformers (with a reasonable energy difference), and previous investigations on the accessible conformations for this kind of molecule points to a two-state model presenting two half-chair structures. Nevertheless, as a decision relying on DFT derived energies alone should not be taken, all possible combinations of 2 conformations out of the 6 were taken into account. The 15 two-membered ensembles were fitted to the RDC data collected in *H*- and *K*-gels and the corresponding Q_C factors and populations were obtained (Table 6.6 and Table 6.7).

Plots of experimental (D^{exp}) *vs.* back-calculated (D^{calc}) $^1D_{CH}$ show that back-calculated RDC are close to experimental values (within 1.5 Hz error) in the ensembles with Q_C lower than 0.68. With the other ensembles, at least one back-calculated value deviated more than 1.5 Hz. The choice of the value 1.5 Hz is somehow arbitrary, or it can be based on the experimental error. The average error for $^1D_{CH}$ is < 1 Hz (Table 6.4), but these errors are bound to the experimental uncertainty on the determination of the center of the corresponding peak. There are other sources of error, e.g. as strong-coupling, that are not taken into account. In this situation, we choose to use a synthetic, homogeneous error.

The fit of two-membered ensembles to RDCs obtained from *H*-gel (Table 6.6 and Figure 6.14) resulted in seven ensembles (out of fifteen) with $Q_C < 0.1$ (Figure 6.13a)). The ensemble containing the two lowest-energy crown-chair conformers **6A**+**6B** furnished a $Q_C = 0.043$, but other three ensembles furnished a Q_C even smaller (Table 6.6), namely **6E**+**6F**, **6C**+**6E** and **6A**+**6E** (from higher to lower Q_C). Note that fits of some

TABLE 6.6: Fit of the experimental one-bond RDCs to the 15 possible two-membered ensembles (*H*-gel). Conformer populations ($p_1 : p_2$) and Q_C factors are shown.

Ensemble	Q_C	$p_1 : p_2$	Ensemble	Q_C	$p_1 : p_2$
6A+**6B**	0.043	32:68	**6B**+**6F**	0.082	57:43
6A+**6C**	0.228	74:26	**6C**+**6D**	0.188	53:47
6A+**6D**	0.082	54:46	**6C**+**6E**	0.025	56:44
6A+**6E**	0.022	44:56	**6C**+**6F**	0.183	58:42
6A+**6F**	0.342	0:100	**6D**+**6E**	0.179	0:100
6B+**6C**	0.083	50:50	**6D**+**6F**	0.126	42:58
6B+**6D**	0.132	100:0	**6E**+**6F**	0.042	48:52
6B+**6E**	0.132	100:0			

ensembles converged to 100% population of one of the conformers. These *pseudo*-ensembles can be discarded, as we have already decided —based on the 1D-NOESY and $^3J_{HH}$— that at least two conformers involving axial and an equatorial disposition of the methyl group should exist in solution. Furthermore, these ensembles usually fit worse to the experimental RDCs (i.e. give high Q_C values).

TABLE 6.7: Fit of the experimental one-bond RDCs to the 15 possible two-membered ensembles (*K*-gel). Conformer populations ($p_1 : p_2$) and Q_C factors are shown.

Ensemble	Q_C	$p_1 : p_2$	Ensemble	Q_C	$p_1 : p_2$
6A+**6B**	0.059	41:59	**6B**+**6F**	0.069	90:10
6A+**6C**	0.296	76:24	**6C**+**6D**	0.240	51:49
6A+**6D**	0.142	58:42	**6C**+**6E**	0.067	58:42
6A+**6E**	0.095	45:55	**6C**+**6F**	0.246	48:52
6A+**6F**	0.314	0:100	**6D**+**6E**	0.153	0:100
6B+**6C**	0.098	100:0	**6D**+**6F**	0.155	41:59
6B+**6D**	0.098	100:0	**6E**+**6F**	0.003	48:52
6B+**6E**	0.098	100:0			

In the case of RDCs obtained from the *K*-gel (Table 6.7 and Figure 6.15), fit resulted in seven ensembles (out of fifteen) with low Q_C (Figure 6.13b)). The low energy ensemble (**6A**+**6B**) furnished a $Q_C = 0.059$, which is slightly higher than that obtained in the fit of *H*-gel data. In this case, only one ensemble composed by the highest energy conformers **6E**+**6F** furnished a lower Q_C of 0.003, which is surprisingly low.

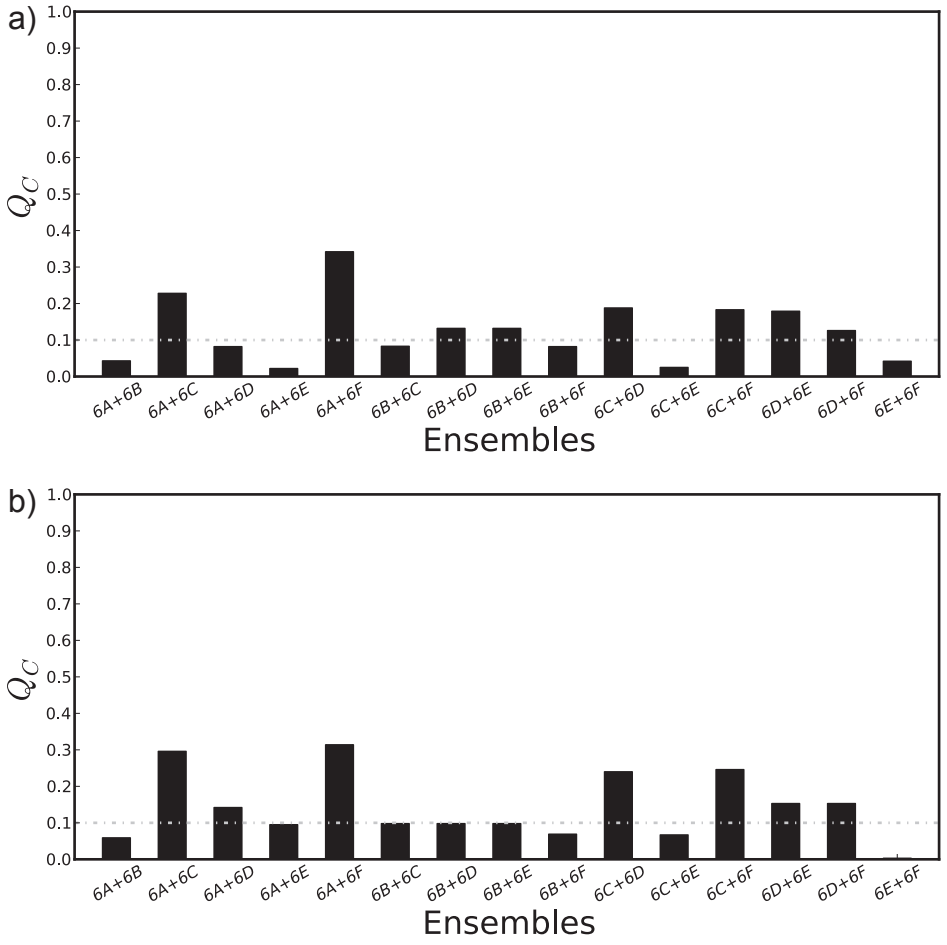


FIGURE 6.13: $^1D_{CH}$ RDC Q_C from two-membered ensembles fit. a) $^1D_{CH}$ from *H*-gel, data is shown in Table 6.6; b) $^1D_{CH}$ from *K*-gel, data is shown in Table 6.7.

Recalling the results with *H*-gel data, the two crown-chair forms **6A**+**6B** resulted in a low Q_C factor of 0.043 (Table 6.6), with a population excess favoring the crystallographic structure **6B** (32 : 68 ratio).

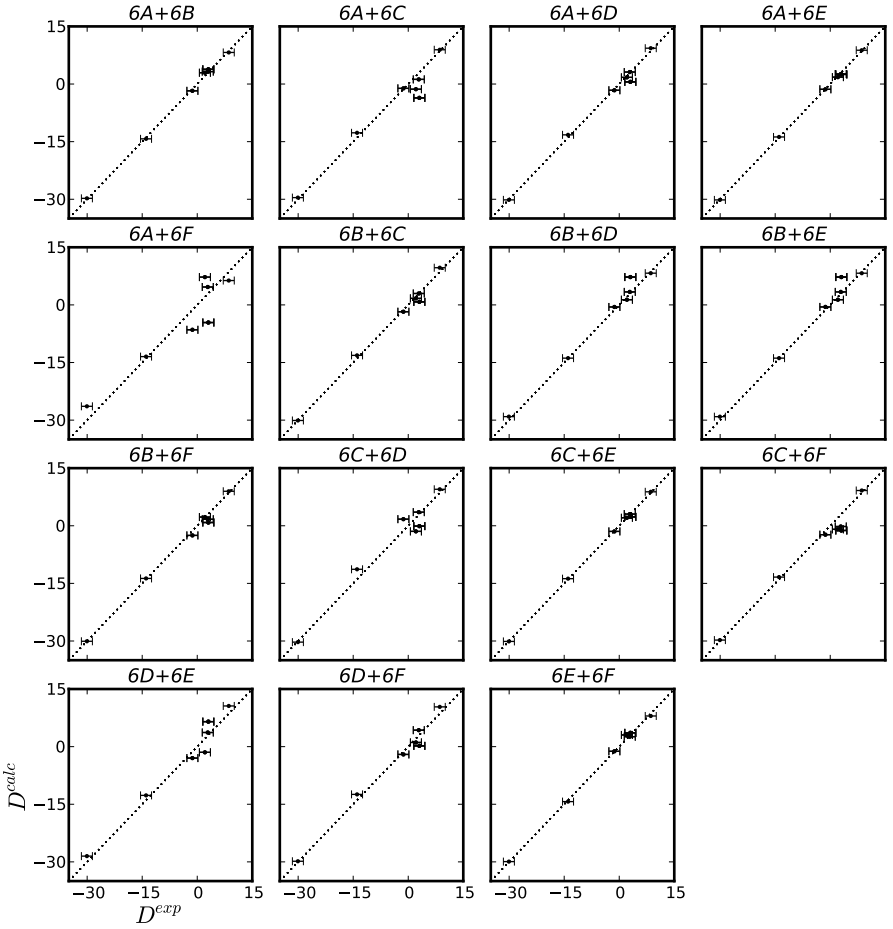


FIGURE 6.14: Plots of experimental (D^{exp}) *vs.* back-calculated (D^{calc}) $^1D_{\mathrm{CH}}$ RDCs resulting from the fit of all possible 2-membered **6** ensembles with *H*-gel experimental data. The error bars are set to 1.5 Hz.

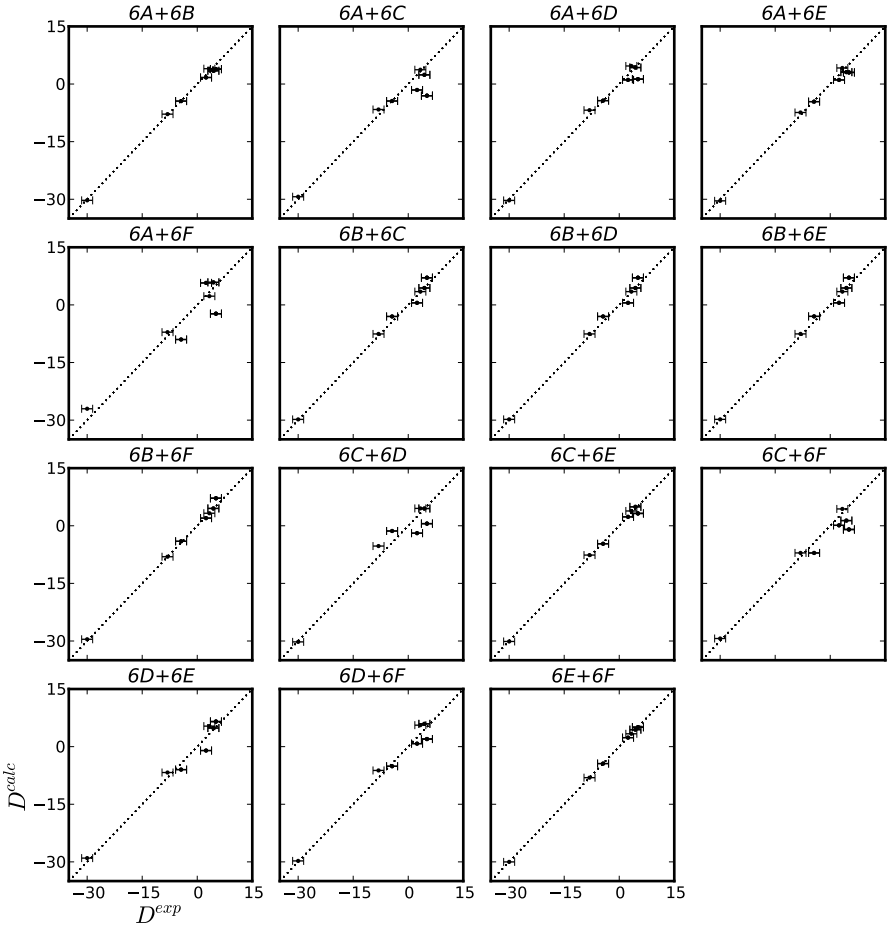


FIGURE 6.15: Plots of experimental (D^{exp}) *vs.* back-calculated (D^{calc}) $^1D_{\mathrm{CH}}$ RDCs resulting from the fit of all possible 2-membered **6** ensembles with *K*-gel experimental data. The error bars are set to 1.5 Hz.

In summary, fit of one-bond RDCs to two-membered ensembles does not provide a convincing answer to the equilibrium composition of lorcaserin in solution. Some ensembles with good fit (low Q_C) contain conformers with high predicted energies. Furthermore, the answer is not unique for two reasons: (*i*) there are several ensembles with low Q_C values within a reasonable range (e.g. < 0.7), and (*ii*) the preferred ensemble is not the same in both *H*- and *K*-gels. In the next section we will show how the addition of other NMR restraints can solve the degeneracy on the solution of the conformational equilibrium of lorcaserin. Additionally, in Section 6.5.2 we will demonstrate how the only inclusion of long-range couplings solves the degeneracy of the solution.

As the fit using the short-range RDC data obtained from the *H*-gel (Table 6.6) resulted in lower quality factors than the fit using *K*-gel data (Table 6.7), long-range data was acquired only the former sample and the rest of the analysis is focused on it.

6.4 RDCs predicted populations in combination with other NMR restraints provide a clear view of the conformational preference of 6 in solution

Analysis of one-bond RDC couplings was unable to select the correct ensemble due to solution degeneracy. In this section, we combine these RDC fits with the NOE discussed in Section 6.2.2 and $^3J_{HH}$ couplings.

With the objective of gaining further conformational restraints for **6**, homonuclear couplings were extracted from 1D-^{1}H spectra. Extensive resonance overlap of several resonances precluded the extraction of their $^3J_{HH}$ couplings. In recent years, following the recommendations of Tolman and co-workers, many groups are extracting $^3J_{HH}$ couplings in cases of severe signal overlap (mainly in aliphatic regions) from the clean subspectra provided by high-resolution F_2-coupled HSQC experiments.[162,230,237] The use of such experimental approach provided clean subspectra for the C5, C1, C4 and C2 carbon frequencies, from which all vicinal $^3J_{HH}$ couplings were extracted (Figure 6.16).

Once experimental $^3J_{HH}$ coupling constants were extracted, conformers **6A-6F** were used for obtaining the *Karplus* predicted couplings with the Haasnoot-Altona equation as implemented in Mspin.[32,42,43]

Table 6.8: Experimental and calculated vicinal $^3J_{HH}$ (Hz) scalar couplings of lorcaserin. Coupling values of each lorcaserin conformer were computed with the Haasnoot-Altona equation.

^{1}H	Exp.	**6A**	**6B**	**6C**	**6D**	**6E**	**6F**
H1–H2b(α)	9.1	3.9	9.9	4.4	8.9	11.1	9.5
H1–H2a(β)	1.0	1.7	1.4	1.4	1.2	6.7	8.3
H5α–H4b(α)	1.7	2.2	0.6	0.5	1.4	9.3	6.3
H5α–H4a(β)	7.9	12.1	5.5	6.2	11.4	1.1	12.0
H5β–H4a(β)	1.9	0.8	1.7	1.3	0.5	10.5	4.1
H5β–H4b(α)	9.6	5.0	11.7	11.4	6.2	7.4	1.1

Determination of configurations α and β is explained later in this Chapter.

Experimental $^3J_{HH}$ were fitted to the two-membered ensembles using the populations determined with the RDC fits (Table 6.6). Ensemble averaged $^3J_{HH}$ are calculated with the expression

$$\langle J_k \rangle = \sum_i^N (p_i J_{ki}), \tag{6.2}$$

where N is the total number of conformers, p_i is the molar fraction of the i^{th} conformer, and J_{ki} is the calculated k^{th} coupling constant in the i^{th} conformer (Table 6.6).

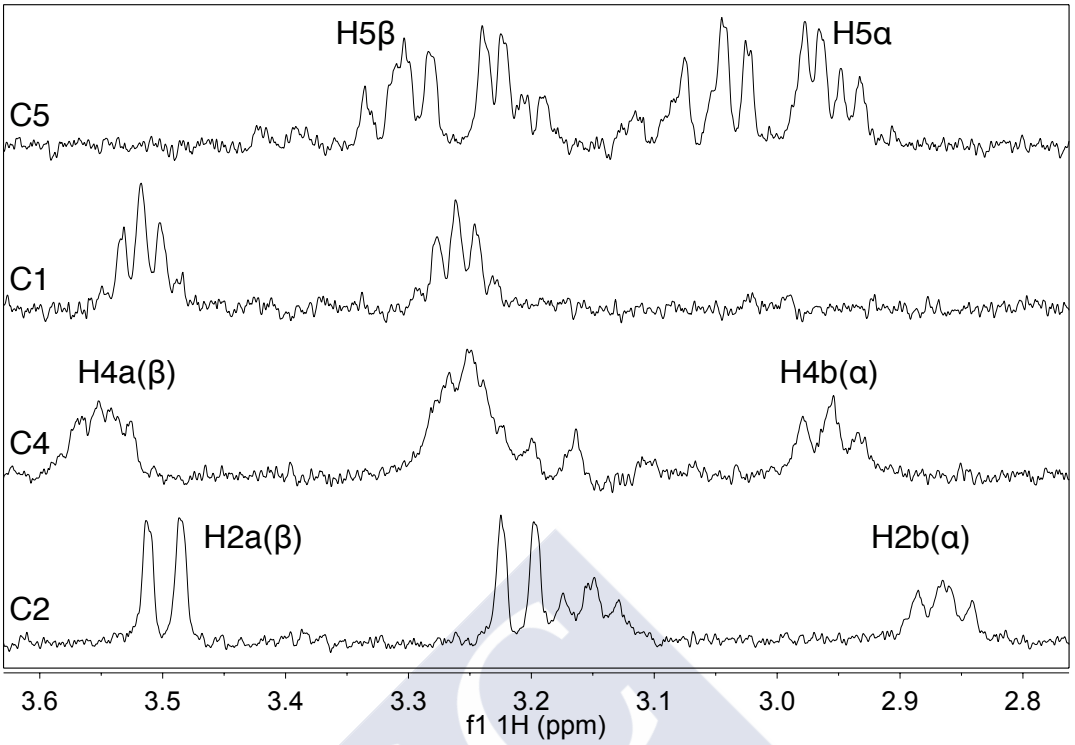


FIGURE 6.16: F_2 projections of the high-resolution F_2-coupled HSQC of Lorcaserin. Each trace is labeled with the corresponding ^{13}C resonance . Vicinal $^3J_{HH}$ couplings can be easily extracted from the clean 1D ^1H subspectra shown. Stereochemical assignment of the geminal protons was determined by other means and is shown here for the sake of clarity.

Later, *RMSD* values between the experimental and *Karplus*-predicted vicinal $^3J_{HH}$ couplings on the two-membered ensembles were computed, using the populations derived from the RDC fit (Table 6.6) to weight the computed couplings, and including couplings involving the H2 and H4 protons permuting the two possible assignments (Table 6.9).

RMSD between experimental and ensemble averaged $^3J_{HH}$ are shown in Table 6.9. The assignment of H5 protons is not swapped with the other methylenes of lorcaserin as they have been already assigned by means of NOE correlations (see Section 6.2.2). The best fit by far corresponds to the assignment H2a/β, H2b/α and H4a/β, H4b/α resulting in a *RMSD* of 0.6 Hz for the **6A**+**6B** ensemble (Table 6.9[a]). Accordingly, this assignment furnished the best fit to the rest of the ensembles, save for the **6E**+**6F** ensemble, which fits best to the inverted assignment of H4 protons (H4a/α, H4b/β, Table 6.9[c]), and **6E**+**6F** ensemble for which the assignment of H4 protons is not possible (Table 6.9, compare [a]-[c] and [b]-[d]).

TABLE 6.9: Fit of the experimental and ensemble-averaged Haasnoot-Altona $^3J_{HH}$ couplings. Populations were derived from $^1D_{CH}$ RDC fit (*H*-gel, see Table 6.6). Root mean square deviation (*RMSD*) is given in Hz.

Ensemble	$RMSD^{[a]}$	$RMSD^{[b]}$	$RMSD^{[c]}$	$RMSD^{[d]}$
6A+**6B**	0.6	4.2	5.9	7.2
6A+**6C**	2.7	3.7	6.3	6.8
6A+**6D**	2.6	4.4	6.3	7.3
6A+**6E**	3.3	4.5	3.3	4.5
6A+**6F**	5.3	5.6	5.7	6.0
6B+**6C**	1.5	4.2	6.2	7.4
6B+**6D**	1.4	5.0	6.2	7.8
6B+**6E**	1.4	5.0	6.2	7.8
6B+**6F**	1.9	4.2	5.1	6.3
6C+**6D**	1.2	4.0	6.2	7.2
6C+**6E**	2.7	4.2	4.2	5.2
6C+**6F**	2.1	3.2	5.3	5.8
6D+**6E**	6.1	7.0	3.4	4.9
6D+**6F**	3.7	4.9	5.5	6.4
6E+**6F**	4.8	5.5	3.7	4.6

Note that the experimental H4α/β resonances are no stereoasigned at this point of the analysis, see text.
[a] The RMSD was calculated assigning H2a resonance to H2β, and H2b to 2α; and H4a resonance to H4β, and H4b to 4α.
[b] The RMSD was calculated assigning H2a resonance to H2α, and H2b to 2β; and H4a resonance to H4β, and H4b to 4α.
[c] The RMSD was calculated assigning H2a resonance to H2β, and H2b to 2α; and H4a resonance to H4α, and H4b to 4β.
[d] The RMSD was calculated assigning H2a resonance to H2α, and H2b to 2β; and H4a resonance to H4α, and H4b to 4β.

We saw in the previous section that 7 ensembles gave good fits to one-bond RDCs (i.e. they had low Q_C values), among these, the **6A**+**6B**, **6A**+**6E**, **6C**+**6E** and **6E**+**6F** have the lowest Q_C and are not distinguishable. Two of them furnished lower Q_C (fitted better to experimental RDCs) than the lowest energy ensemble **6A**+**6B**, namely **6A**+**6E** and **6E**+**6F**. However, ensembles **6A**+**6E** and **6E**+**6F** fit poorly to $^3J_{HH}$, as evidenced by the high RMSD values of 3.3 and 4.8 Hz, respectively, which are clearly larger than that of ensemble **6A**+**6B** (RMSD = 0.6 Hz). Ensemble **6C**+**6E** has a bit smaller $^3J_{HH}$ RMSD (2.7 Hz). Fit to $^3J_{HH}$ values is slightly better than the one of ensembles **6A**+**6E** and **6E**+**6F**, but still worse than that of **6A**+**6B**.

Actually, the combination of RDC fit obtained conformer population with the *Karplus* computation of ensemble-averaged $^3J_{\mathrm{HH}}$ couplings is sufficient to select one ensemble that can fulfill all experimental NMR structural restraints (Figure 6.17).

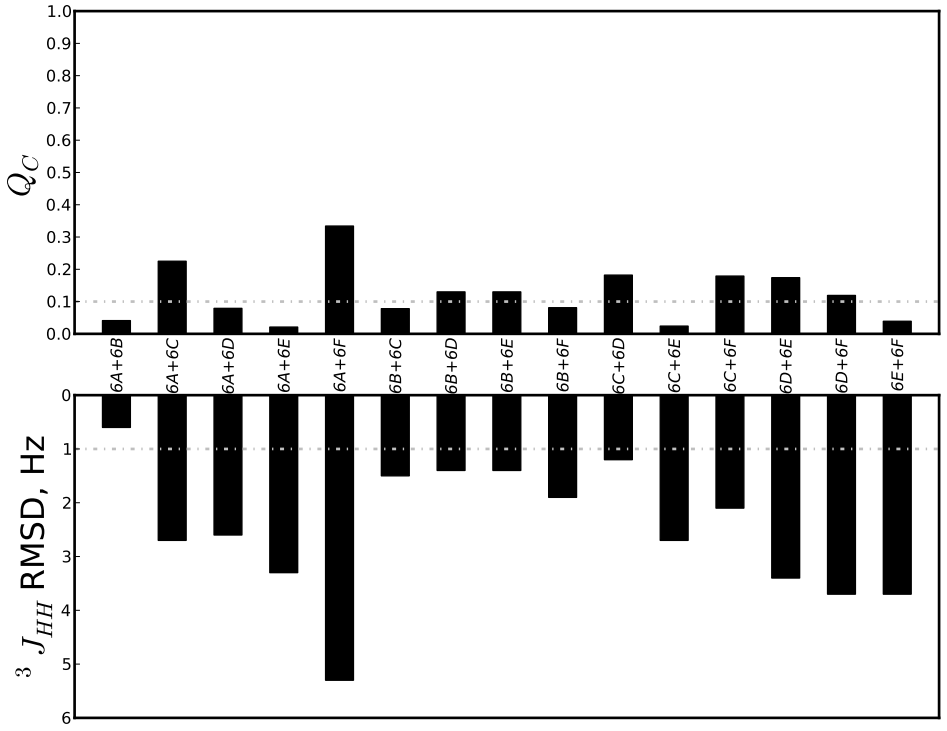


FIGURE 6.17: RDC quality factor Q_C and $^3J_{\mathrm{HH}}$ RMSD values from comparison of experimental and computed data.
All the assignments of H2 and H4 protons were considered (Table 6.9). Note that every ensemble *RMSD* is the lowest one, despite the "correct assignment" of the diasterotopic C4 and C2 protons determined from the best-fit ensemble **6A**+**6B**.

The lowest RDC Q_C factors corresponded to the **6A**+**6B** (both crown-chair), the **6A**+**6E** (crown-chair/boat), the **6C**+**6E** (twist / boat), and the **6E**+**6F** (boat/twist-boat) ensembles. The fact that **6A**+**6B** furnished the best RDC / $^3J_{\mathrm{HH}}$ combined fit is in strong agreement with the DFT computations where both crown-chair conformations are very similar in energy and appreciably more stable than all other conformations ($\Delta\Delta G > 2$ kcal/mol).

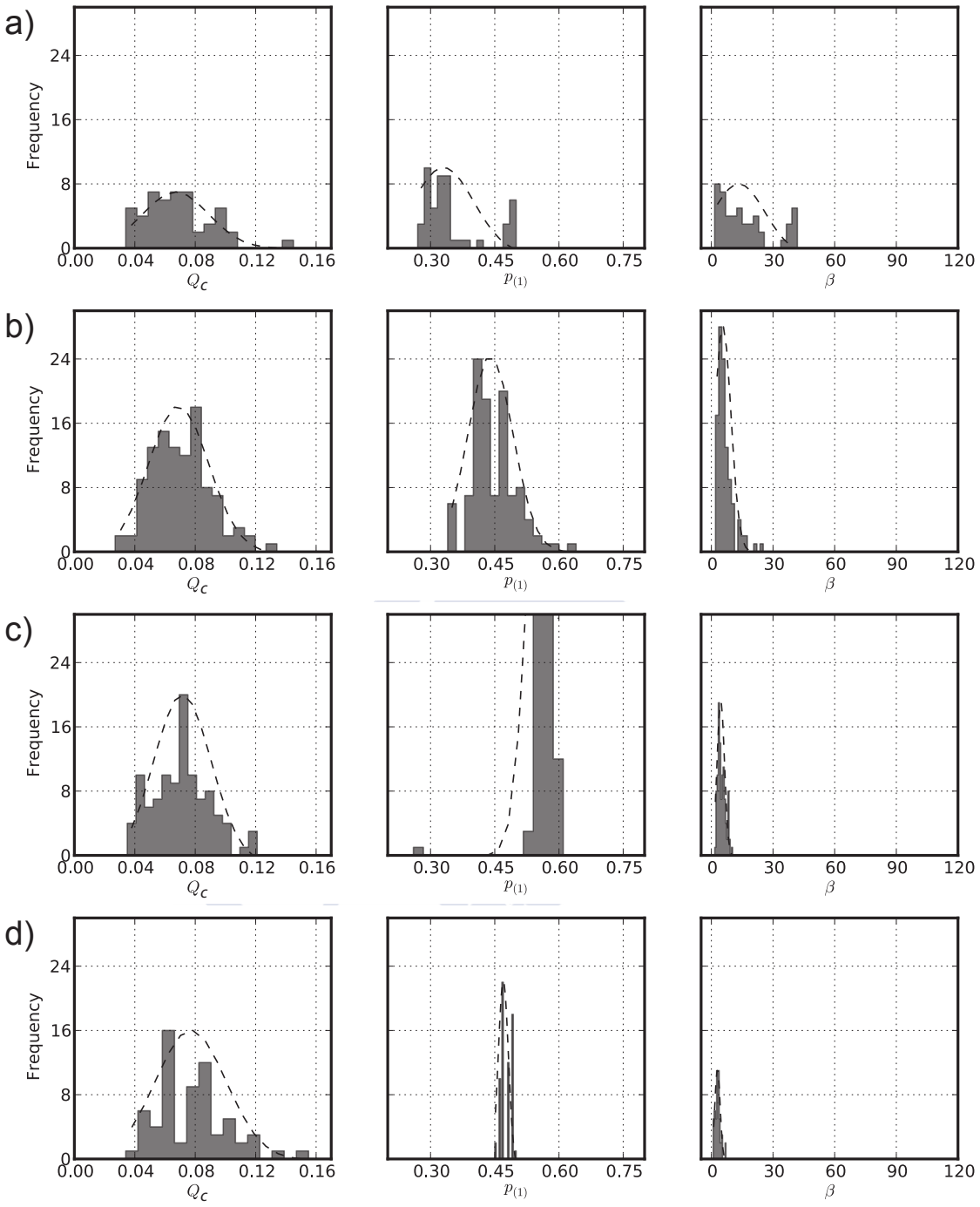


FIGURE 6.18: Summary plots of Monte Carlo-based error estimation of fits of $^1D_{\mathrm{CH}}$ RDCs to lorcaserin two-membered ensembles. The histograms show, in normalized frequency scale; left, the distribution of Q_{ang}; center, the distribution of p_1; and, right, the distribution of the generalized angle β respect to the alignment tensor determined with the experimental data. The dashed line over the distributions represents the normalized Probability Distribution Function (centered in the median). a) **6A**+**6B**; b) **6A**+**6E**; c) **6C**+**6E**; b) **6E**+**6F**. Statistics of the bootstrap analysis can be seen in Table 6.10.

To test the impact of experimental errors in the computed populations, the ensembles

were further studied with a bootstrapping analysis with a Monte Carlo filter. We generated 512 *decoys* allowing a conservative experimental error of 1.5 Hz. The Monte Carlo filter was set to the generator standard deviation (1.5 Hz). The histograms with the fit statistics of the ensembles that passed such a filter are shown in Figure 6.18. Only four ensembles passed the Monte Carlo filter, namely **6A**+**6B**, **6A**+**6E**,**6C**+**6E** and **6E**+**6F**. The final population for the "correct" **6A**+**6B** ensemble does not differ very much from that determined from the experimental data. As can be clearly seen from the Q_C histograms (Figure 6.18), all these ensembles are indistinguishable in terms of fit quality factor.

TABLE 6.10: Statistics of Monte Carlo error estimate of $^1D_{CH}$ RDCs fitted to all two-membered ensembles of lorcaserin. Only the four ensembles that passed the Monte Carlo filter are shown.

		Q_C		p_1		β	
value	accept	μ	σ	μ	σ	μ	σ
6A+**6B**	49	0.069	0.022	0.35	0.07	16	12
6A+**6E**	107	0.071	0.020	0.45	0.05	6	3
6C+**6E**	104	0.070	0.019	0.56	0.03	5	1
6E+**6F**	65	0.078	0.023	0.48	0.01	3	1

μ represents the arithmetic mean
σ is the biased standard deviation estimator.
p_1 represents the population (molar fraction) of the first conformer of the corresponding ensemble.
β is the generalized rotation (in °) of the tensors from that determined from the experimental data.

Fit of one-bond RDCs to the candidate lorcaserin conformers did not suffice to determine the conformational preference of lorcaserin in solution. Nevertheless, RDC fits provided the populations of every conformer in the ensembles, allowing the computation of ensemble-averaged Haasnoot-Altona vicinal $^3J_{HH}$ couplings, thus making possible to determine the pair **6A**+**6B** as the correct one, with a population ratio of 32 : 68. Once the conformational preference of lorcaserin is determined, we can now ensure the assignment of the H2 and H4 protons. With the given population, the assignment H2a = H2β / H4a = H4β fits clearly better ($RMSD = 0.6$) than the others ($RMSD > 4$, see Table 6.9). Monte Carlo computations confirmed the population stability for the **6A**+**6B** (35 : 65, $\pm$7).

6.5 Introduction of long-range RDC improves analysis of equilibrium exchange

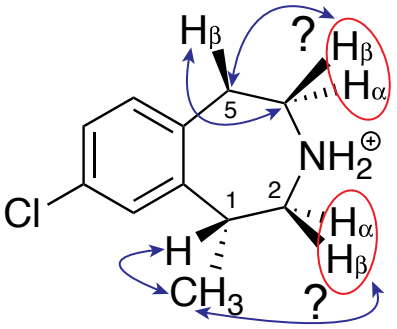


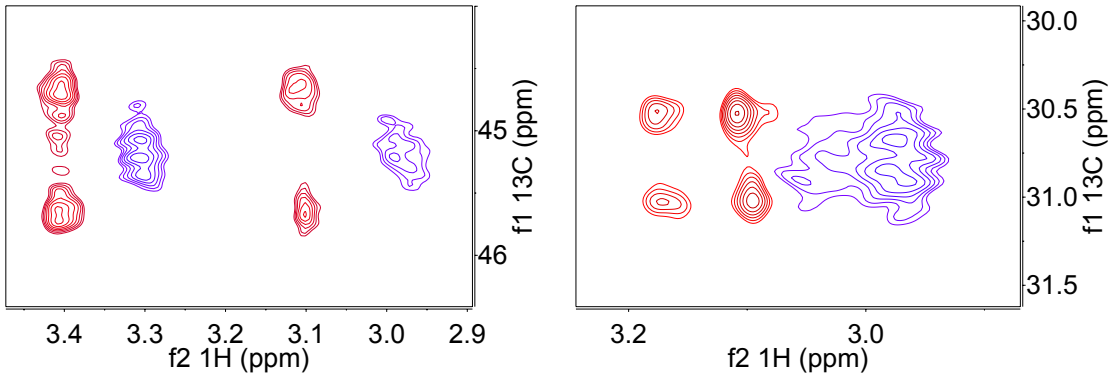
FIGURE 6.19: Long-range RDCs measured in lorcaserin. Coupled C–H pairs are indicated by blue arrows. Red ovals indicate the uncertainty on the stereospecific asignment of C2 and C4 protons.

The previous section described that one-bond RDC did not suffice to solve the question on their own and $^{3}J_{HH}$ / NOE data were also needed. This illustrates the importance of having enough experimental data. In principle, only five experimental couplings are needed to determine the alignment tensor, as it has five independent components and one more is usually needed to discriminate between alternative structures. Besides, at least another coupling is needed if there is another unknown (e.g. a population in the case of a mixture of two conformers). This is true at least in mathematical terms, i.e. there are as many observables as unknowns and the system of equations is determined. However, there may be cases, as the example just described, where more than one answer fits the experimental data. This is most likely due to the fact that five or six couplings do not suffice to sample enough bond orientations in the 3D space. Therefore, more experimental couplings are desirable to better define the system, preferably with vectors that sample as many directions —and as many moieties of the molecule— as possible.

In this section, long-range RDC will be used to address the same question. It will be described how this RDC data set (one-bond and long-range) suffices to determine the populations in equilibrium with no need to resort to other data such as HH vicinal couplings or NOE.

6.5.1 Extraction of long-range $^{n}D_{CH}$ couplings

Long-range C–H couplings were determined from a set of selective *J*-Scaled (SJS) HSQC experiments.[182] Spectra were recorded in isotropic ($^{n}J_{CH}$) and anisotropic ($^{n}T_{CH}$) conditions (Figure 6.20 and Figure 6.22 and Table 6.4). RDCs were determined as the difference $^{n}D_{CH} = {}^{n}T_{CH} - {}^{n}J_{CH}$. Four $^{n}D_{CH}$ couplings were measured from the *H*-gel sample only (Table 6.11).



(a) C4–H5β (b) C5–H4a

FIGURE 6.20: Lorcaserin SJS-HSQC inverting H5β and H4a, respectively showing the long range couplings between C4–H5β and C4–H4a. Red, isotropic; blue, *H*-gel.

The individual long-range couplings between C4–H5β (Figure 6.20*(a)*) and C4–H4a (Figure 6.20*(b)*) were extracted from two different SJS-HSQC experiments.

As mentioned in Chapter 5[182] the total coupling is compatible with more than one possible dipolar coupling. Usually, one of them has an unfeasible magnitude, but it can happen that the two possibilities have similar magnitude and, therefore, both must be taken into account. In Figure 6.21, we show posible combinations of ${}^{n}J_{\mathrm{CH}}$ and ${}^{n}D_{\mathrm{CH}}$ values.

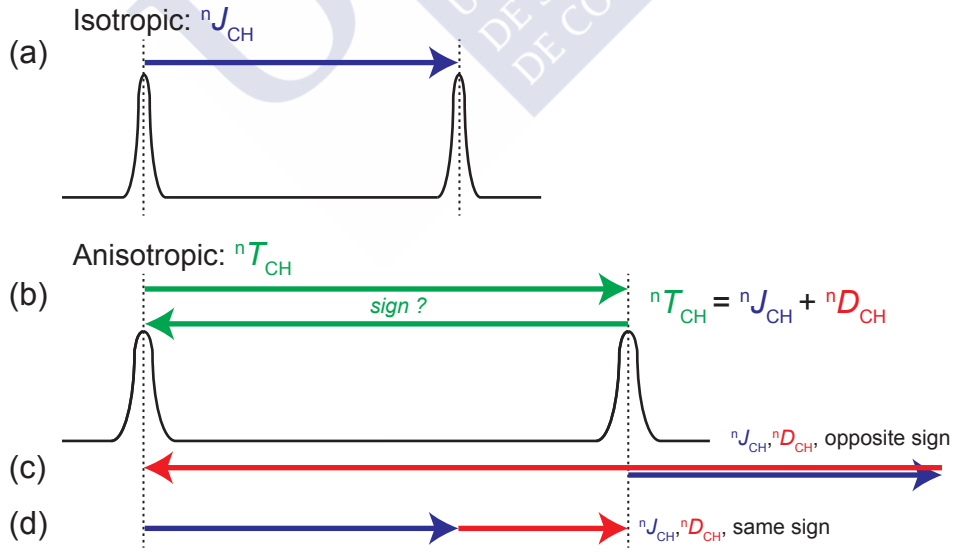


FIGURE 6.21: Schematic representation of different ${}^{n}D_{\mathrm{CH}}$ compatible with the same ${}^{n}T_{\mathrm{CH}}$. The arrow head indicates the sign; right, positive; left, negative. (a) Isotropic spectrum, ${}^{n}J_{\mathrm{CH}}$; (b) anisotropic spectrum, ${}^{n}T_{\mathrm{CH}}$ can have the (c) opposite sign as ${}^{n}J_{\mathrm{CH}}$ or (d) the same sign as ${}^{n}J_{\mathrm{CH}}$.

Additionally, two different long-range couplings involving C10 (Me) were observed (Figure 6.22). In principle, care should be taken in no inverting at the same time two different protons H_A and $H_{A'}$ coupled to the same carbon C_B. Nevertheless, it is still possible to extract the individual couplings from the doublet of doublets pattern with enough resolution.

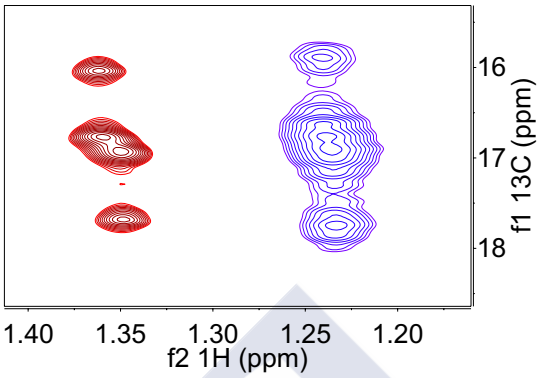


FIGURE 6.22: Expansion of the SJS-HSQC resulting from simultaneous inversion of H1 and H2a. Expansion of the C10 multiplet is shown. red, isotropic; blue, *H*-gel. C10 appears as doublet of doublets in F_1 due to coupling to H1 and H2a, see text.

In Figure 6.23, a simulated version of Figure 6.22 is presented to make clear the splitting pattern originating the observed peak pattern. In this case, one of the couplings is larger than the other, facilitating its assignment to the $^2J_{C10H1}$ and the smaller one to the $^3J_{C10H2a}$.

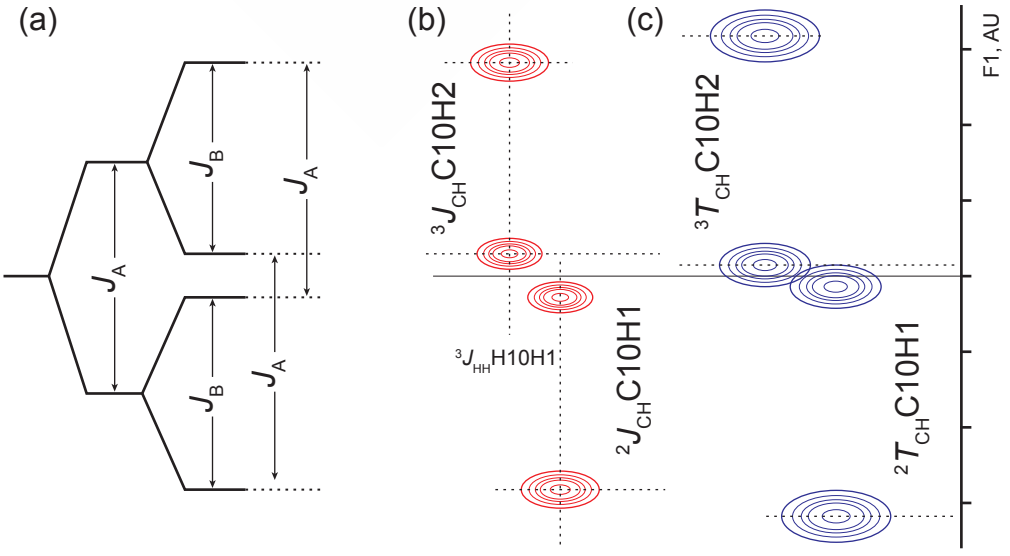
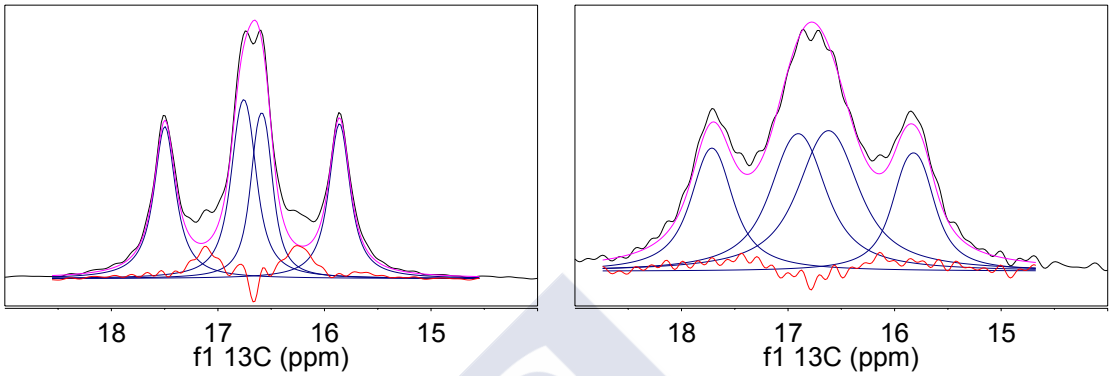


FIGURE 6.23: Simulated C10–H1 / C10–H2a SJS-HSQC peak. (a), splittings due to the two different $^nJ_{CH}$; (b), Isotropic; (c) Anisotropic.

The extraction of the couplings involving C10 was impaired by the overlap of the lines of the multiplet (Figure 6.22). Precise analysis of the peaks was facilitated by deconvolution of the peaks with MestReNova.[42] Briefly, the traces of the C10 peak were summed separately for the isotropic and anisotropic experiments. Then the projection spectrum was extracted and deconvoluted using default parameters (Figure 6.24).



(a) Isotropic (b) Anisotropic

FIGURE 6.24: Deconvolution of the C10 multiplet. The projection spectrum of the C10 multiplet in the SJS-HSQC inverting the H1/H2a resonances was deconvoluted with MestreNova.

Initially, the long range couplings C5–H4a and C10-H2a were excluded from the fit because their stereochemical assignment was not known *a priori*.

An issue with long-range couplings is the choice of merit function. The Cornilescu Q_C factor could be used. However, we should be aware that dipolar couplings have a *radial* and an *angular* component. It is worth noting that we are mainly interested in the angular component, as it contains the orientational information of each vector relative to the alignment tensor. The dependence of dipolar coupling on r^{-3} has the undesirable effect that the angular information of long-range couplings is scaled down relative to one-bond couplings. Therefore, another merit function that compensates this scaling by distance, thus giving equal weight to the orientational information contained in the ${}^1D_{\mathrm{CH}}$ and ${}^nD_{\mathrm{CH}}$ couplings, is needed. This compensation is achieved by defining the *angular quality factor* Q_{ang}, where each ${}^nD_{\mathrm{CH}}$ is multiplied by r^3 (i.e. D_j is substituted by its reduced form d_j, see Chapter 1, Section 1.4.1).[2]

Q_{ang} is the angular version of the well known Cornilescu Q_C[44] factor (Equation (1.70)), and is defined as follows

$$Q_{ang} = \sqrt{\frac{\sum(d_j^{calc} - d_j^{exp})^2}{\sum(d_j^{exp})^2}}. \tag{6.3}$$

[2]Dr. Ad Bax and Dr. Roberto R. Gil, personal communication.

TABLE 6.11: Scalar ($^{n}J_{CH}$) and dipolar ($^{n}D_{CH}$) long-range C–H couplings of lorcaserin, in Hz. The RDCs were acquired from the *H*-gel sample only.

Vector	$^{n}J_{CH}$	n	$\lvert\, ^{n}T_{CH} \,\rvert$	$^{n}D_{CH}$[1]	$^{n}D_{CH}$[2]	$\lvert\, ^{n}T_{CH} \,\rvert$	$^{n}D_{CH}$[3]	$^{n}D_{CH}$[4]	Error[a]
C10–H1	−5.74	2	6.97	−1.23	+12.71	4.91	+0.83	+10.65	0.26
C4–H5β[b]	−6.49	2	1.22	+5.27	+7.71	–	–	–	0.43
C5–H4a[c]	−3.17	2	1.28	+1.82	+4.38	–	–	–	0.26
C10–H2a[c]	+4.58	3	4.91	+0.33	−9.49	6.97	+2.39	−11.55	0.15

[a] Error was estimated as $\kappa \times LW/SN$ in the H-gel spectrum, where κ is the J-evolution scaling factor, LW the linewidth in Hz, and SN the signal-to-noise ratio.
[b] The stereochemistry β of H5 was assigned based on NOE (Section 6.2.2).
[c] The configuration of C4 and C2 protons is yet unknown. Letters a and b refer to the low-field and high-field resonances, respectively.

Note that Q_{ang} is equivalent to Q_C when all internuclear distances r_{IS} are equal, as happens —in good approximation— when only $^1D_{CH}$ RDC are used.

6.5.1.1 DFT computed long-range $^nJ_{CH}$ resemble the experimental ones

To ensure the accuracy of the extracted long-range $^nJ_{CH}$ couplings, we compared the experimental values with the DFT-computed ones (Table 6.12). C–H J couplings were computed using the PBE0[189] functional in combination with the specialized pcJ-1 basis set[238] on the M052X/6-31+G** previous geometries.

TABLE 6.12: DFT-computed $^nJ_{CH}$.

vector	experimental	**6A**	**6B**	**1A+6B**[a]
C10–H1	−5.74	−6.54	−6.03	−6.21
C4–H5α	−6.49	−7.34	−1.91	−3.81
C4–H5β	−6.49	−2.25	−7.46	−5.64
C5–H4α	−3.17	−3.02	−3.24	−3.16
C5–H4β	−3.17	−3.35	−2.97	−3.10
C10–H2α	4.58	1.36	0.14	0.57
C10–H2β	4.58	6.91	2.56	4.08

[a] Ensemble averaged couplings, taking into account the $^1D_{CH}$ fit-derived populations (35 : 65, Table 6.10).

DFT-computed couplings corroborated the assignment of H2a proton (H2a = H2β ΔJ_{C10H2a} = 0.50 Hz / H2a = H2α ΔJ_{C10H2a} > 4 Hz), and that of H5a to H5β (see Table 6.12). Note however, the distinction between the possible assignments of H4 protons is not possible by comparison of the DFT-computed coupling of each assignment, because both of them furnished very close values. Additionally, it can be easily detected that the DFT-derived J couplings do not lie in the determined experimental error range (1σ, see Table 6.11), supporting our decision to take into account not the experimental error but a larger homogeneous one.

6.5.2 Long-range RDC fits to ensembles

First, we back-calculated the long-range couplings from the $^1D_{CH}$ tensor. In Figure 6.25 the fit between the initially back-calculated RDCs with the final determined values is shown. We will show in the next how by a fit-and-check procedure the long-range RDCs could be assigned.

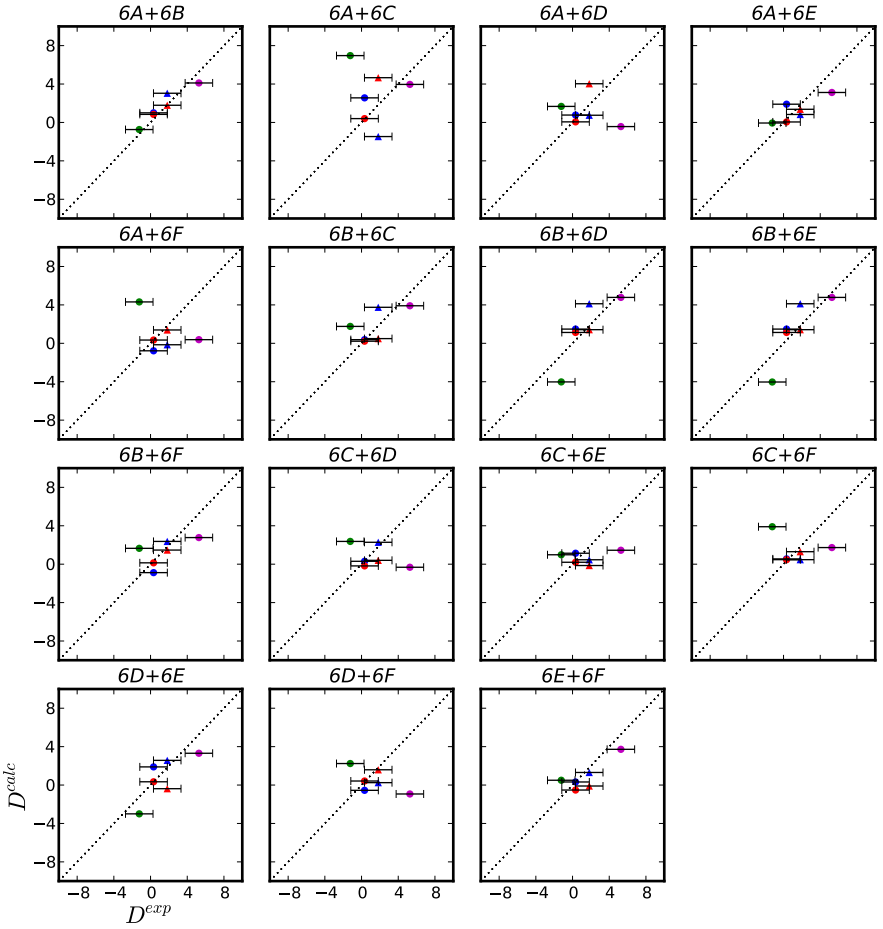


FIGURE 6.25: Plots of back-calculated ${}^{n}D^{calc}$ vs experimental ${}^{n}D^{exp}$. Long-range ${}^{n}D_{\mathrm{CH}}$ were back-calculated using the tensors and population weights determined with one-bond RDC only. The different couplings can be identified as follows; ${}^{2}D_{\mathrm{C10–H1}}$, green circle; ${}^{2}D_{\mathrm{C4–H5}\beta}$ magenta circle; the unassigned methylenic proton couplings are labeled as follows; ${}^{3}D_{\mathrm{C10–H2a}}$ couplings are marked with a circle and ${}^{2}D_{\mathrm{C5–H4a}}$ couplings with a triangle. Markers of couplings back-calculated from Cx to Hα are shaded in blue, while markers of corresponding couplings to Hβ are shaded in red.

6.5.2.1 Analysis of the stereoassigned ${}^{n}D_{\mathrm{CH}}$

All experimental and computational evidences support the existence of lorcaserin as a mixture of conformers in fast-exchange in solution. We consider demonstrated that no single conformer of lorcaserin fulfills alone the experimental restraints, therefore we will fit long-range RDCs directly to ensembles of lorcaserin conformers.

As explained in the previous section, the uncertainty on the determination of ${}^{n}T_{\mathrm{CH}}$ sign forces us to consider two possible ${}^{n}D_{\mathrm{CH}}$ couplings compatible with that measured anisotropic splitting. Additionally, as there is no direct evidence for the assignment of

the largest C10–H1/2a coupling to C1–H1 we should test the opposite assignment as well.

Results of the fit to ensembles are shown in Table 6.13 and Table 6.14. First, only one coupling was introduced, to test its impact on the calculated alignment tensor.

Populations derived from $^{1}D_{\mathrm{CH}}$ fit to the pair ensembles made possible to find one solution compatible with DFT calculations and other NMR evidences such as NOE correlations and vicinal $^{3}J_{\mathrm{HH}}$ couplings. We can assume that the $^{1}D_{\mathrm{CH}}$ tensor is a good-enough description of the aligned lorcaserin. Therefore, the correct long-range RDC assignment will be the one less perturbing the original $^{1}D_{\mathrm{CH}}$ tensor and populations. The perturbation of the $^{1}D_{\mathrm{CH}}$ alignment tensor was evaluated by analyzing Q_{ang}, p_1 and the generalized angle β in fits in which one possible $^{n}D_{\mathrm{CH}}$ was introduced.

An additional requirement was set in form of Monte Carlo filter, in which the back-calculated RDCs were evaluated to resemble the experimental ones within a set error (a very conservative 1.5 Hz error, both for $^{1}D_{\mathrm{CH}}$ and $^{n}D_{\mathrm{CH}}$). It is important to note that only RDC fits to ensemble **6A**+**6B** passed through Monte Carlo filter requirements.

We introduced the long-range RDCs $^{2}D_{\mathrm{C10–H1}}$ and $^{2}D_{\mathrm{C4–H45}\beta}$ one at a time in the fit. The two alternative values of these RDCs (Table 6.11) were used in different fits. The results are shown in Table 6.13 and Table 6.14. As the introduction of the long-range couplings can be seen as a refinement of $^{1}D_{\mathrm{CH}}$ results, the analysis of this step will focus mainly on those ensembles that gave a good fit to the short-range couplings, namely **6A**+**6B**, **6A**+**6D**, **6A**+**6E**, **6B**+**6C**, **6B**+**6F**, **6C**+**6E** and **6E**+**6F**.

Fit of the long-range RDC $^{2}D_{\mathrm{C10–H1}} = -1.23$ Hz (Table 6.13) resulted in lower Q_{ang} (better fit) than that of $^{2}D_{\mathrm{C10–H1}} = +12.71$, which causes a high penalty on the fits save for those ensembles already discarded by $^{1}D_{\mathrm{CH}}$ fit (**6A**+**6C**, **6A**+**6D** and **6A**+**6F**). Three ensembles gave good fit with the former value, i.e. low Q_{ang} $(0.045 - 0.094)$, namely **6A**+**6B**, **6A**+**6E** and **6E**+**6F** (Table 6.13). Populations and tensor orientation are little perturbed compared to the one-bond fit (Figure 6.26). On the contrary, the value +12.71 resulted in poor fit of all 15 ensembles. Moreover, p_1 and β are too different from the determined with one-bond data. This confirms that this value is not compatible with the one-bond RDC. In Figure 6.26, we show a comparison of the obtained tensors (scaled by their eigenvalues) when fitting short-range couplings, and the two candidate values for the $^{2}D_{\mathrm{C10–H1}}$ RDC. The introduction of $^{2}D_{\mathrm{C10–H1}} = -1.23$ Hz causes a little distortion on the alignment tensor orientation and size, whereas the inclusion of $^{2}D_{\mathrm{C10–H1}} = +12.71$ Hz causes a large rotation and a change in tensor size.

TABLE 6.13: Long-range RDC fit to two-membered ensembles. The possible values of the C10–H1 coupling were fitted separately. Results of the fit are compared with the $^1D_{CH}$ reference fit through differences in the quality of the fit (Q_{ang}), population of the first conformer of the ensemble (p_1) and rotation of the alignment tensor (β).

	$^1D_{CH}$		C10H1[1] (−1.23 Hz)			C10H1[2] (+12.71 Hz)		
Ensemble	Q_c	p_1	Q_{ang}	p_1	β	Q_{ang}	p_1	β
6A+6B	0.043	32	0.045	32	10.25	0.146	100	34.64
6A+6C	0.228	74	0.662	33	48.31	0.137	94	8.24
6A+6D	0.082	54	0.285	20	43.93	0.146	100	41.61
6A+6E	0.022	44	0.083	32	6.91	0.144	98	25.47
6A+6F	0.342	0	0.698	0	57.23	0.144	100	51.28
6B+6C	0.083	50	0.181	75	39.83	0.297	68	57.67
6B+6D	0.132	100	0.263	47	52.03	0.321	27	128.42
6B+6E	0.132	100	0.286	0	37.01	0.598	100	126.39
6B+6F	0.082	57	0.197	74	32.31	0.207	72	50.14
6C+6D	0.188	53	0.287	28	83.16	0.235	35	46.46
6C+6E	0.025	56	0.189	27	38.48	0.309	100	72.48
6C+6F	0.183	58	0.507	63	49.64	0.288	70	71.23
6D+6E	0.179	0	0.258	66	47.68	0.322	84	116.14
6D+6F	0.126	42	0.264	74	30.73	0.333	97	82.83
6E+6F	0.042	48	0.094	53	13.20	0.345	0	52.92

[1-2] C10–H1 couplings corresponds to the equally labeled columns of Table 6.11
p_1, represents the population (molar fraction) of the first conformer of the corresponding ensemble.
β, is the generalized rotation (in °) of the tensors, respect to the to the one determined with one-bond data.

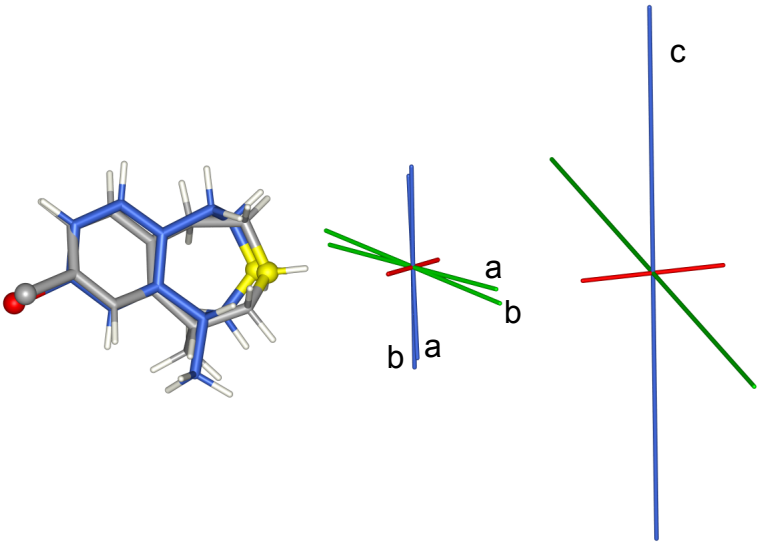


FIGURE 6.26: Graphical representation of the alignment tensor variation upon the introduction of $^2D_{\text{C10–H1}}$ RDC, see text. (a) Short-range tensor; (b) inclusion of $^2D_{\text{C10–H1}} = -1.23$ tensor; (c) inclusion of $^2D_{\text{C10–H1}} = +12.71$ tensor.

Following, $^2D_{\text{C4–H5}\beta}$ was introduced in the fit (Figure 6.14). The **6A**+**6B** ensemble furnished the best fit, followed by the **6E**+**6F**. The quality of the fit of the two possible values ($^2D_{\text{C4–H5}\beta} = +5.27/+7.71$) is indistinguishable in terms of penalty function value ($Q_{ang} = 0.032$). However, the comparison of population ratios (p_1) and tensor orientation (β) of the two fits with that of the short-range RDCs pointed to $^2D_{\text{C4–H5}\beta} = +5.27$ as the more compatible value. The population distortion is larger when $^2D_{\text{C4–H5}\beta}$ is assigned to +7.27 ($p_1 = 44$, $^1D_{\text{CH}}$ $p_1 = 32$) than that obtained when assigned to +5.27 ($p_1 = 37$, $^1D_{\text{CH}}$ $p_1 = 32$). Rotation of the alignment tensor described by the generalized angle β is more severe with the introduction in the fit of $^2D_{\text{C4–H5}\beta} = +7.71$ ($\beta = 34°$), than the resulting with the alternative assignment $^2D_{\text{C4–H5}\beta} = +5.27$ ($\beta = 20°$), that is very close to the $^1D_{\text{CH}}$ tensor. In summary, this suggests that the value $^2D_{\text{C4–H5}\beta} = +5.27$ is compatible with the previous data set ($^1D_{\text{CH}} + {}^2D_{\text{C10–H1}} = -1.23$) and can be regarded as the correct one. Actually, back-calculation of the $^2D_{\text{C4–H5}\beta}$ coupling with the previous data set ($^1D_{\text{CH}} + {}^2D_{\text{C10–H1}} = -1.23$) furnished values of 4.6 and 5.5 Hz for **6A**+**6B** and **6E**+**6F** ensembles, respectively, i.e. much closer to the "compatible" value $^2D_{\text{C4–H5}\beta} = +5.27$ than to the "incompatible" $^2D_{\text{C4–H5}\beta} = +7.71$.

In the following we will present a protocol based on Monte Carlo bootstrap analysis of the previous long-range RDCs combinations that permits the distinction between "good-fitting" and "bad-fitting" $^nD_{\text{CH}}$ RDCs based on their behavior when experimental errors are taken into consideration.

TABLE 6.14: Long-range RDC fit to two-membered ensembles. The possible values of C4–H5β coupling were fitted in conjunction with the one-bond and $^{2}D_{\text{C10–H1}} = -1.23$ Hz. Results of these fits are compared with the $^{1}D_{\text{CH}}$ reference fit through differences in Q_{ang}, population for the first conformer of the ensemble (p_1) and rotation of the alignment tensor (β).

	$^{1}D_{\text{CH}}$		C10H1[1] (−1.23 Hz) C4H5β[1] (+5.27 Hz)			C10H1[1] (−1.23 Hz) C4H5β[2] (+7.71 Hz)		
Ensemble	Q_c	p_1	Q_{ang}	p_1	β	Q_{ang}	p_1	β
6A+6B	0.043	32	0.032	37	20.13	0.032	44	63.73
6A+6C	0.228	74	0.560	64	44.71	0.510	70	53.75
6A+6D	0.082	54	0.471	67	37.43	0.444	72	46.12
6A+6E	0.022	44	0.142	34	16.86	0.189	64	47.29
6A+6F	0.342	0	0.648	61	69.50	0.575	91	50.36
6B+6C	0.083	50	0.140	76	37.31	0.199	78	39.50
6B+6D	0.132	100	0.241	100	39.31	0.283	100	37.14
6B+6E	0.132	100	0.236	100	39.58	0.276	100	37.29
6B+6F	0.082	57	0.158	75	27.11	0.234	78	44.82
6C+6D	0.188	53	0.317	48	115.421	0.316	49	88.35
6C+6E	0.025	56	0.199	28	55.87	0.252	30	36.67
6C+6F	0.183	58	0.558	43	52.70	0.559	36	64.15
6D+6E	0.179	0	0.319	0	21.13	0.390	0	26.67
6D+6F	0.126	42	0.615	31	62.64	0.598	20	62.09
6E+6F	0.042	48	0.065	53	13.16	0.120	53	64.45

[1-2] Long-range couplings corresponds to the equally labeled columns of Table 6.11.
p_1, represents the population (molar fraction) of the first conformer of the corresponding ensemble.
β, is the generalized rotation (in °) of the tensors, respect to the one determined with one-bond data.

The bootstrapping methodology was explained in Chapter 1, Section 1.4.5. In this case, 2048 decoys were generated with a standard deviation of 1.5 Hz. Once every decoy is fitted to the candidate structures, the back-calculated RDCs were analyzed —following Losonczi proposed methodology— with an acceptance Monte Carlo filter matched to the assumed experimental error (1.5 Hz). From this point on, only the accepted data sets (i.e. those whose back-calculated RDCs deviate from experimental ones less than the allowed filter error) are analyzed. Figure 6.27 shows the results of the bootstrap procedure followed by a Monte Carlo filter when fitted to ensemble **6A**+**6B** can be seen.

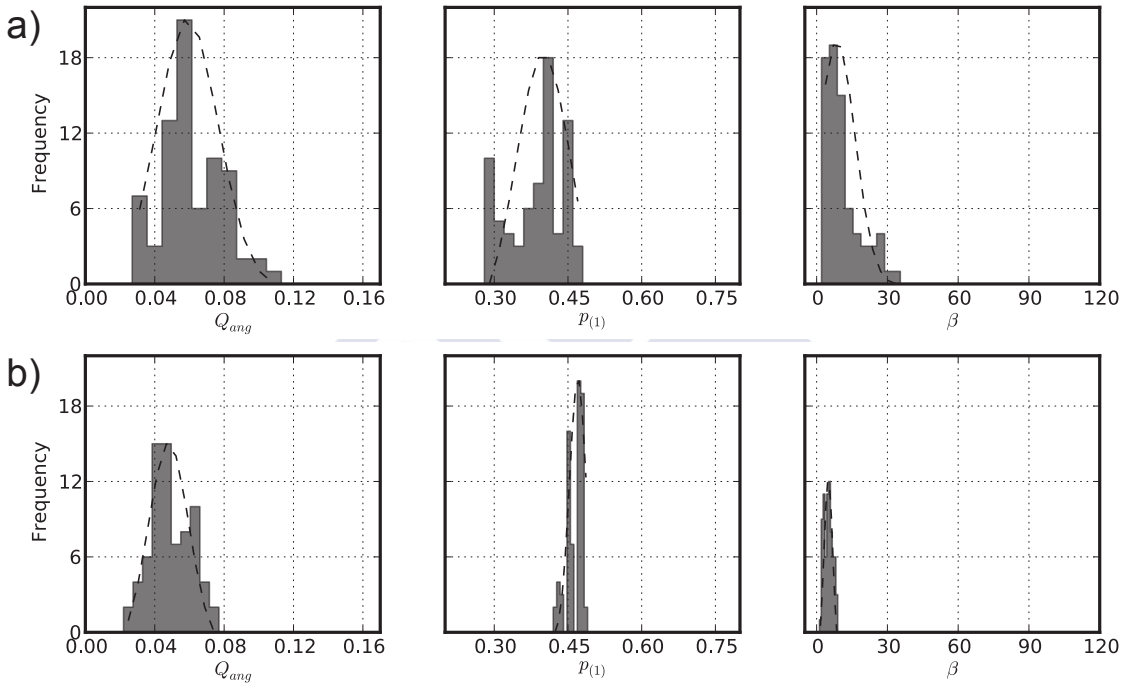


FIGURE 6.27: Summary plots of Monte Carlo filter results for the combination C10–H1 $= -1.23$ and C4H5β fitted to ensemble **6A**+**6B**. Only accepted data-sets are taken into account. Statistics of the bootstrap analysis can be seen in Table 6.15. The histograms show, in normalized frequency scale: left, the distribution of Q_{ang}; center, the distribution of p_1; right, the distribution of the generalized angle β respect to the tensor determined with experimental data. The dashed line over the distributions represents the normalized Probability Distribution Function (centered in the median). a) C4H5$\beta = +5.27$ Hz; b) C4H5$\beta = +7.71$ Hz.

The C4H5$\beta = +5.27$ value furnished broader distributions of Q_{ang}, p_1 and β than the C4H5$\beta = +7.71$ value. Apart from that, the C4H5$\beta = +5.27$ value furnished a higher Q_{ang} mean value than that furnished by C4H5$\beta = +7.71$ (Table 6.15). Nevertheless, when taking into account their standard deviation, the Q_{ang} values are mutually compatible. This is not the case for the population of conformer **6A** (p_1) between the two possible assignments. When assigning C4H5β to +7.71, the resultant population ($p_{6A} = 46\% \pm 2$) is not compatible with the population determined by $^1D_{\mathrm{CH}}$ RDC

Monte Carlo analysis ($p_{6A} = 35\% \pm 7$, Table 6.10). On the contrary, when assigning C4H5β to +5.27, the resulting population ($p_{6A} = 38\% \pm 6$) is compatible with populations determined with one-bond data only ($p_{6A} = 35\% \pm 7$, Table 6.10) and with $^2D_{\text{C10–H1}} = -1.23$ RDC ($p_{6A} = 32\%$, Table 6.13).

TABLE 6.15: Statistics of Monte Carlo filter results for the candidate combinations C10–H1 = −1.23 and C4H5β fitted to ensemble **6A**+**6B**.

		Q_{ang}		p_1		β	
value	accept	μ	σ	μ	σ	μ	σ
C4H5β = +5.27	74	0.061	0.018	0.38	0.06	10	7
C4H5β = +7.71	73	0.049	0.012	0.46	0.02	4	1

μ represents the arithmetic mean.
σ is the biased standard deviation estimator.
p_1 represents the population (molar fraction) of the first conformer of the corresponding ensemble.
β is the generalized rotation (in °) of the tensors from that determined from the experimental data.

We decided to adopt the C4H5β = +5.27 value, as this is the only assignment that results in an ensemble population within 1σ deviation —with a probability of 68.2%— of the Monte Carlo exploration of populations using $^1D_{\text{CH}}$ data. Additionally, the back-calculation of long-range couplings with the $^1D_{\text{CH}}$-determined alignment tensor (Figure 6.25) furnished a value C4H5β = +4.61, i.e. very close to the experimental one. From here on, the combination of $^1D_{\text{CH}}$ with long-range $^2D_{\text{CH}}$ C10–H1 = −1.23 and C4H5β = +5.27 will be used to test the inclusion of the non-stereoassigned long-range couplings involving methylenes.

6.5.3 Analysis of the long-range RDCs from non-stereoassigned methylenes C2 and C4

The definitive tensor determined in the previous section was used to back-calculate $^3D_{\text{C10–H2a}}$ and $^2D_{\text{C5–H4a}}$ RDCs for the two possible stereospecific assignments of the methylenic protons (Table 6.16). It can be appreciated at first sight in Table 6.16 that the different assignments of the long-range coupling furnish almost indistinguishable couplings in many ensembles. The correlation between the experimentally determined values and the back-calculated ones can be seen in Figure 6.28. It can be easily appreciated the indistinguishable values of the α or β long-range couplings involving these protons. Therefore, it is no possible to assign the diasterotopic methylenic protons based on back-calculated long-range RDCs involving them. Fortunately, the majority

of the ensembles had very similar back-calculated values for these couplings, making possible to rule out one of the two possible ${}^{n}D_{\mathrm{CH}}$ values shown in Table 6.11, the so-assigned value is shown in bold type in Table 6.16.

TABLE 6.16: Back-calculated long-range couplings involving the non-stereoassigned methylenes C2 and C4.

Ensemble	C10–H2α	C10–H2β	C5–H4α	C5–H4β
6A+6B	1.62	1.25	2.77	2.37
6A+6C	4.78	0.7	0.33	3.82
6A+6D	2.41	−0.3	2.49	8.12
6A+6E	2.56	−0.12	1.42	0.74
6A+6F	1.53	0.89	2.65	3.46
6B+6C	0.75	0.7	4.56	1.49
6B+6D	0.54	0.56	4.75	0.81
6B+6E	0.54	0.56	4.78	0.81
6B+6F	0.16	0.79	3.98	1.49
6C+6D	−3.05	3.84	−6.92	7.87
6C+6E	1.88	−0.13	3.04	0.31
6C+6F	2.84	2.4	1.26	0.37
6D+6E	2.08	0.19	2.91	−0.68
6D+6F	0.02	1.04	3.5	2.47
6E+6F	1.07	−0.18	2.54	−0.48
Experimental	$+\mathbf{0.33}/-9.49$		$+1.82/+4.38$	

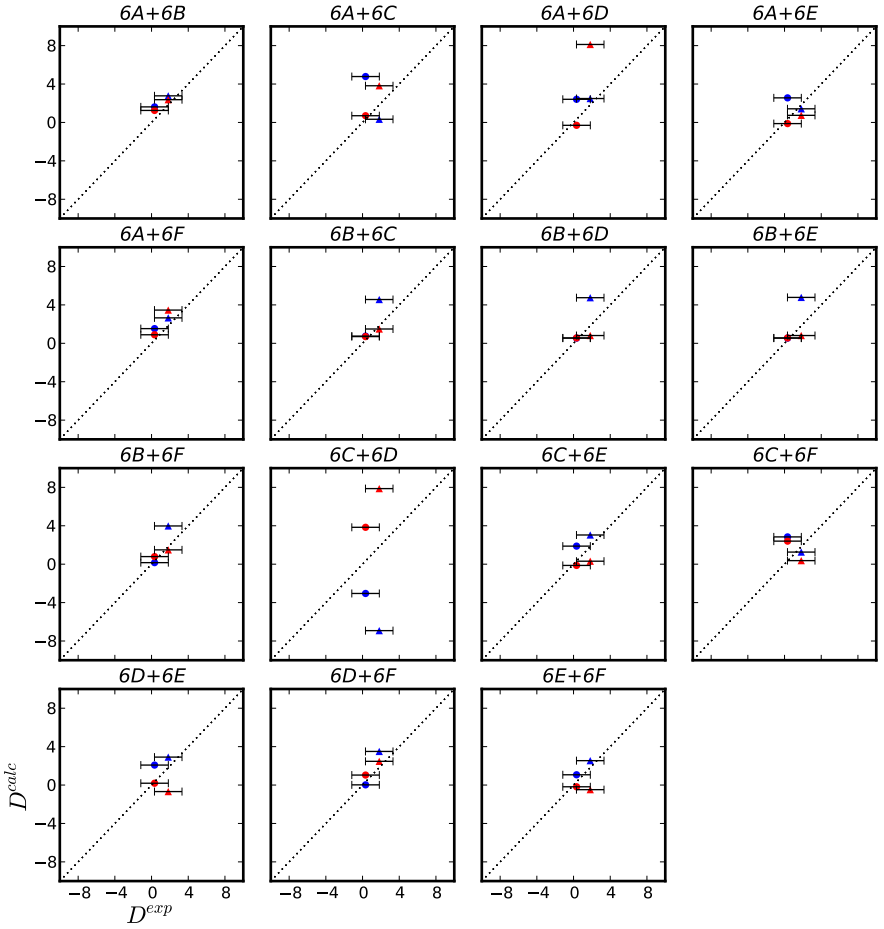


FIGURE 6.28: Plots of back-calculated ${}^{n}D_{\mathrm{CH}}$ RDCs involving the non-stereoassigned methylenic protons Table 6.16. ${}^{3}D_{\mathrm{C10–H2a}}$ couplings are marked with a circle and ${}^{2}D_{\mathrm{C5–H4a}}$ couplings with a triangle. Markers of couplings from Cx to Hα are colored in blue, while markers of corresponding couplings to Hβ are colored in red.

In and equivalent way as in the case of the C4H5β RDC, Monte Carlo-filtered bootstrap procedure will serve to analyze whether these two unassigned couplings can be introduced in the fit, or not.

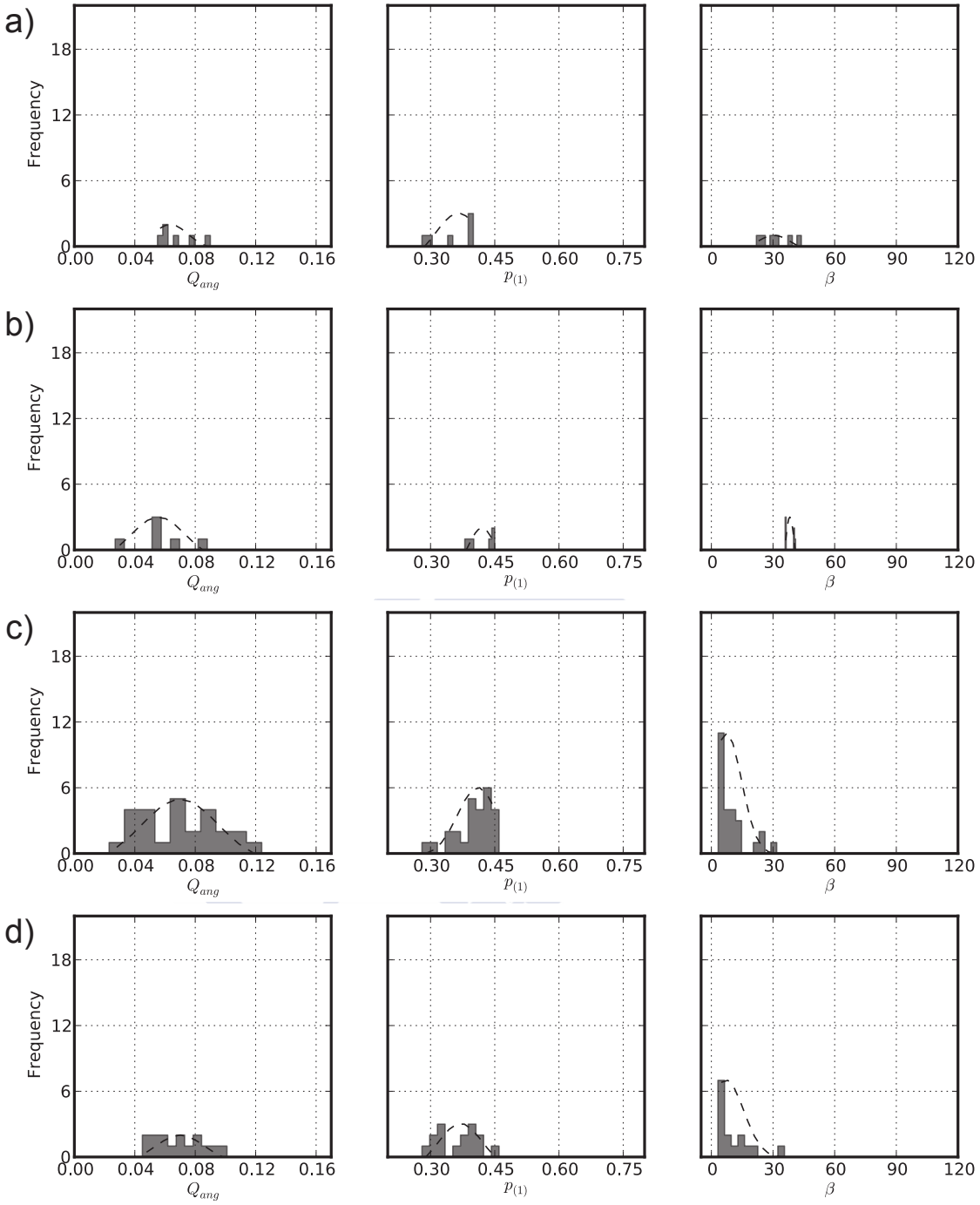


FIGURE 6.29: Summary plots of Monte Carlo filter results for the combination C10–H1 $= -1.23$ and C4H5$\beta = +5.27$ fitted to ensemble **6A**+**6B**. Only accepted data-sets are taken into account. Statistics of the bootstrapping analysis can be seen in Table 6.17. The histograms show, in normalized frequency scale: left, the distribution of Q_{ang}; center, the distribution of p_1; right, the distribution of the generalized angle β respect to the determined with experimental data. The dashed line over the distributions represents the normalized Probability Distribution Function (centered in the median). a), C10–H2α; b) C10–H2β; c) C5–H4α; d) C5–H4β.

From the bootstrapping histograms (Figure 6.29) it can be appreciated a general deterioration in the fit when including C10–H2a either assigned to H2α or H2β. In Table 6.17, the introduction of these coupling is compared with the previous fit. There is a sizable reduction in the number of decoys passing the Monte Carlo filter and a severe rotation of the alignment tensor described by the generalized angle β. This deterioration of the fit can be attributed to the small size of the coupling and bad peak shape, among other factors, which suggest not to use this coupling in the fits. Regarding the other coupling C5–H4a, its fit was much better and compatible with fits with the previous data set (C10–H1 $= -1.23$ Hz and C4H5β $= +5.27$ Hz). Differences in the fit quality estimators between the two stereoassignments (H4α/β) were negligible, thus making impossible to assign this coupling.

TABLE 6.17: Statistics of Monte Carlo filter results for the introduction of C10–H2a$(\alpha/\beta) = +0.33$ and C5H4a$(\alpha/\beta) = +1.82$ fitted to ensemble **6A**+**6B**.

		Q_{ang}		p_1		β	
value	accept	μ	σ	μ	σ	μ	σ
C4H5$\beta = +5.27$	74	0.061	0.018	0.38	0.06	10	7
C10–H2α	6	0.068	0.012	0.35	0.05	31	8
C10–H2β	8	0.057	0.018	0.42	0.03	38	2
C5–H4α	26	0.071	0.025	0.40	0.05	9	7
C5–H4β	15	0.069	0.016	0.36	0.05	10	8

μ represents the arithmetic mean
σ is the biased standard deviation estimator.
p_1 represents the population (molar fraction) of the first conformer of the corresponding ensemble.
β is the generalized rotation (in °) of the tensors from that determined from the experimental data.

These results make impossible the assignment of the diasterotopic protons of C2 and C4 by RDC fits, and makes reasonable the exclusion of their long-range couplings for the calculation of the alignment tensor. Still, we can assign the configuration of the C2 and C4 protons based on *Karplus* computed ensemble-averaged $^3J_{HH}$ couplings, and then include the consequently assigned dipolar coupling into the RDC fit. As the population ratio does not deviate significantly from the determined from $^1D_{CH}$, we can rely on the same assignment (see Table 6.9).

In summary, bootstrapping calculations with a Monte Carlo filter fitting $^1D_{CH}$ and $^2D_{C10–H1}$ and $^2D_{C4–H5\beta}$ long-range RDCs, selected the ensemble **6A**+**6B** with a population ratio 38 : 62 ($\sigma = \pm 6$).

6.5.4 RDC fit to three-membered ensembles

One-bond and long-range RDC fits pointed to a conformational ensemble of two members: **6A**+**6B**, which fulfills all available NMR structural restraints as well as computational findings. Nevertheless, we also considered ensembles of three conformations.

Ensembles that contain the pair **6A**+**6B** fit clearly better than the rest (Figure 6.30). Moreover, the contribution of the third conformer to **6A**+**6B** was negligible ($p_3 < 1\%$) with **6C**, **6D**, **6F**, and **6E** conformers (Figure 6.30). Therefore, these results support the validity of the calculation with two-membered ensembles in the previous section.

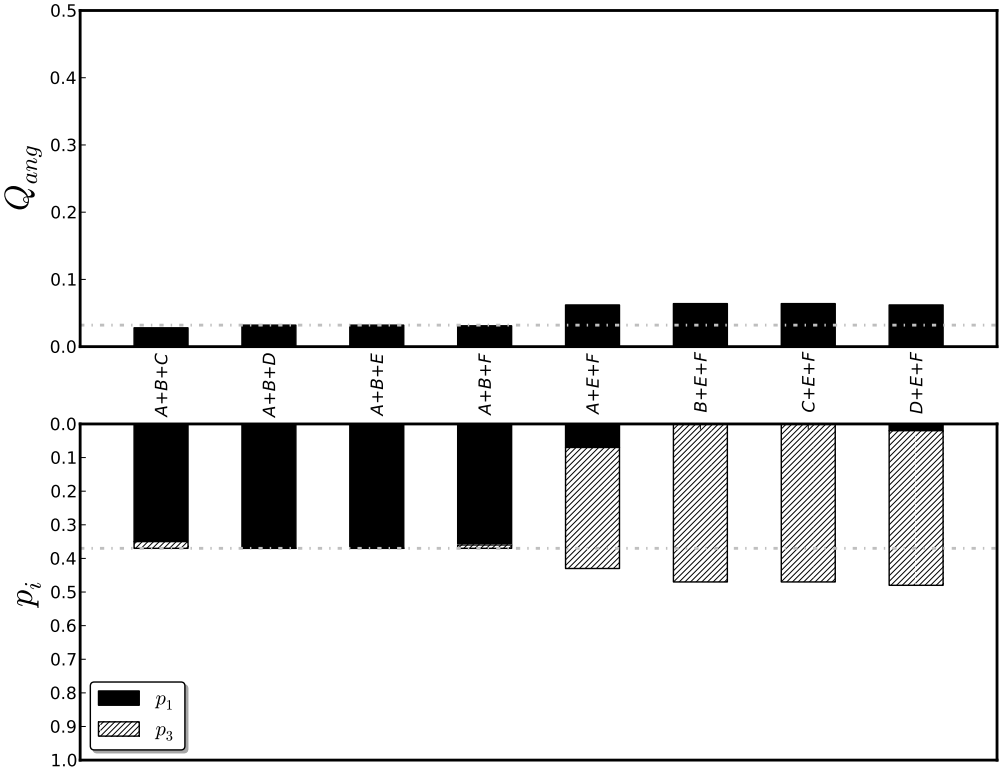


FIGURE 6.30: RDC quality factor Q_{ang} (top) and population (bottom) from 3-membered ensembles fit. Population p_3 corresponds to the population of the third conformer different than **6A** and **6B**. A light gray dash-dotted line indicates the reference population of conformer **6A**.

6.5.5 RDC fit to four-membered ensembles

Moreover, fit of 4-membered ensembles resulted in a preferential selection of of those containing **6A**+**6B** (Figure 6.31). Importantly, the population of other conformers than **6A** or **6B** in most of the ensembles is less than 1σ ($\sigma = 6\%$) as determined previously for ensemble **6A**+**6B**.

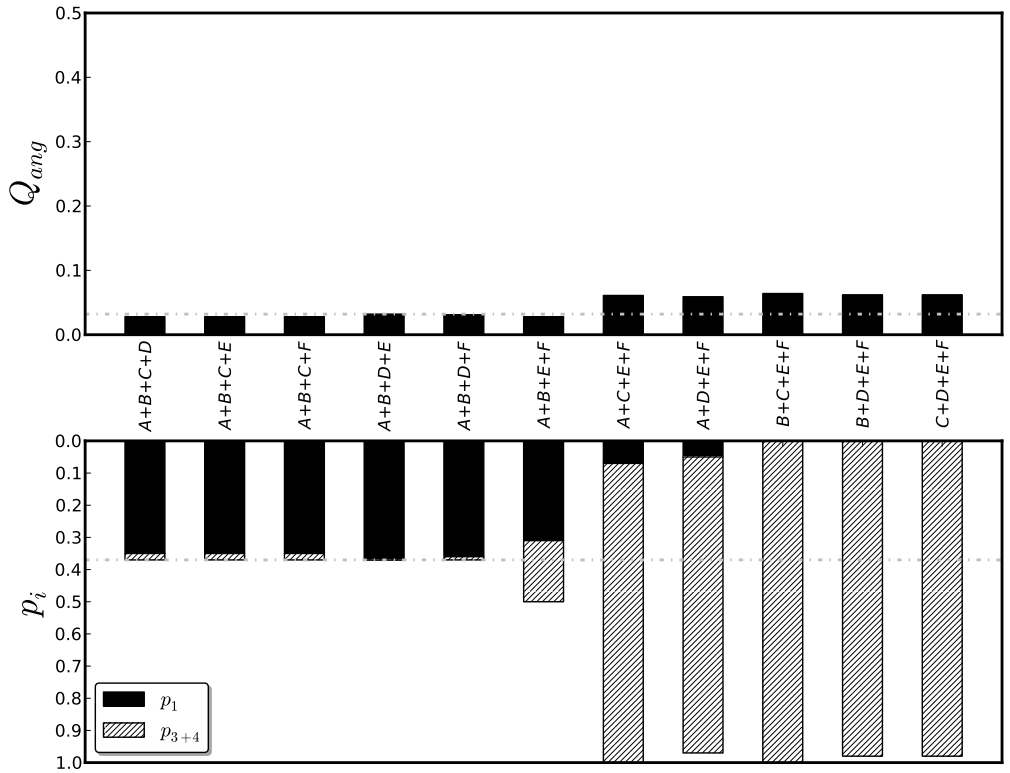


FIGURE 6.31: RDC quality factor Q_{ang} (top) and population (bottom) from 4-membered ensembles fit. Population p_{3+4} corresponds to the sum of the populations of the third and fourth conformers different than **6A** and **6B**. A light gray dash-dotted line indicate the reference population of conformer **6A**.

6.5.6 Model selection

It is usually under question how many conformers should be taken into account for a particular analysis. We have shown that ensembles containing two members provide a model that can explain all the experimental and computational restraints. Additionally, in the previous sections we fitted the data to ensembles composed by three and four members, obtaining compatible results.

Ensembles with different number of members (models) can be evaluated in terms of the well-known Akaike's Information Criterion (AIC),[239] which was successfully used by Griesinger and co-workers for their conformational analysis of lactose.[240] AIC is expressed as

$$AIC = \chi^2 + 2k \ , \tag{6.4}$$

where k is the number of fitted parameters —alignment tensor elements (5) and population weight parameters ($N-1$)— and χ^2 is calculated as follows

$$\chi^2 = \sum_j^n \left(\frac{D_j^{exp} - D_j^{calc}}{\Delta D_j^{exp}} \right)^2 , \tag{6.5}$$

in which ΔD_j^{exp} is the experimental error of each measurement.

In Table 6.18, the corresponding χ^2 and AIC scores for the different models proposed are shown.

TABLE 6.18: Quality of the fit of the different models proposed of lorcaserin.

Model	χ^2	k	AIC
6A	410.34	5	420.3
6B	50.72	5	60.7
6A+**6B**	1.24	6	13.2
6A+**6B**+**6C**	0.93	7	14.9
6A+**6B**+**6D**	1.24	7	15.2
6A+**6B**+**6E**	1.24	7	15.2
6A+**6B**+**6F**	1.19	7	15.2
6A+**6B**+**6C**+**6D**	0.93	8	16.9
6A+**6B**+**6C**+**6E**	0.93	8	16.9
6A+**6B**+**6C**+**6F**	0.93	8	16.9
6A+**6B**+**6D**+**6E**	1.24	8	17.2
6A+**6B**+**6E**+**6F**	1.19	8	17.2

The AIC values of the single conformations are far higher ($AIC > 60$) than that of ensembles ($AIC < 20$). Akaike's Information Criterion allows the comparison of different models with different number of elements. In our case, this makes possible a direct comparison between the 2-, 3- and 4-membered ensembles. The model with two conformers fits clearly better ($AIC = 13.2$) than 3-membered ($AIC = 14.9$) and 4-membered ($AIC = 16.9$) ones. It is important to stress that **6A**+**6B**+**6C** and **6A**+**6B**+**6C**+**6D** —the lowest AIC ensembles of 3 and 4 members— have a very low population of conformers different that **6A** or **6B**. Therefore, a two-conformer model suffices to explain the observed data at their present level of accuracy.

6.6 Conclusion

Lorcaserin hyrochloride (**6**) was analyzed with classical NMR techniques that determined the existence of conformational equilibrium while in solution. NOE (Figure 6.7, and Figure 6.5), vicinal $^{3}J_{HH}$ (Table 6.8), and previous computational studies (Table 6.1 and Table 6.2) strongly indicate the presence of more than one conformer. This findings are further supported by the single structure fit of RDCs obtained in three PMMA-acrylamide gels, in which no single conformer furnished a Q_C quality factor lower than 0.1 (Table 6.5). The evidence of conformational equilibrium is also supported by the results of our conformational search (Table 6.3 and Figure 6.3) pointing to two low-energy conformers (one of them previously detected by X-ray crystallography) with Boltzmann-derived populations around 50 : 50.

The determination of the lorcaserin conformational preference in solution by $^{1}D_{CH}$ analysis alone was not possible (Table 6.6). Nevertheless, the combination of classical NMR structural restraints (NOE and $^{3}J_{HH}$, Figure 6.17) with RDC fits allowed us to determine that lorcaserin exists in solution as a mixture of conformers **6A** and **6B** with populations of 35 : 65 ($\sigma = 7$) favoring the X-ray determined conformer (**6B**).

By using SJS-HSQC experiments we measured three $^{2}D_{CH}$ and one $^{3}D_{CH}$ long-range ^{13}C–^{1}H RDCs (Table 6.11 and Figure 6.19), two of them involving H4/H2 that were not stereospecifically assigned. Due to the novelty of the use of long-range RDCs we, predicted (*ab initio*) their $^{n}J_{CH}$ couplings, which resembled quite accurately the experimental values (Table 6.12).

When the long-range RDCs are small, two different values of that dipolar coupling, are compatible with the measured total coupling (Table 6.11, see Figure 6.21). This uncertainty on the value of long-range RDC was overcame with a combination of RDC back-calculation from the $^{1}D_{CH}$-determined tensor, and the introduction of each value at once, testing the rotation of the aligment tensor and the deviation of population probabilities as well as the quality of the fit (Section 6.5.2.1). This protocol allowed us to determine the correct value of long-range couplings and to include them in the fit.

The $^{2}J_{CH}$ and $^{3}J_{CH}$ involving unassigned diasterotopic protons did not enhance the structural discrimination of the fit, making impossible their use to assign those protons and to include them in the fits (see Section 6.5.2.1 and Section 6.5.3).

Fit of all RDCs $^{1}D_{CH}$ and $^{2}D_{CH}$ resulted in a clear distinction of ensemble **6A**+**6B** (Table 6.14) over the other ensembles. Importantly, the the population determined with one-bond data 35 : 65 ($\pm$7), which is very close to that determined with $^{1}D_{CH}$+$^{2}D_{CH}$ (38 : 62 $\pm$6).

Furthermore, we considered 3- and 4-membered ensembles to test the stability and validity of our results (Figure 6.30 and Figure 6.31, respectively). This analysis resulted is a preferred selection of ensembles containing **6A**+**6B**. Interestingly, the contribution of conformers different from **6A** or **6B** was negligible for most of the ensembles.

Finally, we conducted a study regarding the number of conformers that explain best the experimental data. Model selection through the use of the Akaike's information criterion pointed to the ensemble composed by **6A** ($38\% \pm 6$) and **6B** ($62\% \pm 6$) as the best solution for the experimental RDCs.

6.7 Materials and Methods

6.7.1 Materials

Racemic **6** was synthesized by Dr. Maria C. de la Fuente according to a previously described procedure.[227]

6.7.2 Gel Preparation

Polymerization

A pre-gel solution containing AMPS, DMAA, and BIS (1 : 1 : 0.034 mol) with total monomer concentration of 0.75 mol/L was prepared in milli-Q water. The pre-gel solution (0.6 mL) was poured into the gel chamber (New Era Enterprises, 6 mm inner diameter) and polymerized at room temperature (30 min.), initiated by 0.0015 g/mL ammonium persulphate and 0.023 g/mL TEMED. After polymerization, gels were extracted from the chamber, washed in a neutralizing solution of 0.02 M HCl, KOH or NaOH, as indicated, and extensively washed with milli-Q water (1 $\times$ overnight, followed by 3 $\times$ 1 hour). The swollen gels were dried in a oven at 40 °C, typically for 2–3 days, and the resulting sticks were stored at room temperature until use.

Gel swelling

A dry polymer stick was placed on the bottom of the NMR tube and a solution of 25 mM lorcaserin hydrochloride was added (0.55 mL). The vertical growth of the gel was restricted by inserting a Shigemi plunger. Typically, gel alignment is assessed by measuring the ^{2}H quadrupolar splitting ($|\Delta\nu_Q|$) of the solvent. However, acrylamide-type gels may not always show $|\Delta\nu_Q|$ when they become anisotropic. Therefore, we

assessed the alignment of the sample by recording a fast high-resolution band-selective F_2-coupled HSQC spectrum of the aromatic region.

6.7.3 NMR

General

All experiments were carried out on a Bruker *Avance III–500* spectrometer operating at 500.13 MHz for ^{1}H, 125.23 MHz for ^{13}C and 62.25 MHz for ^{2}H and equipped with a BBFO Plus Smart probe room temperature probe with Z-only gradients.

NMR samples contained 25 mM of **6** in D_2O. To make sure that **6** was indeed in its protonated state, 1 equiv. of NaOH and excess HCl were successively added to the D_2O solution. Addition of alkali caused significant changes in the ^{1}H and ^{13}C chemical shifts, as well as partial precipitation of **6** due to deprotonation of the nitrogen atom. Addition of excess acid to the basified sample redissolved the precipitate and recovered the original spectra. All ^{1}H and ^{13}C resonances were assigned, save for the α and β diastereotopic proton labels, on the basis of HSQC, HMBC and NOESY experiments. The assignment spectra can be seen in Appendix A, Section A.5.1 and the assignment in Appendix A, Section A.5.2.

NOE

1D NOESY experiments were measured using the Bruker standard pulse sequence (*selnogpzs*) containing the ZQ-filter block as proposed by Keeler and co-workers in order to minimize zero-quantum contributions from scalar-coupled protons.[241] Experiments were recorded at mixing times (τ_{mix}, d8) of 200 and 400 ms.

RDC measurement: JS-HSQC

One-bond C–H couplings ($^1J_{CH}$ and $^1T_{CH} = {}^1J_{CH} + {}^1D_{CH}$) were extracted from F_1-coupled *J*-Scaled HSQC spectra (JS-HSQC).[52] All spectra were acquired as 512* (^{13}C) × 499* (^{1}H) data matrices, where N* refers to N complex pairs, and spectral widths of 9432 × 5000 Hz, respectively, using 8 transients per FID and 1.5 s delay between scans, with a total acquisition time of 54 ms in the F_1 dimension and a *J*-amplification factor, κ, of 3.

RDC measurement: SJS-HSQC

Long-range C–H couplings ($^nJ_{CH}$ and $^nT_{CH} = {}^nJ_{CH} + {}^nD_{CH}$) were obtained from a set of Selective *J*-Scaled HSQC experiments (SJS-HSQC).[182] A total of three SJS-HSQC spectra were recorded, where the following proton resonances were selectively inverted, respectively: $3.20 - 3.27$ ppm (H5β), $3.37 - 3.47$ ppm (H1 / H2β / H4β), and $3.03 - 3.18$ ppm (H2α / H4α / H5α). Pulses for the selective inversion of protons had the profile of the center lobe of a *sinc* shape and 30 ms (H5β and H1 / H2β / H4β) or 10 ms (H2α / H4α / H5α) duration. SJS-HSQC experiments were acquired as 512* (^{13}C) $\times$ 499* (^{1}H) data matrices, where N* refers to N complex pairs, and spectral widths of 9432 $\times$ 5000 Hz, respectively, using 16 transients per FID and 1.5 s delay between scans, with a total acquisition time of 54 ms in the F_1 dimension and a *J*-amplification factor, κ, of 20.

High-Resolution F_2-coupled HSQC

Spectra were acquired using a modified version of the pulse program *hsqcetgpsisp2.2* from the Bruker pulse program library in which the decoupling (*cpd*) statement during the acquisition was removed. Spectra were acquired as 64* (^{13}C) $\times$ 1024* (^{1}H) data matrices, where N* refers to N complex pairs, using 32 transients per FID and 1 s delay between scans, with a total acquisition time of 17 ms in the F_1 dimension (^{13}C) and 2.05 s in the F_2 dimension (^{1}H). Spectral width was 1 ppm and 30 ppm in F_2 and F_1, respectively.

6.7.4 Conformational Search

The conformational space of protonated lorcaserin **6** was explored using the GMMX[130] stochastic conformational search procedure as implemented in PCMODEL[111] using the MM3 force field.[150,151] Six different conformations were obtained within an energy cutoff of 5 kcal/mol. The geometries of the so-obtained conformations were then refined at the DFT level using the M052X functional[132] in combination with the 6-31+G** basis set and the solvation polarizable continuum model (PCM)[124] using Gaussian03[167] water parameters. Analytical frequencies were computed to verify the nature of stationary points and to obtain thermochemical parameters. All relative energies are reported as differences in Gibbs ($\Delta G_{298.15K}$) free energies. DFT derived energies are shown in Appendices (Appendix B, Section B.5.2). Note that the chloride counterion was neither included in the force field nor in the DFT calculations.

6.7.5 RDC fit

RDC fit was performed using the RDCFIT software, which makes use of Scipy,[114,115] Numpy,[116] and Openbabel[112,113] libraries. Optimization of the alignment matrix function was done by means of the Powell minimization algorithm.[117]

Input and output data are reported in Appendices. Molecular coordinates of conformers **1A-C** (Appendix B, Section B.5.1), RDC input tables in RDCFIT-ready format (Appendix B, Section B.5.3), and program outputs (Appendix B, Section B.5.4).

Chapter 7

Conclusions

The main goal of this thesis has been the development of new methodologies for the structural elucidation of small molecules by using NMR residual dipolar couplings and their application to the configurational and conformational analysis of flexible systems.

Resembling the objectives, the main conclusions of this thesis dissertation are the following:

- Synthesis and alignment properties of AMPS-acrylamide gels.

    1. Synthesis of AMPS-acrylamide gels was modified to polymerize these gels at room temperature by the addition of a second radical initiator to the pre-gel mixture. AMPS-acrylamide gel synthesized in this way do not present the stiffness of the heat-polymerized ones. Additionally, long polymerization times are not needed, but still can be used if desired without any drawback in the final gel.
    2. This modified gel was tested in different swelling conditions leading to different swelling and alignment properties. These investigations permitted the identification of a "maturation period" in which no macroscopic changes in the gel matrix lead to changes in the alignment properties of the gel.
    3. As was demonstrated for other alignment media, the alignment induced by AMPS-acrylamide gels can be tuned by the addition of ionic additives such as Na^+. Additionally, the presence of Na^+ ions modify the ion-pairs that are known to occur in the gel matrix, leading to stronger alignment.

4. We observed less swelling of AMPS-acrylamide gels when decreasing the polarity of the solvent or increasing the ionic force. Furthermore, the modification of the ionic force was concomitant with alignment tensor rotation.

- Application of RDC-enhanced NMR to the conformational analysis of a biologically relevant amine (salsolidine) under quasi-physiological conditions.

  1. In the case of MCI, the single-tensor approximation was used to discard high-energy conformations, assessing the validity of the methodology.
  2. Salsolidine was analyzed in terms of the single-tensor approach, finding two almost equally populated conformers. Additionally, other NMR structural restraints, such as chemical shifts and $^{3}J_{HH}$ couplings, pointed to the same solution, demonstrating the power of the methodology.
  3. A further methodological advance was made by the inclusion of averaged RDCs in the fit. Extending the previously developed methodology for the inclusion of methyl and phenyl groups RDCs, new matrix expansion elements were introduced for the inclusion of methyelene RDCs as a half-sum, thus circumventing both the assignment problem and the impossibility of taking accurate values for the individual protons in F_1-coupled HSQC experiments.

- Application of RDCs in combination with chiroptic methods (ECD) and high-level computations (TD-DFT) to determine the absolute configuration and the conformation of the newly isolated natural alkaloid 19-(OH)-(−)-Eburnamonine.

  1. The relative configuration as well as the conformation around the new stereogenic center of 19-OH-(−)-Eburnamonine was determined by RDC analysis. Relative configuration was independently confirmed by DFT chemical shift prediction.
  2. Once the relative configuration and conformation of 19-(OH)-(−)-Eburnamonine was known, its absolute configuration was determined by TD-DFT prediction of the ECD spectra and optical rotation and comparison with the experimental values.

- Enhancement of the structural information of RDCs by measuring long-range couplings with high accuracy, solving the relative configuration of the 10-Epi-8-deoxycumambrin B natural product with five stereogenic centers relying only on RDC analysis. In collaboration with Dr. Ad Bax and Dr. Jinfa Ying from NIH (Bethesda, USA).

1. The use of one-bond ($^1D_{CH}$) fits in combination with other NMR structural restraints and high-level computations could not determine the relative configuration of **5**. The introduction in the fit of the averaged RDCs from methyl and methylene groups using the previously developed methodology permitted the configuration determination in **5** but with low confidence.
2. The newly developed SJS-HSQC permitted the measurement of long-range couplings with high sensitivity and accuracy. These new restraints led to the complete structural discrimination of the five stereogenic centers of **5**.

- Combination of long-range couplings with the single-tensor approximation in order to determine the conformational equilibrium of lorcaserin, a 1-substituted 3-benzazepine targeted to obesity treatment.

  1. The conformational state of lorcaserin was determined by $^1D_{CH}$ RDC fit, in conjuction with NOE and ensemble averaged $^3J_{HH}$ (constrained to RDC-derived populations). The determined conformational state (**A**+**B**, 35 : 65, $\pm$7) fulfilled all experimental evidences as well as DFT-derived energies.
  2. The inclusion of long-range RDCs to the fit permitted the discrimination of the previously determined ensemble on the basis of RDCs alone. The inclusion of long-range couplings also reduced the statistical uncertainty of the RDC-derived populations.
  3. Additionally, the inclusion of long-range couplings permitted the extension of the analysis of conformational equilibrium to ensembles made of 3 and 4 conformers, selecting always the same pair as with the previous two-membered ensembles approach.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Appendix A

Appendix A: Assignment NMR spectra and experimental and computed NMR parameters

A.1 MCI

A.1.1 NMR Spectra

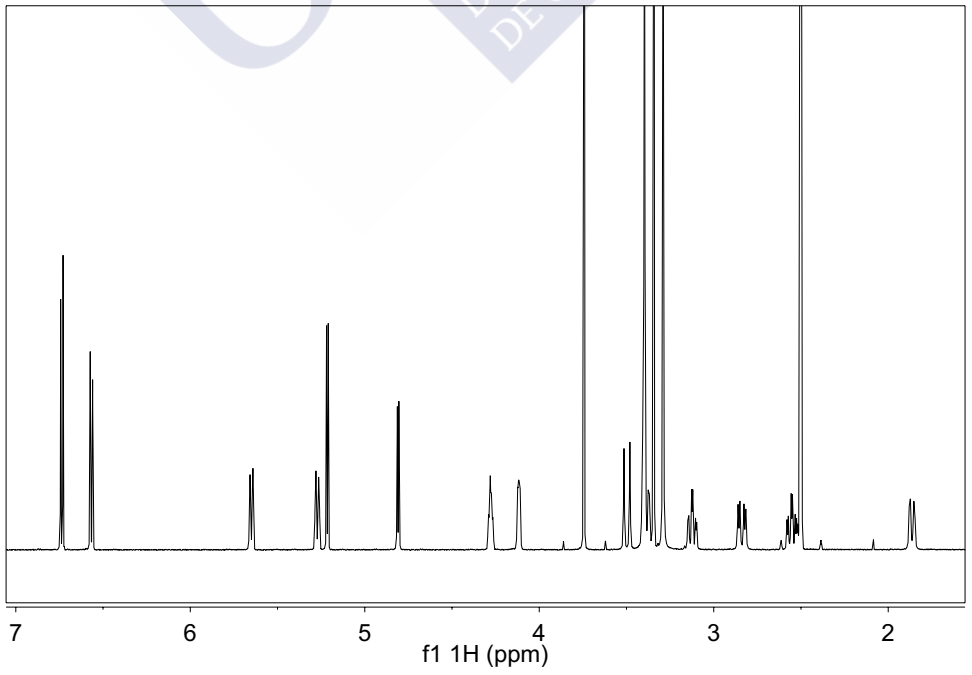


FIGURE A.1: 1D ^{1}H NMR of MCI **1**.

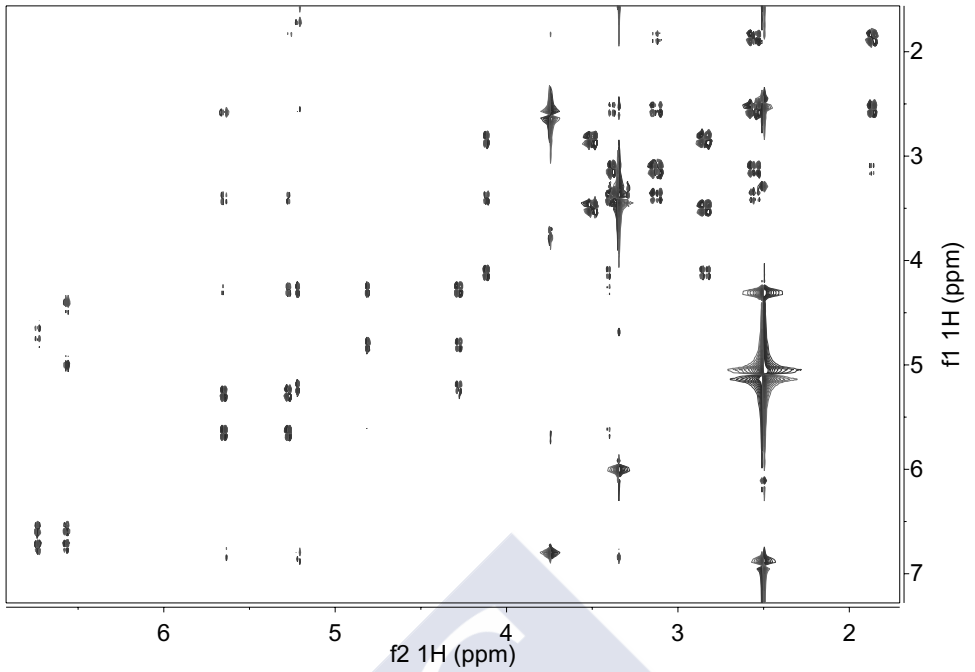


FIGURE A.2: 2D ^{1}H–^{1}H Double Quantum Filtered COSY spectrum of MCI **1**.

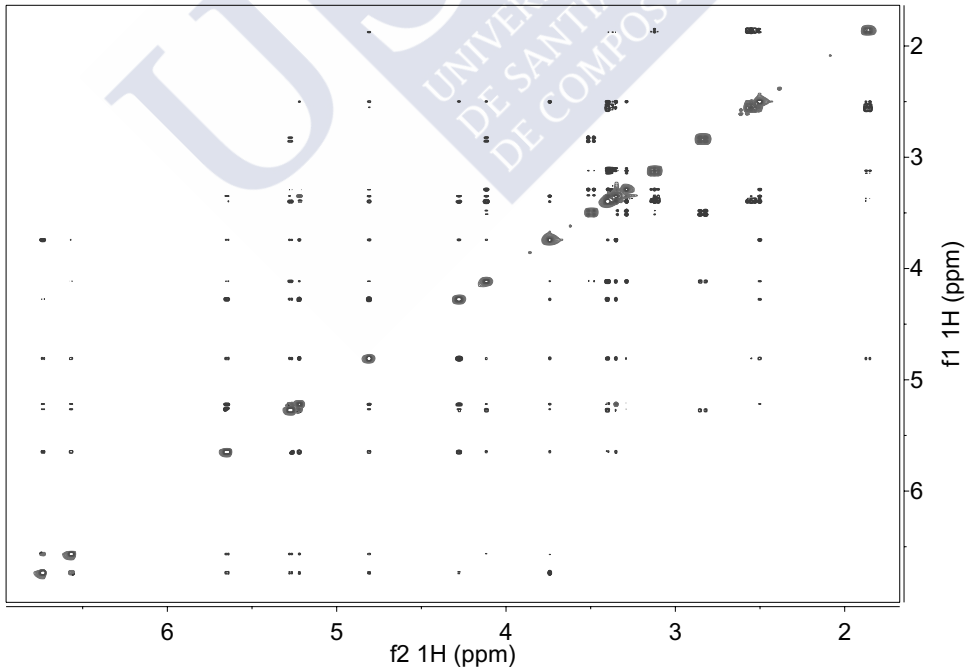


FIGURE A.3: 2D ^{1}H–^{1}H NOESY NMR spectrum of MCI **1**.

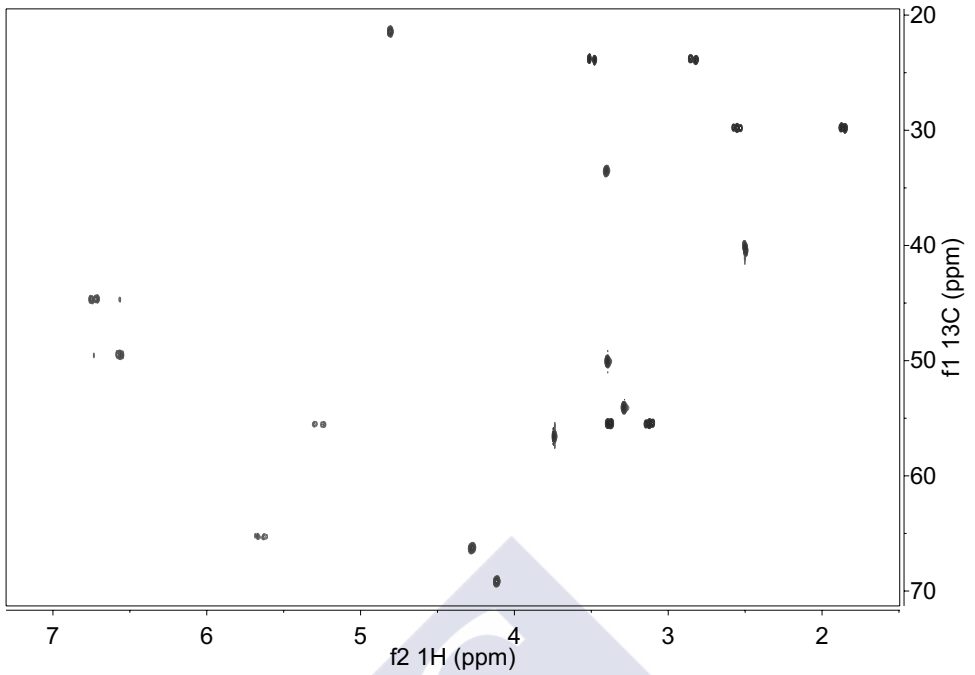


FIGURE A.4: 2D ^{1}H–^{13}C HSQC NMR spectrum of MCI **1**.

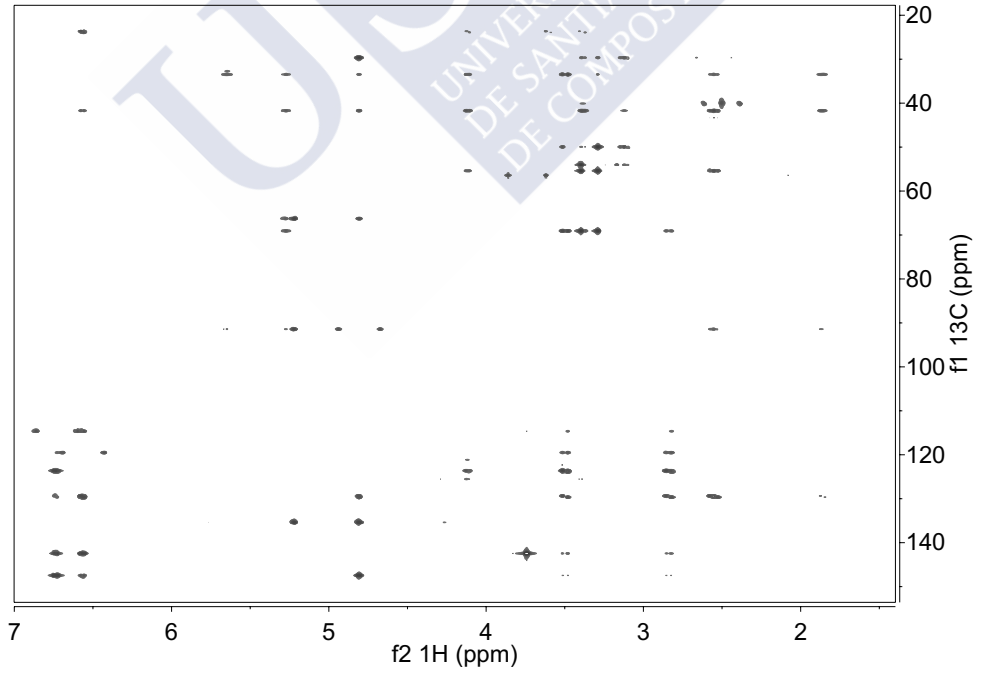


FIGURE A.5: 2D ^{1}H–^{13}C HMBC NMR spectrum of MCI **1**.

A.1.2 NMR assignment

TABLE A.1: MCI (**1**) NMR assignment

δ^{13}C	δ^{1}H	Assignment
23.89	3.50	10*a*
23.89	2.83	10*b*
29.80	2.55	15*a*
29.80	1.86	15*b*
33.54	3.40	14
41.71	–	13
50.08	3.39	22
54.06	3.29	23
55.43	3.37	16*a*
55.43	3.12	16*b*
56.56	3.74	21
66.28	4.28	6
69.13	4.11	9
91.44	4.81	5
–	5.22	20OH
114.63	6.73	2
119.51	6.57	1
123.69	–	11
125.56	5.28	8
129.49	–	12
135.30	5.64	7
142.42	–	3
147.44	–	4

A.2 Salsolidine

A.2.1 NMR Spectra

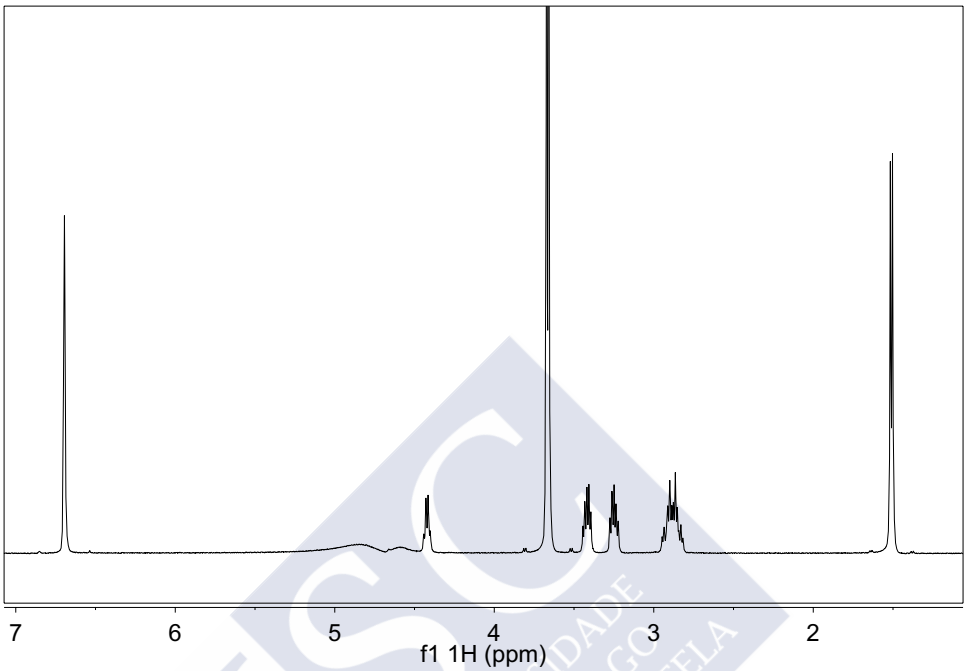


FIGURE A.6: 1D ^{1}H NMR spectrum of salsolidine **2**.

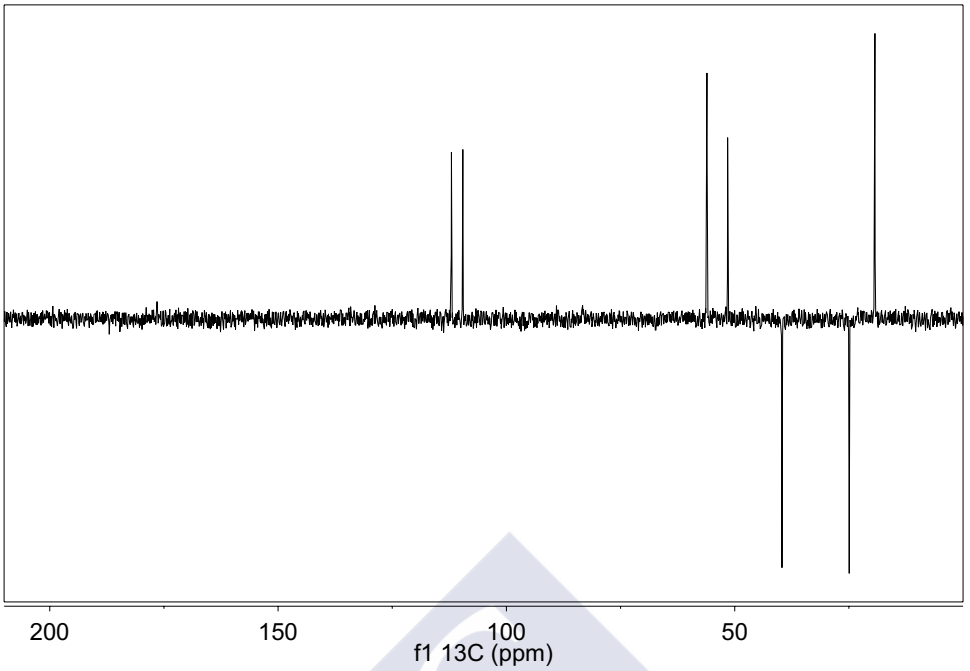


FIGURE A.7: 1D ^{13}C DEPT (135) NMR spectrum of salsolidine **2**.

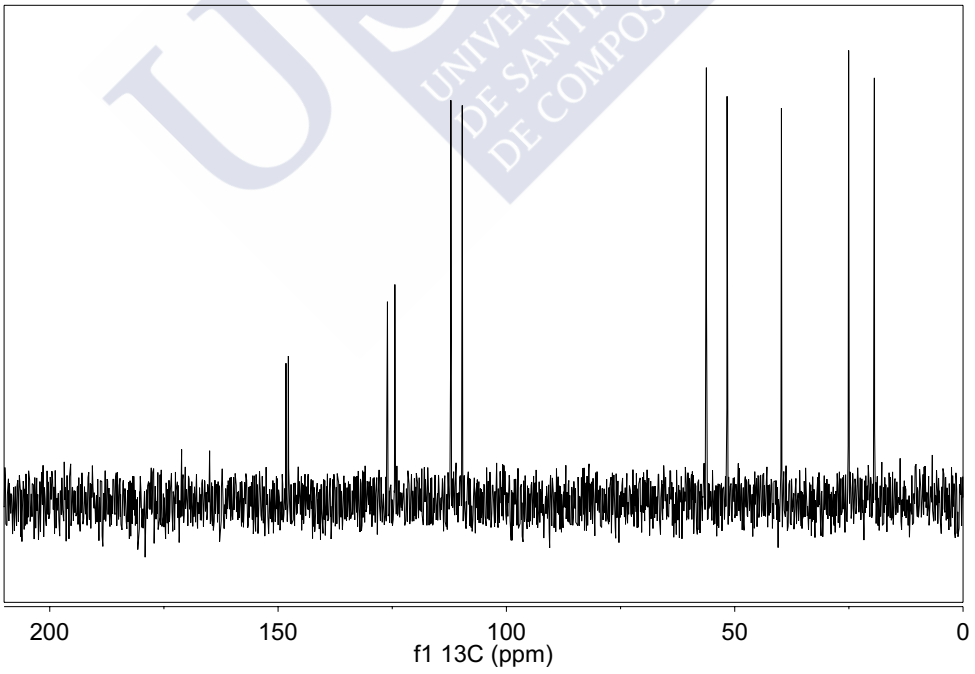


FIGURE A.8: 1D ^{13}C NMR spectrum of salsolidine **2**.

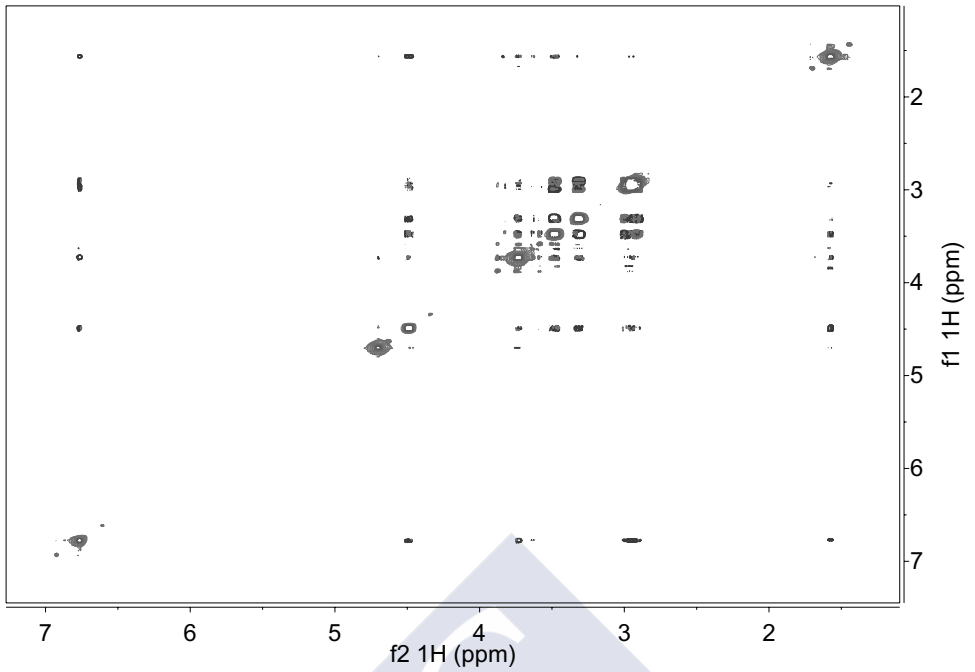


FIGURE A.9: 2D ^{1}H–^{1}H NOESY NMR spectrum of salsolidine **2**.

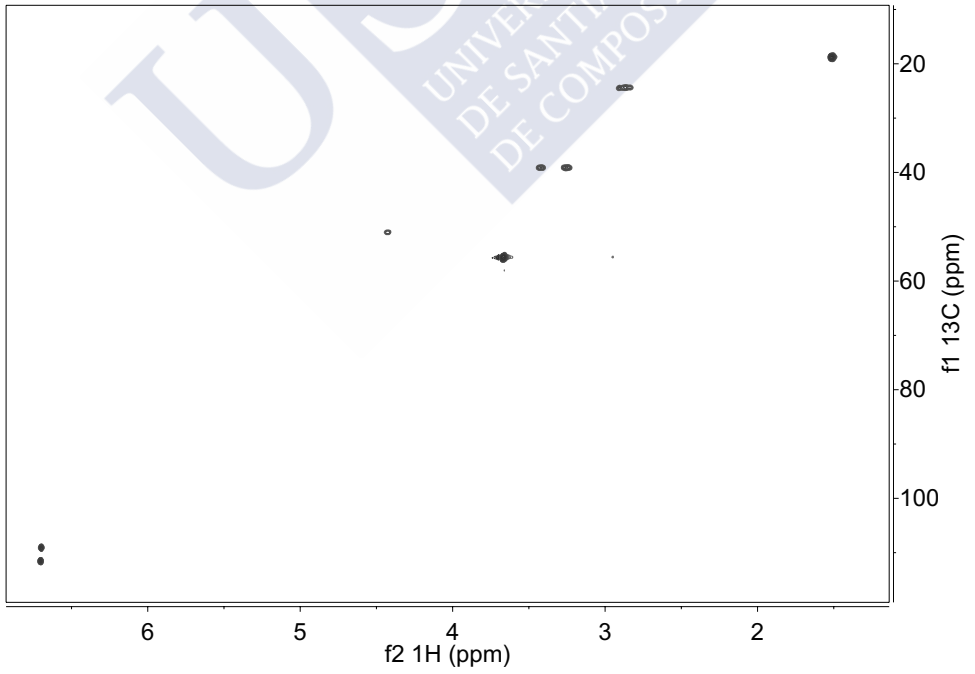


FIGURE A.10: 2D heteronuclear ^{1}H–^{13}C HSQC NMR spectrum of salsolidine **2**.

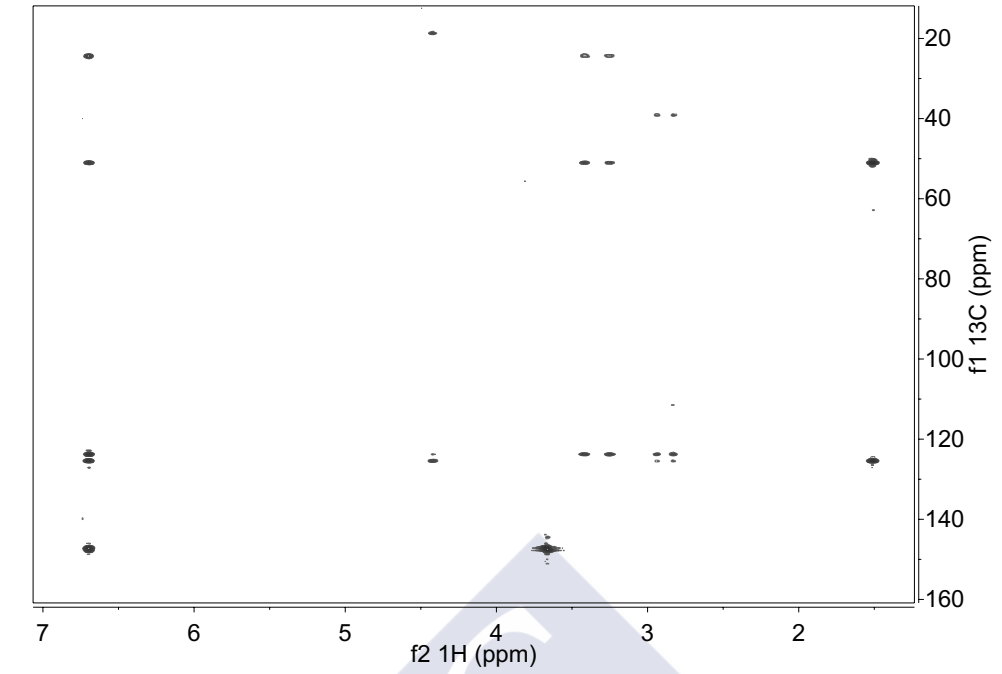


FIGURE A.11: 2D heteronuclear ^{1}H–^{13}C HMBC NMR spectrum of salsolidine **2**.

A.2.2 NMR assignment

TABLE A.2: Salsolidine (**2**) NMR assignments

$\delta^{13}C$	$\delta^{1}H$	Assignment
19.4	1.5	10
25.04	2.86	4
25.04	2.9	4
39.77	3.25	3
39.77	3.41	3
51.67	4.42	1
56.2	3.66	6/7
56.2	3.67	6/7
56.31	3.66	6/7
56.31	3.67	6/7
109.67	6.69	8
112.16	6.7	5

A.2.3 Other experimental and computed NMR parameters

TABLE A.3: Salsolidine (**2**) ^{13}C and ^{1}H experimental and computed chemical shifts.

Resonance	Experimental	**2A**	**2B**	**2C**	**2A**+**2B**[a]
C1	51.67	61.39	60.23	58.08	60.81
C3	39.77	48.85	43.77	47.32	46.31
C4	25.04	29.59	28.43	28.66	29.01
C4a	124.67	113.62	112.49	119.76	113.06
C5	112.16	106.69	106.47	106.64	106.58
C6	148.28	149.48	149.45	150.43	149.47
C7	147.76	148.37	148.25	148.44	148.31
C8	109.67	101.87	103.20	101.70	102.54
C8a	126.25	113.61	112.61	114.90	113.11
C9	19.40	20.60	22.69	16.74	21.65
H1	4.42	4.92	4.77	4.39	4.85
H3β	3.42	3.71	3.86	3.07	3.79
H3α	3.25	3.65	3.55	3.77	3.60
H4β	2.91	3.42	3.15	3.04	3.29
H4α	2.85	3.20	3.37	3.16	3.29
H5	6.70	6.77	6.69	7.06	6.73
H8	6.69	6.54	6.36	6.85	6.45
H9	1.50	1.76	1.79	1.83	1.78

[a] **2A**+**2B** ensemble 50% populated for each conformer.

TABLE A.4: Salsolidine (**2**) experimental and Haasnoot–Altona computed $^{3}J_{HH}$[a]

Structure	H3β–H4β	H3β–H4α	H3α–H4β	H3α–H4α
2A	5.9	1.2	12.3	3.6
2B	3.8	12.2	1.2	6.1
2C	4.1	12.1	1.2	6
2A+**2B**[b]	4.9	6.7	6.7	4.9
Experimental [c]	6.3	6.2	6.5	6.2

[a] Haasnoot-altona equation as implemented in MSPIN.
[b] Ensemble averaged $^{3}J_{HH}$ couplings assigning a 50% population to each conformer.
[c] Experimental $^{3}J_{HH}$ couplings for the extensively averaged H3/H4 averaged resonances derived from NUMARIT analysis in Spinworks.

A.3 19-OH-(−)-Eburnamonine

A.3.1 NMR Spectra

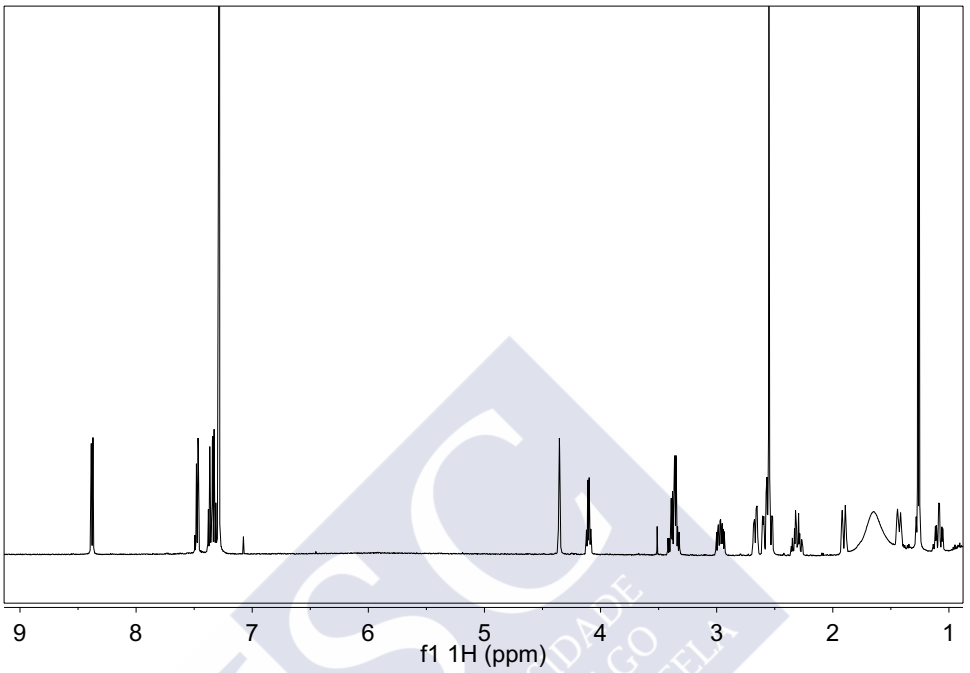


FIGURE A.12: 1D ^{1}H NMR spectrum of 19-OH-(−)-Eburnamonine **4**.

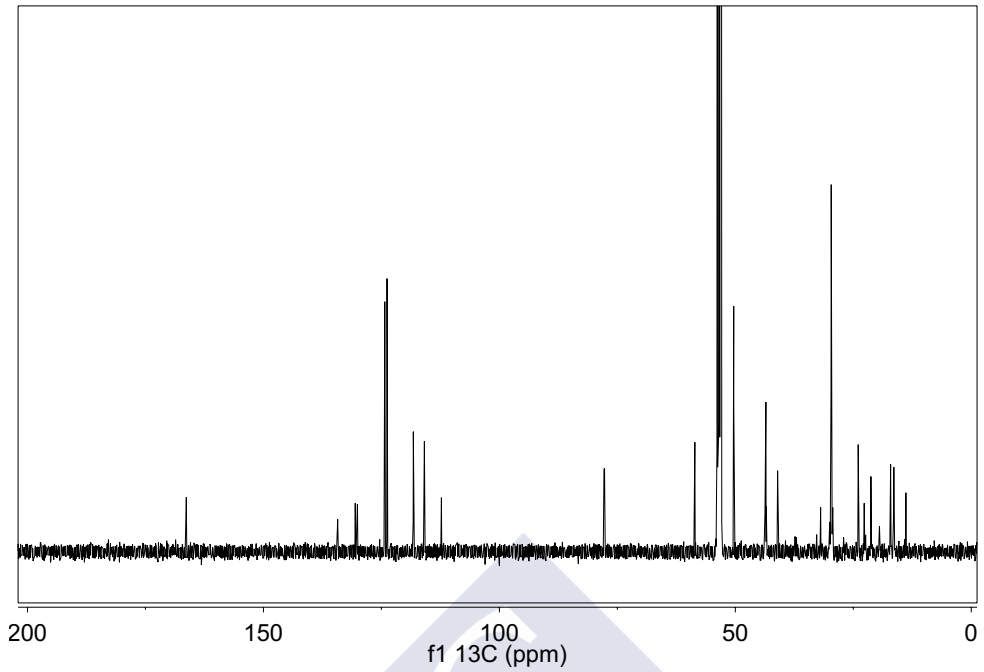


FIGURE A.13: 1D ^{13}C NMR spectrum of 19-OH-(−)-Eburnamonine **4**.

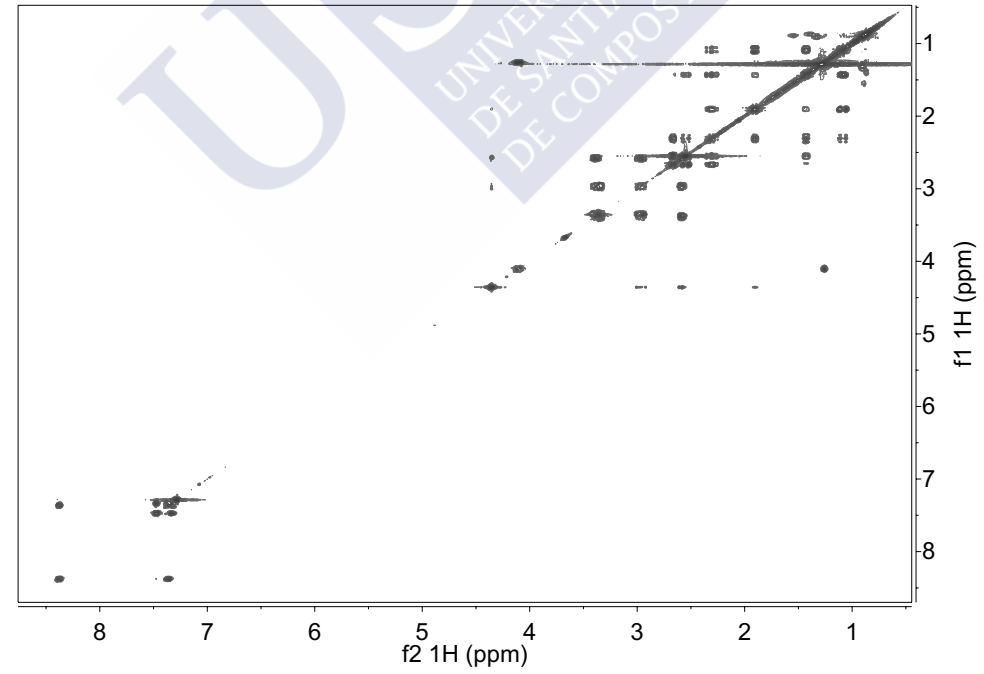


FIGURE A.14: 2D ^{1}H–^{1}H COSY (COSY45) NMR spectrum of 19-OH-(−)-Eburnamonine **4**.

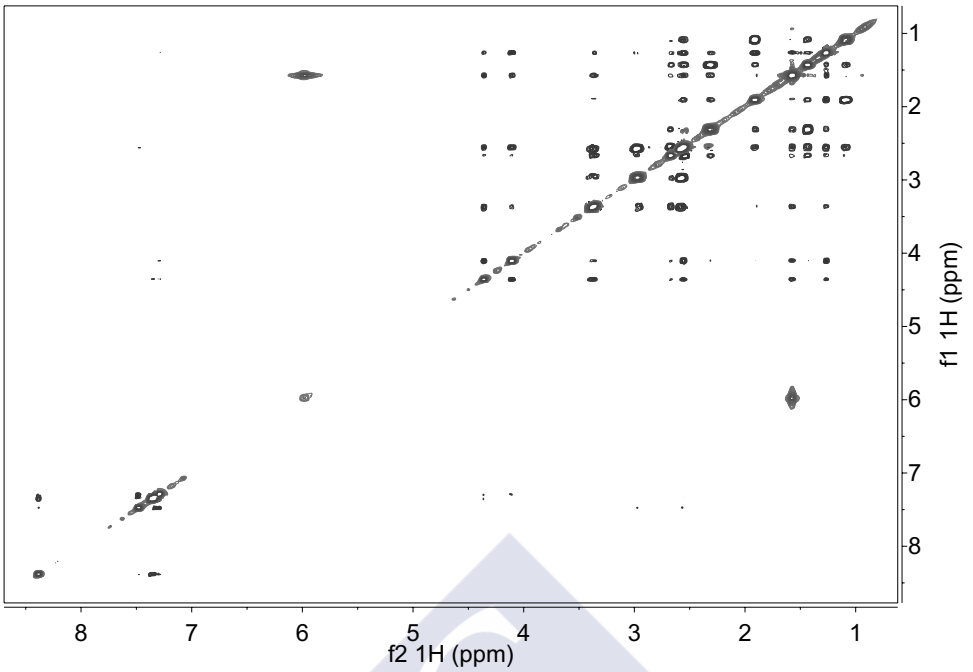


FIGURE A.15: 2D ^{1}H–^{1}H NOESY NMR spectrum of 19-OH-(−)-Eburnamonine **4**.

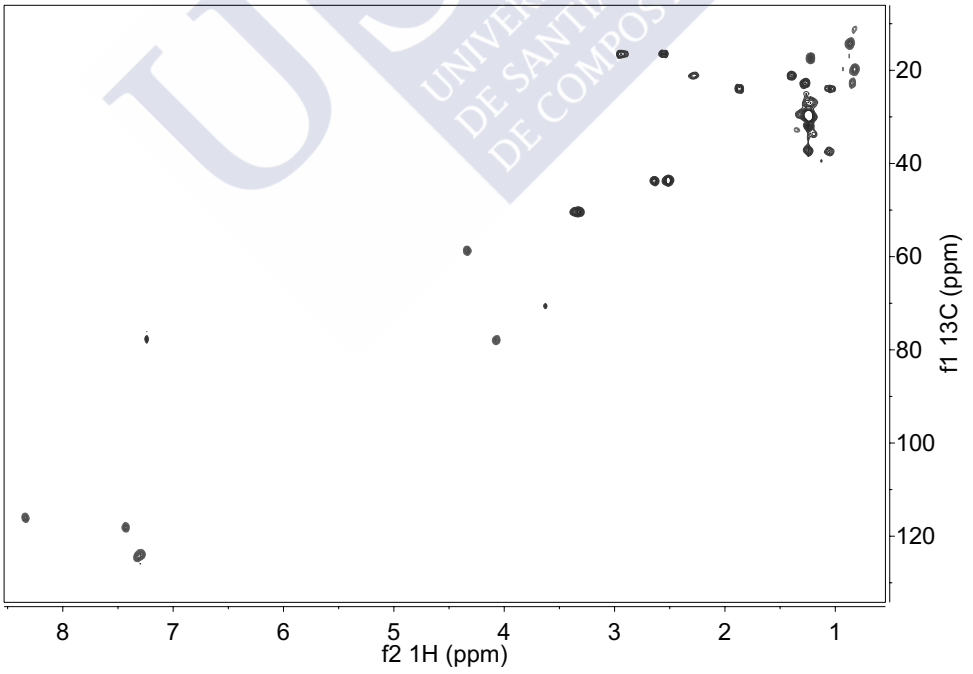


FIGURE A.16: 2D ^{1}H–^{13}C edited HSQC NMR spectrum of 19-OH-(−)-Eburnamonine **4**.

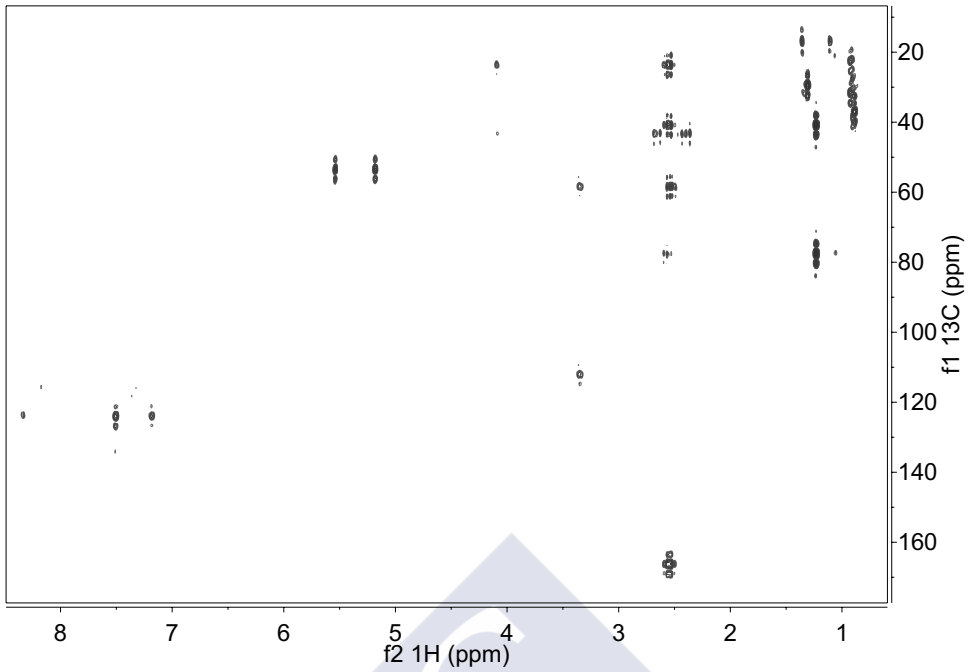


FIGURE A.17: 2D heteronuclear ^{1}H–^{13}C HMBC NMR spectrum of 19-OH-(−)-Eburnamonine **4**.

A.3.2 NMR assignment

TABLE A.5: 19-OH-(−)-Eburnamonine (**4**) NMR assignment

δ^{13}C	δ^{1}H	Assignment
16.7	2.94	6α
16.7	2.56	6β
17.5	–	18
21.3	2.28	14α
21.3	1.40	14β
24.1	1.88	15α
24.14	1.06	15β
41.2	–	20
43.8	2.52	17
43.9	2.65	3α
43.9	2.52	3β
50.6	3.35	5α
50.6	3.33	5β
58.9	4.33	21
78.2	4.08	19
78.2	5.95	19OH
112.6	–	7
116.5	8.35	12α
118.4	7.45	9α
124.3	7.31	10
124.9	7.33	11
130.1	–	2
130.3	–	8
134.5	–	13
166.5	–	16

A.3.3 Other experimental and computed NMR parameters

TABLE A.6: Boltzmann-averaged 19-OH-(−)-eburnamonine diastereomers computed ^{13}C and ^{1}H chemical shifts *in vacuo*. Values in ppm (referenced to TMS).

Resonance	Experimental	***SSS***	***SSR***	***SRS***	***SRR***
C10	124.08	124.06	123.78	122.08	122.04
C11	124.67	123.75	123.36	121.69	121.59
C12	116.29	116.44	115.85	114.39	114.39
C13	134.30	133.59	132.60	129.98	130.03
C8	129.92	127.19	126.77	125.72	125.90
C9	118.25	118.11	117.77	116.23	116.17
C7	112.41	112.98	114.07	111.09	111.39
C2	130.13	133.18	131.50	128.96	129.72
C21	58.73	64.54	66.61	59.48	52.61
C5	50.40	55.09	54.30	51.79	51.72
C6	16.50	26.01	25.59	19.28	19.51
C16	166.32	165.65	164.13	162.42	163.28
C17	43.66	46.15	47.42	44.29	43.58
C20	41.04	46.17	44.25	42.22	41.79
C15	23.95	41.16	33.23	24.31	32.02
C14	21.16	26.85	26.89	23.55	24.33
C3	43.71	57.98	56.93	44.60	44.16
C19	78.06	75.77	73.25	81.90	77.57
C18	17.35	20.77	18.64	16.95	20.39
H10	7.31	7.47	7.50	7.53	7.52
H11	7.33	7.44	7.47	7.51	7.52
H12	8.35	8.52	8.53	8.62	8.61
H9	7.45	7.59	7.63	7.67	7.67
H21	4.33	3.29	3.42	4.25	4.51
H5b	3.35	2.41	2.45	3.38	3.35
H5a	3.31	2.91	2.91	3.24	3.19
H6b	2.56	2.67	2.75	2.50	2.49
H6a	2.94	2.94	3.01	3.05	3.02
H17b	2.52	2.30	2.28	2.53	2.81
H17a	2.52	2.58	2.42	2.33	2.15
H15a	1.06	1.83	2.09	1.04	1.30
H15b	1.88	1.36	1.18	1.84	1.59
H14b	2.28	1.47	1.45	2.53	2.59
H14a	1.40	2.45	2.75	1.31	1.39
H3b	2.65	2.32	2.38	2.49	2.42
H3a	2.52	2.91	2.89	2.56	2.48
H19	4.08	4.40	3.78	4.26	4.07
H18	1.24	0.80	1.04	1.15	1.25

TABLE A.7: Boltzmann-averaged 19-OH-(−)-eburnamonine diastereomers computed ^{13}C and ^{1}H chemical shifts taking solvation into account (IEFPCM chloroform). Values in ppm (referenced to TMS).

Resonance	Experimental	***SSS***	***SSR***	***SRS***	***SRR***
C10	124.08	124.23	123.78	122.08	122.04
C11	124.67	123.86	123.36	121.69	121.59
C12	116.29	116.55	115.85	114.39	114.39
C13	134.30	133.67	132.60	129.98	130.03
C8	129.92	127.32	126.77	125.72	125.90
C9	118.25	118.22	117.77	116.23	116.17
C7	112.41	113.09	114.07	111.09	111.39
C2	130.13	133.24	131.49	128.96	129.72
C21	58.73	64.60	66.61	59.48	52.61
C5	50.40	55.21	54.30	51.79	51.72
C6	16.50	26.11	25.59	19.28	19.51
C16	166.32	165.67	164.13	162.42	163.28
C17	43.66	46.67	47.42	44.29	43.58
C20	41.04	45.80	44.25	42.22	41.79
C15	23.95	41.86	33.23	24.31	32.03
C14	21.16	27.15	26.89	23.55	24.33
C3	43.71	58.11	56.93	44.60	44.16
C19	78.06	76.78	73.25	81.90	77.57
C18	17.35	21.45	18.63	16.95	20.39
H10	7.31	7.48	7.50	7.53	7.52
H11	7.33	7.44	7.47	7.51	7.52
H12	8.35	8.52	8.53	8.62	8.61
H9	7.45	7.59	7.63	7.67	7.67
H21	4.33	3.32	3.42	4.25	4.51
H5b	3.35	2.39	2.45	3.38	3.35
H5a	3.31	2.91	2.91	3.24	3.19
H6b	2.56	2.68	2.75	2.50	2.49
H6a	2.94	2.95	3.01	3.05	3.02
H17b	2.52	2.31	2.28	2.53	2.81
H17a	2.52	2.54	2.42	2.33	2.15
H15a	1.06	1.84	2.09	1.04	1.30
H15b	1.88	1.38	1.18	1.84	1.59
H14b	2.28	1.49	1.45	2.53	2.59
H14a	1.40	2.57	2.75	1.31	1.39
H3b	2.65	2.32	2.38	2.49	2.42
H3a	2.52	2.91	2.89	2.56	2.48
H19	4.08	4.31	3.78	4.26	4.07
H18	1.24	0.76	1.04	1.15	1.25

A.4 10-Epi-8-deoxicumambrin B

A.4.1 NMR Spectra

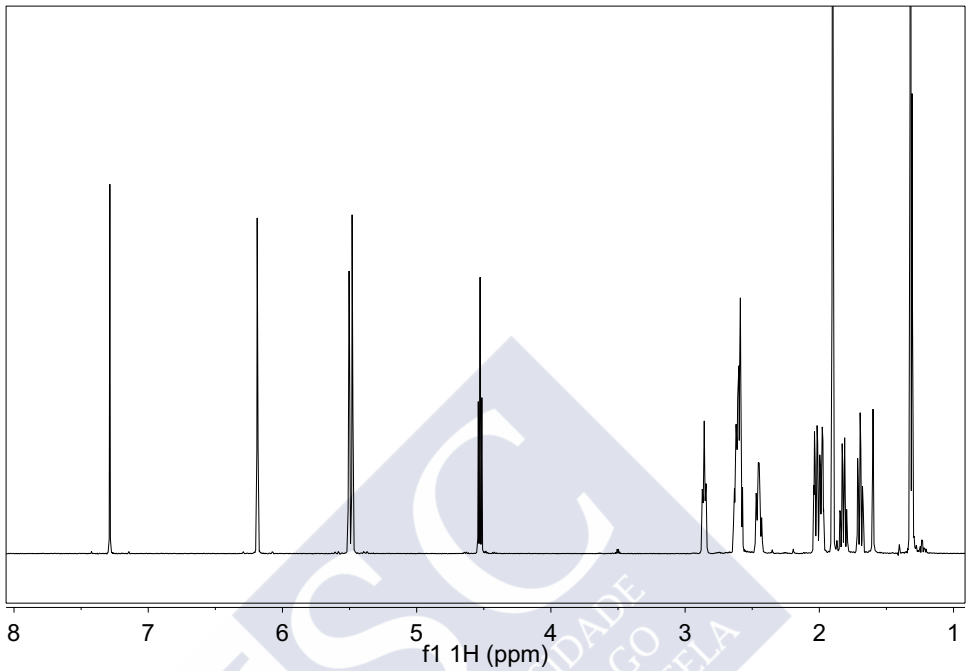


FIGURE A.18: 0-Epi-8-deoxicumambrin B **5** 1D ^{1}H NMR spectrum.

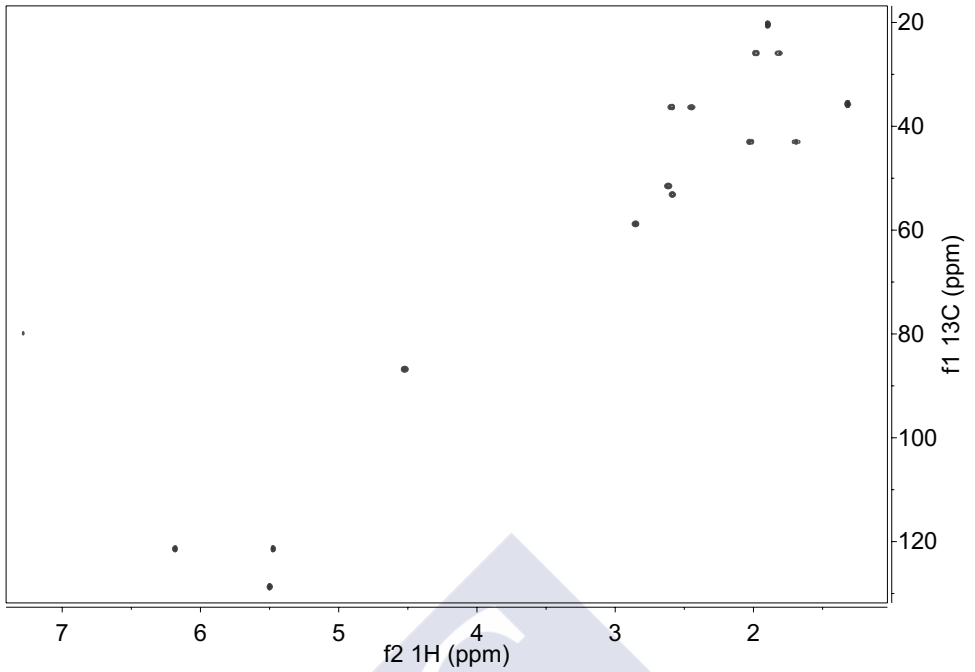


FIGURE A.19: 10-Epi-8-deoxicumambrin B **5** 2D heteronuclear ^{1}H–^{13}C HSQC NMR spectrum.

A.4.2 NMR assignment

TABLE A.8: 10-Epi (**5**) NMR assignment

$\delta^{13}C$	$\delta^{1}H$	Assignment
20.4	1.9	15
25.9	1.98	8*a*
25.9	1.82	8*b*
35.7	1.31	14
36.3	2.59	2*b*
36.3	2.45	2*a*
42.9	2.02	9*b*
42.9	1.69	9*a*
51.5	2.62	7
53.1	2.59	1
58.8	2.85	5
86.8	4.52	6
121.4	6.18	13*a*
121.4	5.48	13*b*
128.7	5.5	3

A.5 Lorcaserin

A.5.1 NMR Spectra

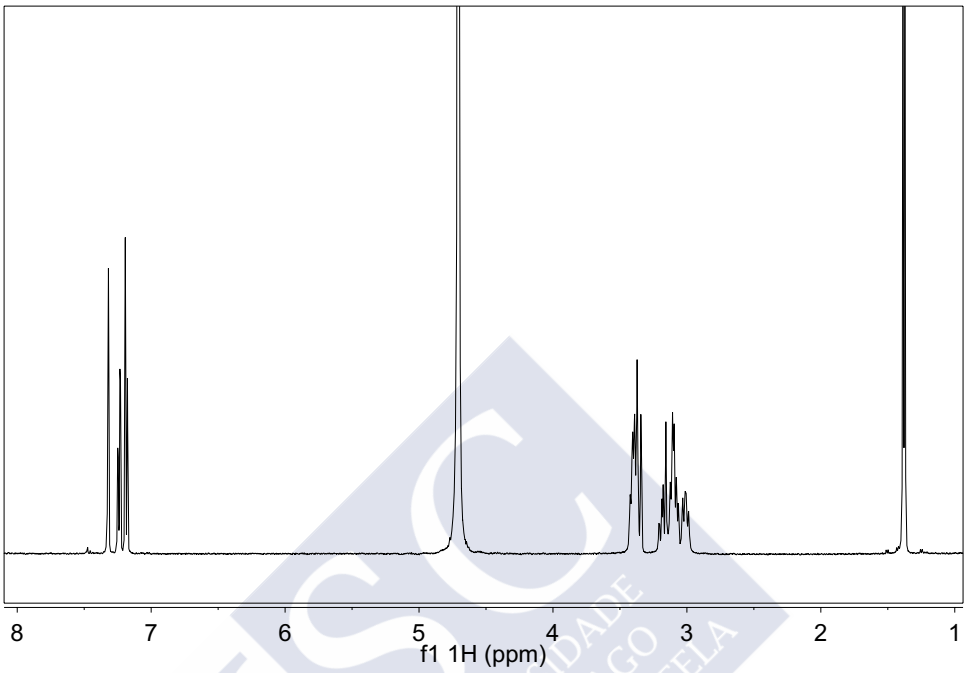


FIGURE A.20: 1D ^{1}H NMR spectrum of lorcaserin **5**.

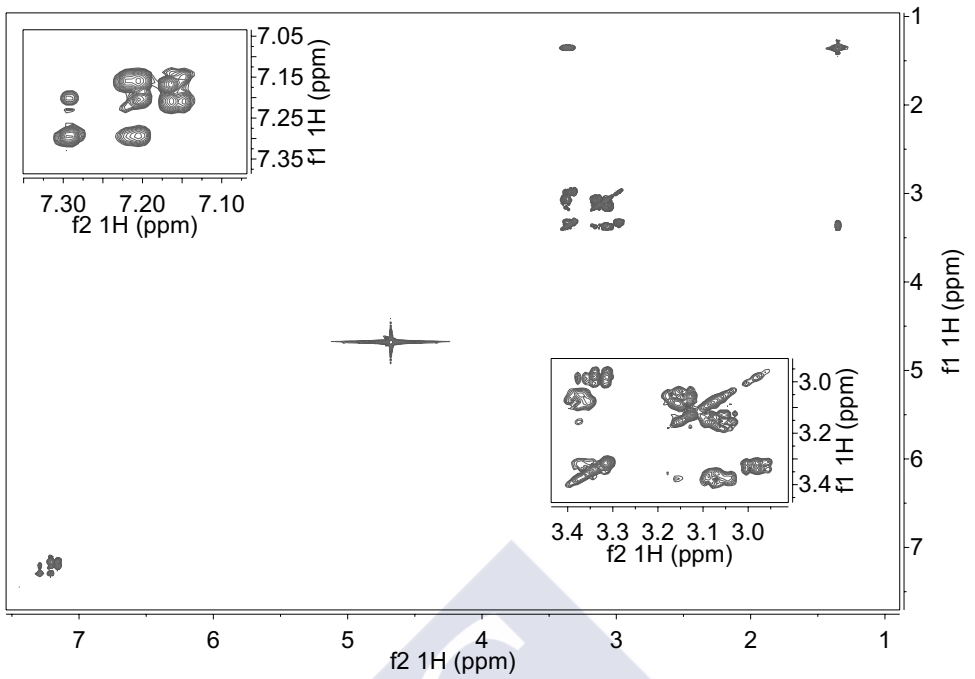


FIGURE A.21: 2D ^{1}H–^{1}H COSY (COSY45) NMR spectrum of lorcaserin **5**.

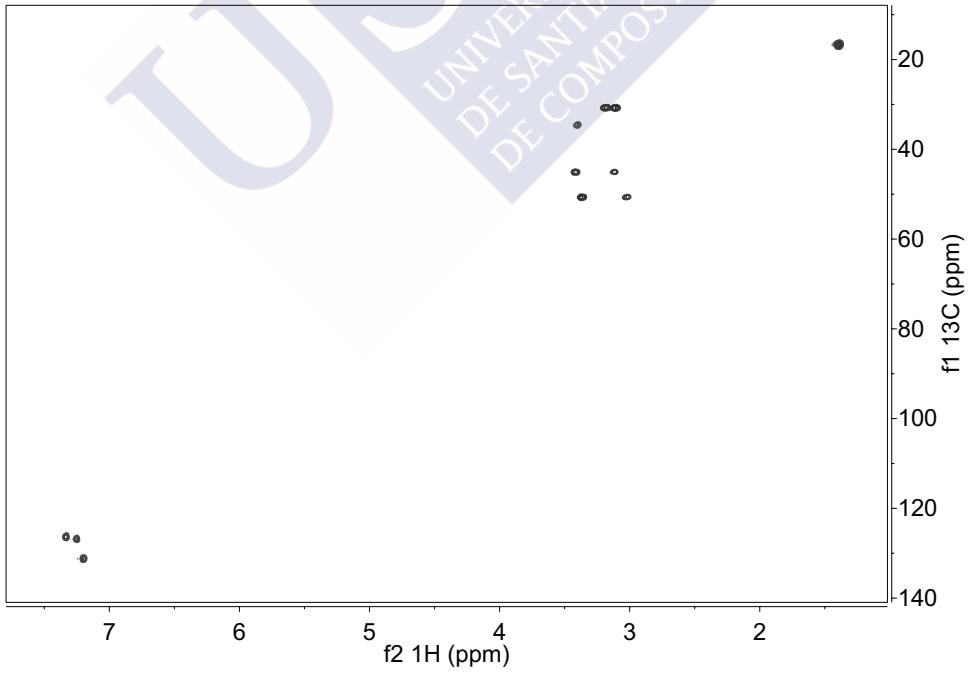


FIGURE A.22: 2D ^{1}H–^{13}C edited HSQC NMR spectra of lorcaserin **5**.

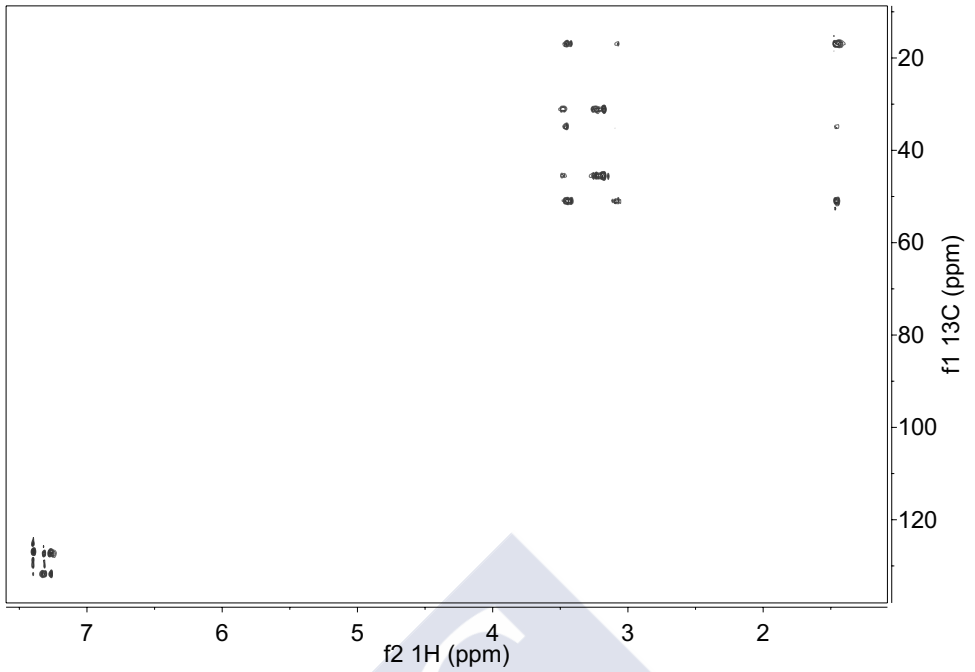


FIGURE A.23: 2D (^{1}H–^{1}H)–^{13}C edited HSQC–TOCSY (DIPSI-2 spinlock) NMR spectrum of lorcaserin **5**.

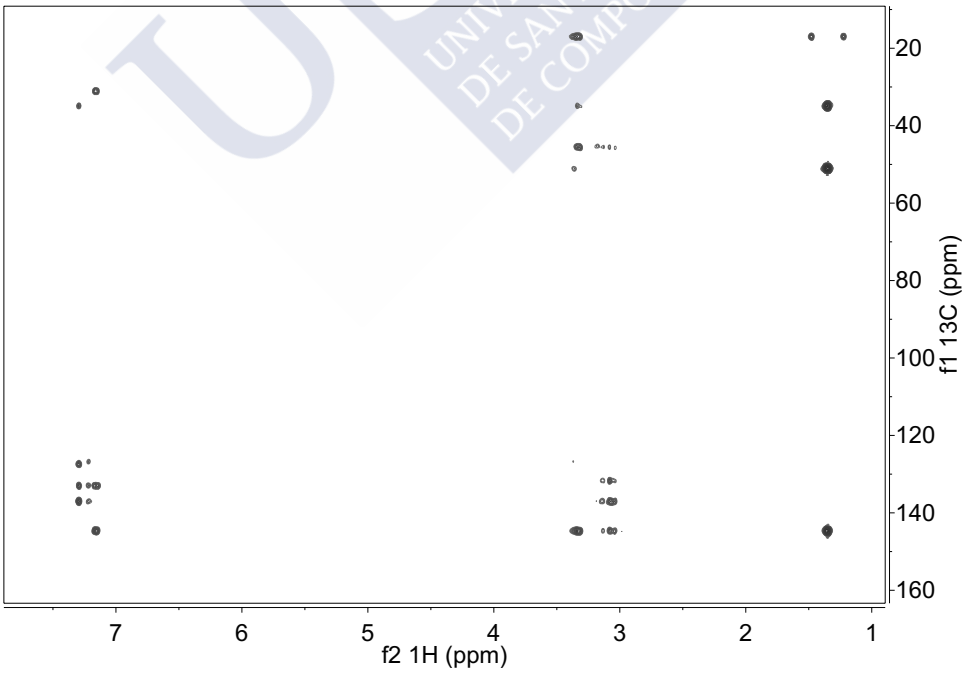


FIGURE A.24: 2D ^{1}H–^{13}C HMBC NMR spectrum of Lorcaserin **5**.

A.5.2 NMR assignment

TABLE A.9: Lorcaserin (**6**) NMR assignment

$\delta^{13}C$	$\delta^{1}H$	Assignment
1.39	19.6	10
3.08	33.8	5α
3.17	33.8	5β
3.39	37.6	1
3.11	48.1	4α
3.40	48.1	4β
3.01	53.7	2α
3.36	53.7	2β
7.32	129.5	9
7.24	130	7
7.18	134.4	6

A.5.3 Other experimental and computed NMR parameters

TABLE A.10: Experimental and calculated vicinal $^{3}J_{HH}$ (Hz) scalar couplings of lorcaserin. Coupling values of each lorcaserin conformer were computed with the Haasnoot-Altona equation.

coupling	Exp.	**6A**	**6B**	**6C**	**6D**	**6E**	**6F**
H1–H2α	9.1	3.9	9.9	4.4	8.9	11.1	9.5
H1–H2β	1.0	1.7	1.4	1.4	1.2	6.7	8.3
H5α–H4α	1.7	2.2	0.6	0.5	1.4	9.3	6.3
H5α–H4β	7.9	12.1	5.5	6.2	11.4	1.1	12.0
H5β–H4β	1.9	0.8	1.7	1.3	0.5	10.5	4.1
H5β–H4α	9.6	5.0	11.7	11.4	6.2	7.4	1.1

Appendix B

Appendix B: DFT optimized structures and energies and RDC fit input and output files

B.1 MCI

B.1.1 DFT-Optimized XYZ coordinates (Å) of MCI conformers

MCI_1A WATER

C	-2.65696	-1.07740	-0.21549
C	-1.63842	-0.16530	-0.49683
C	-0.34247	-0.41130	-0.11072
C	0.01818	-1.47001	0.70655
C	-0.98852	-2.39251	1.00429
C	-2.28939	-2.21349	0.52245
C	0.53351	0.72338	-0.55165
C	1.48292	1.02506	0.62492
C	2.27259	-0.26708	0.91684
C	1.36982	-1.42313	1.37005
C	-0.53035	1.79249	-0.86973
C	-0.73888	2.81747	0.26426
C	-0.36747	2.28782	1.62483
C	0.69016	1.49473	1.81264
O	-3.90876	-0.77404	-0.64512
N	3.13890	-0.63188	-0.32204
C	1.35518	0.35512	-1.79451
C	2.25877	-0.82886	-1.54306
C	3.91040	-1.89287	-0.07505
O	-1.78893	1.06358	-1.08633
O	-2.04058	3.36061	0.21061
H	-0.77663	-3.24850	1.63945
H	-3.04553	-2.94970	0.77363
H	2.18636	1.82110	0.34891
H	3.02001	-0.07868	1.69419
H	1.21155	-1.25475	2.44305
H	1.89389	-2.38129	1.33536
H	-0.34123	2.31412	-1.81118
H	-0.05818	3.65453	0.04629
H	-0.95262	2.65817	2.46274
H	0.99260	1.18366	2.81040
C	-4.93422	-1.69772	-0.30969
C	4.15268	0.44509	-0.57798
H	0.68727	0.08468	-2.61853
H	1.92551	1.22675	-2.13606
H	1.67619	-1.73357	-1.34785
H	2.92768	-1.03030	-2.38223
H	4.62692	-2.01875	-0.88527
H	3.23785	-2.74614	-0.06061
H	4.43678	-1.80473	0.87513
H	-2.62891	2.61109	0.04183
H	-5.85427	-1.29316	-0.72871
H	-5.04156	-1.79927	0.77637
H	-4.74076	-2.68561	-0.74319
H	3.66730	1.37670	-0.85198
H	4.79186	0.12128	-1.39778
H	4.74651	0.58549	0.32441

MCI_1A DMSO

C	-2.65696	-1.07740	-0.21549
C	-1.63842	-0.16530	-0.49683
C	-0.34247	-0.41130	-0.11072
C	0.01818	-1.47001	0.70655
C	-0.98852	-2.39251	1.00429
C	-2.28939	-2.21349	0.52245
C	0.53351	0.72338	-0.55165
C	1.48292	1.02506	0.62492
C	2.27259	-0.26708	0.91684
C	1.36982	-1.42313	1.37005
C	-0.53035	1.79249	-0.86973
C	-0.73888	2.81747	0.26426
C	-0.36747	2.28782	1.62483
C	0.69016	1.49473	1.81264
O	-3.90876	-0.77404	-0.64512
N	3.13890	-0.63188	-0.32204
C	1.35518	0.35512	-1.79451
C	2.25877	-0.82886	-1.54306
C	3.91040	-1.89287	-0.07505
O	-1.78893	1.06358	-1.08633
O	-2.04058	3.36061	0.21061
H	-0.77663	-3.24850	1.63945
H	-3.04553	-2.94970	0.77363
H	2.18636	1.82110	0.34891
H	3.02001	-0.07868	1.69419
H	1.21155	-1.25475	2.44305
H	1.89389	-2.38129	1.33536
H	-0.34123	2.31412	-1.81118
H	-0.05818	3.65453	0.04629
H	-0.95262	2.65817	2.46274
H	0.99260	1.18366	2.81040
C	-4.93422	-1.69772	-0.30969
C	4.15268	0.44509	-0.57798
H	0.68727	0.08468	-2.61853
H	1.92551	1.22675	-2.13606
H	1.67619	-1.73357	-1.34785
H	2.92768	-1.03030	-2.38223
H	4.62692	-2.01875	-0.88527
H	3.23785	-2.74614	-0.06061
H	4.43678	-1.80473	0.87513
H	-2.62891	2.61109	0.04183
H	-5.85427	-1.29316	-0.72871
H	-5.04156	-1.79927	0.77637
H	-4.74076	-2.68561	-0.74319
H	3.66730	1.37670	-0.85198
H	4.79186	0.12128	-1.39778
H	4.74651	0.58549	0.32441

MCI_1B WATER

C	-2.61745	-1.10543	-0.18396
C	-1.61603	-0.17821	-0.47714
C	-0.32160	-0.37184	-0.05731
C	0.04177	-1.38622	0.81579
C	-0.94734	-2.32338	1.12641
C	-2.23981	-2.20241	0.60561
C	0.53893	0.75142	-0.57045
C	1.51214	1.10846	0.57112
C	2.27812	-0.18441	0.92519
C	1.37028	-1.26398	1.51001
C	-0.55364	1.78909	-0.90894
C	-0.79729	2.83622	0.20191
C	-0.33197	2.39966	1.56474
C	0.75849	1.65254	1.75057
O	-3.86390	-0.85264	-0.65831
N	3.10248	-0.68001	-0.33745
C	1.34438	0.32574	-1.82771
C	2.82150	0.18901	-1.55058
C	2.82406	-2.10644	-0.70704
O	-1.78809	1.02036	-1.12011
O	-2.13714	3.28167	0.18716
H	-0.72766	-3.14524	1.80269
H	-2.98153	-2.94944	0.86717
H	2.23635	1.86742	0.24534
H	3.05396	0.04595	1.65798
H	1.18575	-0.94203	2.54363
H	1.89408	-2.21647	1.62042
H	-0.37716	2.29722	-1.85927
H	-0.19165	3.71345	-0.07198
H	-0.87626	2.81629	2.40864
H	1.12360	1.43508	2.75229
C	-4.87065	-1.79808	-0.32567
C	4.55960	-0.56229	-0.01670
H	0.93784	-0.59969	-2.24878
H	1.24947	1.07936	-2.61478
H	3.35627	-0.26593	-2.38573
H	3.28764	1.15488	-1.35116
H	3.36701	-2.32373	-1.62604
H	1.75523	-2.24599	-0.86206
H	3.17833	-2.76339	0.08362
H	-2.67542	2.49306	0.02963
H	-5.78813	-1.43838	-0.78895
H	-5.01182	-1.86722	0.75902
H	-4.62977	-2.79265	-0.71828
H	4.78709	-1.20946	0.82945
H	4.78362	0.47446	0.23225
H	5.13804	-0.86911	-0.88787

MCI_1B DMSO

C	-2.61709	-1.10571	-0.18399
C	-1.61583	-0.17828	-0.47715

C	-0.32136	-0.37178	-0.05741
C	0.04220	-1.38623	0.81557
C	-0.94669	-2.32362	1.12610
C	-2.23917	-2.20277	0.60536
C	0.53892	0.75172	-0.57047
C	1.51207	1.10885	0.57115
C	2.27815	-0.18400	0.92510
C	1.37050	-1.26355	1.51006
C	-0.55397	1.78920	-0.90880
C	-0.79777	2.83622	0.20214
C	-0.33228	2.39963	1.56494
C	0.75833	1.65273	1.75069
O	-3.86355	-0.85291	-0.65802
N	3.10224	-0.67997	-0.33754
C	1.34434	0.32624	-1.82786
C	2.82145	0.18874	-1.55094
C	2.82335	-2.10637	-0.70673
O	-1.78812	1.02026	-1.11995
O	-2.13756	3.28150	0.18743
H	-0.72692	-3.14546	1.80238
H	-2.98076	-2.94989	0.86697
H	2.23636	1.86782	0.24546
H	3.05411	0.04643	1.65778
H	1.18574	-0.94130	2.54357
H	1.89473	-2.21579	1.62080
H	-0.37772	2.29748	-1.85911
H	-0.19224	3.71361	-0.07162
H	-0.87669	2.81605	2.40883
H	1.12348	1.43520	2.75240
C	-4.87024	-1.79855	-0.32587
C	4.55933	-0.56243	-0.01681
H	0.93742	-0.59888	-2.24921
H	1.24970	1.08007	-2.61478
H	3.35581	-0.26678	-2.38606
H	3.28833	1.15431	-1.35183
H	3.36647	-2.32430	-1.62550
H	1.75446	-2.24546	-0.86187
H	3.17698	-2.76316	0.08436
H	-2.67591	2.49297	0.02952
H	-5.78763	-1.43876	-0.78923
H	-5.01163	-1.86800	0.75879
H	-4.62920	-2.79296	-0.71878
H	4.78655	-1.20923	0.82976
H	4.78338	0.47448	0.23161
H	5.13793	-0.86977	-0.88772

MCI_1C WATER

C	-2.56675	-1.04477	-0.16835
C	-1.54490	-0.18090	-0.55741
C	-0.28529	-0.28051	-0.01173
C	0.03518	-1.15321	1.01316
C	-0.97388	-2.03523	1.41163
C	-2.23833	-1.99462	0.81368
C	0.58527	0.77654	-0.63957
C	1.61312	1.22925	0.39837
C	2.31521	-0.05206	0.90430
C	1.37035	-0.96739	1.68029
C	-0.52991	1.75978	-1.03197
C	-1.00343	2.70161	0.12545
C	-0.18034	2.63291	1.37480
C	0.98008	1.99289	1.51626
O	-3.78567	-0.87128	-0.73971
N	3.02866	-0.77951	-0.31254
C	1.32940	0.22491	-1.88451
C	2.79520	-0.02984	-1.61450
C	2.59343	-2.20190	-0.51281
O	-1.64776	0.89411	-1.41316
O	-2.34533	2.45216	0.51354
H	-0.79319	-2.75135	2.20874
H	-2.99833	-2.68829	1.15654
H	2.36950	1.88157	-0.05768
H	3.14545	0.22460	1.55744
H	1.85370	-1.91056	1.94947
H	1.21759	-0.45795	2.64140
H	-0.29031	2.34162	-1.92432
H	-0.92135	3.73241	-0.26277
H	-0.58864	3.19957	2.20880
H	1.52050	2.04802	2.45952
C	-4.82181	-1.73893	-0.30175
C	4.50209	-0.78404	-0.04651
H	0.83592	-0.67844	-2.25642
H	1.28917	0.95093	-2.70252
H	3.26035	-0.62757	-2.40018
H	3.36112	0.89916	-1.53613
H	1.51649	-2.24483	-0.66503
H	2.87480	-2.79925	0.35060
H	3.10856	-2.58212	-1.39387
H	-2.82446	2.20992	-0.29014
H	-5.71361	-1.44673	-0.85392
H	-5.00630	-1.63109	0.77329
H	-4.58525	-2.78643	-0.52085
H	4.69404	-1.33885	0.87138
H	4.84372	0.24489	0.06052
H	5.01185	-1.25948	-0.88451

MCI_1C DMSO

C	-2.56688	-1.04476	-0.16852
C	-1.54493	-0.18103	-0.55764
C	-0.28533	-0.28084	-0.01203
C	0.03520	-1.15389	1.01250
C	-0.97384	-2.03593	1.41087
C	-2.23838	-1.99489	0.81320
C	0.58519	0.77637	-0.63960
C	1.61298	1.22913	0.39845
C	2.31505	-0.05218	0.90434
C	1.37022	-0.96795	1.67987
C	-0.53003	1.75961	-1.03196
C	-1.00373	2.70124	0.12556
C	-0.18071	2.63243	1.37493
C	0.97976	1.99257	1.51641
O	-3.78580	-0.87081	-0.73942
N	3.02912	-0.77909	-0.31246
C	1.32926	0.22456	-1.88449
C	2.79512	-0.02990	-1.61468
C	2.59497	-2.20186	-0.51245
O	-1.64769	0.89400	-1.41336
O	-2.34540	2.45167	0.51369
H	-0.79322	-2.75212	2.20793
H	-2.99853	-2.68830	1.15624
H	2.36950	1.88146	-0.05751
H	3.14521	0.22427	1.55774
H	1.21730	-0.45876	2.64109
H	1.85377	-1.91114	1.94883
H	-0.29037	2.34160	-1.92424
H	-0.92156	3.73214	-0.26249
H	-0.58937	3.19860	2.20905
H	1.51993	2.04749	2.45982
C	-4.82227	-1.73778	-0.30101
C	4.50245	-0.78171	-0.04609
H	0.83592	-0.67909	-2.25584
H	1.28866	0.95009	-2.70293
H	3.26015	-0.62786	-2.40029
H	3.36107	0.89914	-1.53680
H	1.51787	-2.24556	-0.66368
H	2.87747	-2.79920	0.35062
H	3.10965	-2.58164	-1.39399
H	-2.82481	2.20927	-0.28977
H	-5.71401	-1.44525	-0.85308
H	-5.00647	-1.62946	0.77404
H	-4.58629	-2.78547	-0.1987
H	4.69463	-1.33351	0.87358
H	4.84293	0.24791	0.05824
H	5.01318	-1.25892	-0.88253

B.1.2 DFT energies of conformers

TABLE B.1: Gibbs free energies of the methylcodeine conformations optimized in (IEFPCM) water, in a.u. units.

Structure	$\Delta G_{298.15K}$ $(kcal/mol)$	$G_{298.15K}$ (a.u.)	SCF Energy $(a.u.)$
1A	0.00	−1018.24090	−1018.60499
1B	7.00	−1018.22978	−1018.59344
1C	7.60	−1018.22873	−1018.59190

TABLE B.2: Gibbs free energies of the methylcodeine conformations optimized in (IEFPCM) DMSO, in a.u. units.

Structure	$\Delta G_{298.15K}$ $(kcal/mol)$	$G_{298.15K}$ (a.u.)	SCF Energy $(a.u.)$
1A	0.00	−1018.24009	−1018.60420
1B	6.90	−1018.22905	−1018.59264
1C	7.70	−1018.22788	−1018.59110

B.1.3 RdcFit formatted RDC input file

RDC input files corresponding to the fully swollen gels.

H-gel swollen in D_2O (*H*-gel/*w*)

```
#RDC from F1-HSQC experiments. Using HP gels
rdc_data {
  5 22  -29.92
#  6 23   -6.14
 11 28   -6.83
 12 29   -3.94
 13 30   14.88
 14 31   -1.74
  9 25   16.16
  8 24   -5.11
#equivalent
 10 26    1.67
 10 27    1.67
#equivalent
```

```
 17 34    5.15
 17 35    5.15
#equivalent
 18 36    2.75
 18 37    2.75
#methyl
# 32 42   -0.17
# 32 43   -0.17
# 32 44   -0.17
#methyl
 33 45    0.54
 33 46    0.54
 33 47    0.54
#methyl
 19 38    9.86
 19 39    9.86
 19 40    9.86
}

standard_error {
1
}

penalty_function {
q
}

gridpoints {
16
}

bootstrapping {
false
}

bootstrappingpoints {
8
}
```

```
optimize_populations {
false
}

populations {
1.0
}

adaptative_grid {
false
}

print_penalty_histogram {
true
}

print_populations_histogram {
false
}
```

***Na*-gel swollen in D_2O (*Na*-gel/*w*)**

```
#RDC from F1-HSQC experiments. Using HP gels
rdc_data {
  5 22  -36.82
#  6 23   -9.98
 11 28   -8.83
 12 29   -7.75
 13 30   17.61
 14 31    0.40
  9 25   23.65
  8 24   -9.27
#equivalent
 10 26    2.56
 10 27    2.56
#equivalent
 17 34    6.79
```

```
 17 35    6.79
#equivalent
 18 36    1.30
 18 37    1.30
#methyl
# 32 42    0.92
# 32 43    0.92
# 32 44    0.92
#methyl
 33 45    1.67
 33 46    1.67
 33 47    1.67
#methyl
 19 38   11.89
 19 39   11.89
 19 40   11.89
}

standard_error {
1
}

penalty_function {
q
}

gridpoints {
16
}

bootstrapping {
false
}

bootstrappingpoints {
8
}

optimize_populations {
```

```
false
}

populations {
1.0
}

adaptative_grid {
false
}

print_penalty_histogram {
true
}

print_populations_histogram {
false
}
```

_H_-gel swollen in 500 mM NaCl D_2O (_H_-gel/_s1_)

```
#RDC from F1-HSQC experiments. Using HP gels
rdc_data {
  5 22  -25.39
#  6 23   -2.28
 11 28    0.02
 12 29   -0.34
 13 30   11.73
 14 31    0.24
  9 25   11.50
  8 24   -0.86
#equivalent
 10 26    1.80
 10 27    1.80
#equivalent
 17 34    5.69
 17 35    5.69
#equivalent
```

```
 18 36    4.99
 18 37    4.99
#methyl
# 32 42   -0.62
# 32 43   -0.62
# 32 44   -0.62
#methyl
 33 45   -0.29
 33 46   -0.29
 33 47   -0.29
#methyl
 19 38    8.01
 19 39    8.01
 19 40    8.01
}

standard_error {
1
}

penalty_function {
q
}

gridpoints {
16
}

bootstrapping {
false
}

bootstrappingpoints {
8
}

optimize_populations {
false
}
```

```
populations {
1.0
}

adaptative_grid {
false
}

print_penalty_histogram {
true
}

print_populations_histogram {
false
}
```

***Na*-gel swollen in 500 mM NaCl D_2O (*Na*-gel/*s1*)**

```
#RDC from F1-HSQC experiments. Using HP gels
rdc_data {
  5 22  -31.46
#  6 23  -11.86
 11 28    1.76
 12 29   -2.03
 13 30   14.20
 14 31    3.99
  9 25   14.86
  8 24   -1.95
#equivalent
 10 26    3.62
 10 27    3.62
#equivalent
 17 34    7.54
 17 35    7.54
#equivalent
```

```
 18 36    4.25
 18 37    4.25
#methyl
# 32 42    0.11
# 32 43    0.11
# 32 44    0.11
#methyl
 33 45    0.30
 33 46    0.30
 33 47    0.30
#methyl
 19 38   10.87
 19 39   10.87
 19 40   10.87
}

standard_error {
1
}

penalty_function {
q
}

gridpoints {
16
}

bootstrapping {
false
}

bootstrappingpoints {
8
}

optimize_populations {
false
}
```

```
populations {
1.0
}

adaptative_grid {
false
}

print_penalty_histogram {
true
}

print_populations_histogram {
false
}
```

***H*-gel swollen in 1.0 M NaCl D_2O (*H*-gel/*s2*)**

```
#RDC from F1-HSQC experiments. Using HP gels
rdc_data {
  5 22  -21.01
#  6 23   -6.09
 11 28    0.10
 12 29    0.67
 13 30   11.43
 14 31   -0.06
  9 25   10.32
  8 24   -0.34
#equivalent
 10 26    1.71
 10 27    1.71
#equivalent
 17 34    5.39
 17 35    5.39
#equivalent
```

```
 18 36    5.24
 18 37    5.24
#methyl
# 32 42   -0.61
# 32 43   -0.61
# 32 44   -0.61
#methyl
 33 45   -0.11
 33 46   -0.11
 33 47   -0.11
#methyl
 19 38    8.64
 19 39    8.64
 19 40    8.64
}

standard_error {
1
}

penalty_function {
q
}

gridpoints {
16
}

bootstrapping {
false
}

bootstrappingpoints {
8
}

optimize_populations {
false
}
```

```
populations {
1.0
}

adaptative_grid {
false
}

print_penalty_histogram {
true
}

print_populations_histogram {
false
}
```

***Na*-gel swollen in 1.0 M NaCl D_2O (*Na*-gel/*s2*)**

```
#RDC from F1-HSQC experiments. Using HP gels
rdc_data {
  5 22  -21.34
#  6 23   -8.45
 11 28    3.87
 12 29   -1.12
 13 30   13.36
 14 31    4.19
  9 25   12.79
  8 24   -1.51
#equivalent
 10 26    3.72
 10 27    3.72
#equivalent
 17 34    6.54
 17 35    6.54
#equivalent
 18 36    5.47
 18 37    5.47
```

```
#methyl
# 32 42    0.11
# 32 43    0.11
# 32 44    0.11
#methyl
 33 45    0.51
 33 46    0.51
 33 47    0.51
#methyl
 19 38   10.01
 19 39   10.01
 19 40   10.01
}

standard_error {
1
}

penalty_function {
q
}

gridpoints {
16
}

bootstrapping {
false
}

bootstrappingpoints {
8
}

optimize_populations {
false
}

populations {
```

```
1.0
}

adaptative_grid {
false
}

print_penalty_histogram {
true
}

print_populations_histogram {
false
}
```

***H*-gel swollen in** $1:1$ **D_2O:DMSO-d_6 (*H*-gel/*m*)**

```
#RDC from F1-HSQC experiments. Using HP gels
rdc_data {
  5 22   -1.86
#  6 23    5.72
 11 28   -1.73
 12 29    1.63
 13 30    1.48
 14 31   -1.38
  9 25    5.86
  8 24    1.21
#equivalent
 10 26   -1.07
 10 27   -1.07
#equivalent
 17 34    3.53
 17 35    3.53
#equivalent
 18 36    1.08
 18 37    1.08
#methyl
# 32 42   -0.61
```

```
# 32 43   -0.61
# 32 44   -0.61
#methyl
 33 45   -0.80
 33 46   -0.80
 33 47   -0.80
#methyl
 19 38    2.92
 19 39    2.92
 19 40    2.92
}

standard_error {
1
}

penalty_function {
q
}

gridpoints {
16
}

bootstrapping {
false
}

bootstrappingpoints {
8
}

optimize_populations {
false
}

populations {
1.0
}
```

```
adaptative_grid {
false
}

print_penalty_histogram {
true
}

print_populations_histogram {
false
}
```

***Na*-gel swollen in** 1 : 1 **D_2O:DMSO-d_6 (*Na*-gel/*m*)**

```
#RDC from F1-HSQC experiments. Using HP gels
rdc_data {
  5 22   -4.73
#  6 23    4.88
 11 28   -3.13
 12 29    1.66
 13 30    1.72
 14 31   -2.76
  9 25    8.69
  8 24    1.89
#equivalent
 10 26   -1.69
 10 27   -1.69
#equivalent
 17 34    5.00
 17 35    5.00
#equivalent
 18 36    1.27
 18 37    1.27
#methyl
# 32 42   -0.70
# 32 43   -0.70
# 32 44   -0.70
```

```
#methyl
# 33 45   ALERT
# 33 46   ALERT
# 33 47   ALERT
#methyl
 19 38    3.87
 19 39    3.87
 19 40    3.87
}

standard_error {
1
}

penalty_function {
q
}

gridpoints {
16
}

bootstrapping {
false
}

bootstrappingpoints {
8
}

optimize_populations {
false
}

populations {
1.0
}

adaptative_grid {
```

```
false
}

print_penalty_histogram {
true
}

print_populations_histogram {
false
}
```

_Na_-gel swollen in DMSO-d_6 (_Na_-gel/_d_)

```
#RDC from F1-HSQC experiments. Using HP gels
rdc_data {
  5 22   -2.47
#  6 23   -1.20
 11 28   -3.09
 12 29    2.09
 13 30   -0.78
 14 31   -2.83
  9 25    3.91
  8 24    1.83
#equivalent
 10 26   -2.56
 10 27   -2.56
#equivalent
 17 34    0.39
 17 35    0.39
#equivalent
 18 36   -0.65
 18 37   -0.65
#methyl
# 32 42   -0.21
# 32 43   -0.21
# 32 44   -0.21
#methyl
 33 45   -0.62
```

```
 33 46   -0.62
 33 47   -0.62
#methyl
 19 38   -0.07
 19 39   -0.07
 19 40   -0.07
}

standard_error {
1
}

penalty_function {
q
}

gridpoints {
16
}

bootstrapping {
false
}

bootstrappingpoints {
8
}

optimize_populations {
false
}

populations {
1.0
}

adaptative_grid {
false
}
```

```
print_penalty_histogram {
true
}

print_populations_histogram {
false
}
```

B.1.4 RDC fit output parameters

H-gel swollen in D_2O (*H*-gel/*w*)

```
Alignment tensor
 7.22367994e-05  -1.80585713e-04  -9.73478384e-06
-1.80585713e-04   1.82620002e-04  -3.44234694e-04
-9.73478384e-06  -3.44234694e-04  -2.54856802e-04
GDO (*1.0e6)= 637.124156664
Alignment tensor eigenvalues
 4.38085960e-04+0.j   2.39058093e-05+0.j  -4.61991769e-04+0.j
Alignment tensor eigenvectors
 0.39622228 -0.89925244  0.18534550
-0.82448712 -0.25964809  0.50279604
 0.40401596  0.35203397  0.84430041

 0.39622228 -0.82448712  0.40401596
-0.89925244 -0.25964809  0.35203397
 0.18534550  0.50279604  0.84430041
Ordered alignment tensor ( |Ax| < |Ay| < |Az| )
Ax = 2.391e-05
Ay = 4.381e-04
Az = -4.620e-04
Principal frame eigenvectors
X'= [-0.899,-0.260, 0.352]
Y'= [ 0.396,-0.824, 0.404]
Z'= [ 0.185, 0.503, 0.844]
```

Experimental and back calculated RDCs

I	J	*RDC(exp)*	*RDC(calc)*	*weight*	*stderr*
11	28	−6.83	−7.7	1	1
12	29	−3.94	−4.04	1	1
13	30	14.88	15.23	1	1
14	31	−1.74	−0.73	1	1
9	25	16.16	15.57	1	1
8	24	−5.11	−4.83	1	1
10	26	1.67	1.25	0.5	1
10	27	1.67	1.25	0.5	1
17	34	5.15	5.66	0.5	1
17	35	5.15	5.66	0.5	1
18	36	2.75	2.91	0.5	1
18	37	2.75	2.91	0.5	1
33	45	0.54	0.51	0.333	1
33	46	0.54	0.51	0.333	1
33	47	0.54	0.51	0.333	1
19	38	9.86	9.45	0.333	1
19	39	9.86	9.45	0.333	1
19	40	9.86	9.45	0.333	1

```
Fitting scores
RMSD =    0.404 Hz
Cornilescu Q factor =    0.064
Cornilescu Q factor with experimental errors =    0.064
Cornilescu Q factor,<r3> averaged =    0.064
Cornilescu Q factor <r3> averaged, with experimental errors =    0.064
Chi^2 =    2.945
Number of input rdcs =  18
Number of non equivalent rdcs= 11
n/Chi^2 3.73542707225
```

Na-gel swollen in D_2O (*Na*-gel/*w*)

```
Alignment tensor
  9.99236914e-05  -2.09430965e-04  -5.51274716e-05
 -2.09430965e-04   2.44259831e-04  -4.38353703e-04
 -5.51274716e-05  -4.38353703e-04  -3.44183523e-04
GDO (*1.0e6)= 816.222138075
Alignment tensor eigenvalues
 5.38229056e-04+0.j   7.12402894e-05+0.j  -6.09469345e-04+0.j
Alignment tensor eigenvectors
 0.35422005 -0.91136597  0.20961926
-0.84614374 -0.21689834  0.48682222
 0.39820714  0.34981021  0.84797635

 0.35422005 -0.84614374  0.39820714
-0.91136597 -0.21689834  0.34981021
```

```
 0.20961926  0.48682222  0.84797635
Ordered alignment tensor ( |Ax| < |Ay| < |Az| )
Ax = 7.124e-05
Ay = 5.382e-04
Az = -6.095e-04
Principal frame eigenvectors
X'= [-0.911,-0.217, 0.350]
Y'= [ 0.354,-0.846, 0.398]
Z'= [ 0.210, 0.487, 0.848]
```

Experimental and back calculated RDCs

I	J	*RDC(exp)*	*RDC(calc)*	*weight*	*stderr*
11	28	−8.83	−10.11	1	1
12	29	−7.75	−7.78	1	1
13	30	17.61	17.98	1	1
14	31	0.4	1.88	1	1
9	25	23.65	22.77	1	1
8	24	−9.27	−8.96	1	1
10	26	2.56	2.02	0.5	1
10	27	2.56	2.02	0.5	1
17	34	6.79	7.28	0.5	1
17	35	6.79	7.28	0.5	1
18	36	1.3	1.51	0.5	1
18	37	1.3	1.51	0.5	1
33	45	1.67	1.58	0.333	1
33	46	1.67	1.58	0.333	1
33	47	1.67	1.58	0.333	1
19	38	11.89	11.67	0.333	1
19	39	11.89	11.67	0.333	1
19	40	11.89	11.67	0.333	1

```
Fitting scores
RMSD =    0.552 Hz
Cornilescu Q factor =    0.065
Cornilescu Q factor with experimental errors =    0.065
Cornilescu Q factor,<r3> averaged =    0.065
Cornilescu Q factor <r3> averaged, with experimental errors =    0.065
Chi^2 =    5.477
Number of input rdcs =  18
Number of non equivalent rdcs= 11
n/Chi^2 2.00854242931
```

H-gel swollen in 500 mM NaCl D_2O (*H*-gel/*s1*)

```
Alignment tensor
  1.34786305e-04  -1.91135419e-04   1.60448981e-05
 -1.91135419e-04   1.16837724e-04  -2.52210601e-04
```

```
  1.60448981e-05  -2.52210601e-04  -2.51624029e-04
GDO (*1.0e6)= 543.997979986
Alignment tensor eigenvalues
  3.78335578e-04+0.j   1.23602338e-05+0.j  -3.90695812e-04+0.j
Alignment tensor eigenvectors
 0.59941389 -0.78626040  0.14999188
-0.73771253 -0.46993762  0.48470492
 0.31061746  0.40118975  0.86172129

 0.59941389 -0.73771253  0.31061746
-0.78626040 -0.46993762  0.40118975
 0.14999188  0.48470492  0.86172129
Ordered alignment tensor ( |Ax| < |Ay| < |Az| )
Ax = 1.236e-05
Ay = 3.783e-04
Az = -3.907e-04
Principal frame eigenvectors
X'= [-0.786,-0.470, 0.401]
Y'= [ 0.599,-0.738, 0.311]
Z'= [ 0.150, 0.485, 0.862]
```

Experimental and back calculated RDCs

I	J	*RDC(exp)*	*RDC(calc)*	*weight*	*stderr*
11	28	0.02	−1.04	1	1
12	29	−0.34	−0.1	1	1
13	30	11.73	12.18	1	1
14	31	0.24	1.47	1	1
9	25	11.5	10.47	1	1
8	24	−0.86	−0.69	1	1
10	26	1.8	1.24	0.5	1
10	27	1.8	1.24	0.5	1
17	34	5.69	6.03	0.5	1
17	35	5.69	6.03	0.5	1
18	36	4.99	4.33	0.5	1
18	37	4.99	4.33	0.5	1
33	45	−0.29	−0.39	0.333	1
33	46	−0.29	−0.39	0.333	1
33	47	−0.29	−0.39	0.333	1
19	38	8.01	8.48	0.333	1
19	39	8.01	8.48	0.333	1
19	40	8.01	8.48	0.333	1

```
Fitting scores
RMSD =    0.531 Hz
Cornilescu Q factor =    0.113
Cornilescu Q factor with experimental errors =    0.113
Cornilescu Q factor,<r3> averaged =    0.113
```

```
Cornilescu Q factor <r3> averaged, with experimental errors =    0.113
Chi^2 =    5.067
Number of input rdcs =  18
Number of non equivalent rdcs= 11
n/Chi^2 2.1707641293
```

_Na_-gel swollen in 500 mM NaCl D_2O (_Na_-gel/_s1_)

```
Alignment tensor
  2.37313528e-04  -2.36980361e-04  -3.60039991e-06
 -2.36980361e-04   1.23505816e-04  -3.15416496e-04
 -3.60039991e-06  -3.15416496e-04  -3.60819344e-04
GDO (*1.0e6)= 716.297728303
Alignment tensor eigenvalues
  4.75531114e-04+0.j   5.70968465e-05+0.j  -5.32627961e-04+0.j
Alignment tensor eigenvectors
 0.67976691 -0.71808380  0.14924007
-0.68720890 -0.55251127  0.47167279
 0.25624377  0.42318666  0.86905246

 0.67976691 -0.68720890  0.25624377
-0.71808380 -0.55251127  0.42318666
 0.14924007  0.47167279  0.86905246
Ordered alignment tensor ( |Ax| < |Ay| < |Az| )
Ax = 5.710e-05
Ay = 4.755e-04
Az = -5.326e-04
Principal frame eigenvectors
X'= [-0.718,-0.553, 0.423]
Y'= [ 0.680,-0.687, 0.256]
Z'= [ 0.149, 0.472, 0.869]
```

Experimental and back calculated RDCs

I	J	*RDC*(*exp*)	*RDC*(*calc*)	*weight*	*stderr*
11	28	1.76	0.78	1	1
12	29	−2.03	−1.43	1	1
13	30	14.2	14.61	1	1
14	31	3.99	5.19	1	1
9	25	14.86	14.09	1	1
8	24	−1.95	−2.44	1	1
10	26	3.62	2.9	0.5	1
10	27	3.62	2.9	0.5	1
17	34	7.54	8.02	0.5	1
17	35	7.54	8.02	0.5	1
18	36	4.25	4.43	0.5	1
18	37	4.25	4.43	0.5	1
33	45	0.3	0.21	0.333	1
33	46	0.3	0.21	0.333	1
33	47	0.3	0.21	0.333	1
19	38	10.87	10.54	0.333	1
19	39	10.87	10.54	0.333	1
19	40	10.87	10.54	0.333	1

```
Fitting scores
RMSD =    0.509 Hz
Cornilescu Q factor =    0.084
Cornilescu Q factor with experimental errors =    0.084
Cornilescu Q factor,<r3> averaged =    0.085
Cornilescu Q factor <r3> averaged, with experimental errors =    0.085
Chi^2 =    4.672
Number of input rdcs =  18
Number of non equivalent rdcs= 11
n/Chi^2 2.35420169465
```

H-gel swollen in 1.0 M NaCl D_2O (*H*-gel/*s*2)

```
Alignment tensor
  1.39978987e-04  -1.93379860e-04   2.30779252e-05
 -1.93379860e-04   1.04261311e-04  -2.47065094e-04
  2.30779252e-05  -2.47065094e-04  -2.44240298e-04
GDO (*1.0e6)= 536.709329092
Alignment tensor eigenvalues
  3.76610907e-04+0.j   5.73480217e-06+0.j  -3.82345709e-04+0.j
Alignment tensor eigenvectors
 0.61956131 -0.77150703  0.14463986
-0.72113885 -0.48667611  0.49305692
 0.31000411  0.40978441  0.85788938

 0.61956131 -0.72113885  0.31000411
-0.77150703 -0.48667611  0.40978441
```

```
 0.14463986  0.49305692  0.85788938
Ordered alignment tensor ( |Ax| < |Ay| < |Az| )
Ax = 5.735e-06
Ay = 3.766e-04
Az = -3.823e-04
Principal frame eigenvectors
X'= [-0.772,-0.487, 0.410]
Y'= [ 0.620,-0.721, 0.310]
Z'= [ 0.145, 0.493, 0.858]
```

Experimental and back calculated RDCs

I	J	*RDC(exp)*	*RDC(calc)*	*weight*	*stderr*
11	28	0.1	−0.77	1	1
12	29	0.67	0.6	1	1
13	30	11.43	12.03	1	1
14	31	−0.06	0.99	1	1
9	25	10.32	9.55	1	1
8	24	−0.34	0.01	1	1
10	26	1.71	1.24	0.5	1
10	27	1.71	1.24	0.5	1
17	34	5.39	6.01	0.5	1
17	35	5.39	6.01	0.5	1
18	36	5.24	4.69	0.5	1
18	37	5.24	4.69	0.5	1
33	45	−0.11	−0.59	0.333	1
33	46	−0.11	−0.59	0.333	1
33	47	−0.11	−0.59	0.333	1
19	38	8.64	8.35	0.333	1
19	39	8.64	8.35	0.333	1
19	40	8.64	8.35	0.333	1

```
Fitting scores
RMSD =    0.479 Hz
Cornilescu Q factor =    0.105
Cornilescu Q factor with experimental errors =    0.105
Cornilescu Q factor,<r3> averaged =    0.106
Cornilescu Q factor <r3> averaged, with experimental errors =    0.106
Chi^2 =    4.138
Number of input rdcs =  18
Number of non equivalent rdcs= 11
n/Chi^2 2.65842313837
```

Na-gel swollen in 1.0 M NaCl D_2O (*Na*-gel/*s2*)

```
Alignment tensor
  2.35716912e-04  -2.22989812e-04   1.31998856e-05
 -2.22989812e-04   1.12394144e-04  -2.75616460e-04
```

```
  1.31998856e-05  -2.75616460e-04  -3.48111056e-04
GDO (*1.0e6)= 664.155452967
Alignment tensor eigenvalues
  4.49759322e-04+0.j   3.74915405e-05+0.j  -4.87250862e-04+0.j
Alignment tensor eigenvectors
 0.70721604 -0.69617011  0.12325846
-0.66455752 -0.59508807  0.45192200
 0.24126496  0.40151882  0.88350091

 0.70721604 -0.66455752  0.24126496
-0.69617011 -0.59508807  0.40151882
 0.12325846  0.45192200  0.88350091
Ordered alignment tensor ( |Ax| < |Ay| < |Az| )
Ax = 3.749e-05
Ay = 4.498e-04
Az = -4.873e-04
Principal frame eigenvectors
X'= [-0.696,-0.595, 0.402]
Y'= [ 0.707,-0.665, 0.241]
Z'= [ 0.123, 0.452, 0.884]
```

Experimental and back calculated RDCs

I	J	*RDC*(*exp*)	*RDC*(*calc*)	*weight*	*stderr*
11	28	3.87	2.76	1	1
12	29	−1.12	−0.78	1	1
13	30	13.36	14.07	1	1
14	31	4.19	5.53	1	1
9	25	12.79	11.7	1	1
8	24	−1.51	−1.57	1	1
10	26	3.72	2.92	0.5	1
10	27	3.72	2.92	0.5	1
17	34	6.54	7.31	0.5	1
17	35	6.54	7.31	0.5	1
18	36	5.47	4.82	0.5	1
18	37	5.47	4.82	0.5	1
33	45	0.51	0.08	0.333	1
33	46	0.51	0.08	0.333	1
33	47	0.51	0.08	0.333	1
19	38	10.01	9.85	0.333	1
19	39	10.01	9.85	0.333	1
19	40	10.01	9.85	0.333	1

```
Fitting scores
RMSD =    0.610 Hz
Cornilescu Q factor =    0.109
Cornilescu Q factor with experimental errors =    0.109
Cornilescu Q factor,<r3> averaged =    0.109
```

```
Cornilescu Q factor <r3> averaged, with experimental errors =    0.109
Chi^2 =    6.708
Number of input rdcs =  18
Number of non equivalent rdcs= 11
n/Chi^2 1.63978166181
```

***H*-gel swollen in** 1 : 1 **D_2O:DMSO-d_6 (*H*-gel/*m*)**

```
Alignment tensor
  1.84104220e-05  -8.88369326e-05  -1.90513458e-05
 -8.88369326e-05   3.67798619e-05  -9.05406718e-05
 -1.90513458e-05  -9.05406718e-05  -5.51902839e-05
GDO (*1.0e6)= 194.017536757
Alignment tensor eigenvalues
  1.35221474e-04+0.j   3.85790707e-06+0.j  -1.39079381e-04+0.j
Alignment tensor eigenvectors
-0.54101570  0.73492480  0.40888575
 0.77922900  0.25513521  0.57245802
-0.31639244 -0.62832441  0.71070687

-0.54101570  0.77922900 -0.31639244
 0.73492480  0.25513521 -0.62832441
 0.40888575  0.57245802  0.71070687
Ordered alignment tensor ( |Ax| < |Ay| < |Az| )
Ax = 3.858e-06
Ay = 1.352e-04
Az = -1.391e-04
Principal frame eigenvectors
X'= [ 0.735, 0.255,-0.628]
Y'= [-0.541, 0.779,-0.316]
Z'= [ 0.409, 0.572, 0.711]
```

Experimental and back calculated RDCs

I	J	$RDC(exp)$	$RDC(calc)$	*weight*	*stderr*
11	28	−1.73	−2.32	1	1
12	29	1.63	1.73	1	1
13	30	1.48	1.58	1	1
14	31	−1.38	−0.7	1	1
9	25	5.86	5.5	1	1
8	24	1.21	1.37	1	1
10	26	−1.07	−1.19	0.5	1
10	27	−1.07	−1.19	0.5	1
17	34	3.53	3.45	0.5	1
17	35	3.53	3.45	0.5	1
18	36	1.08	1.09	0.5	1
18	37	1.08	1.09	0.5	1
33	45	−0.8	−0.85	0.333	1
33	46	−0.8	−0.85	0.333	1
33	47	−0.8	−0.85	0.333	1
19	38	2.92	3.12	0.333	1
19	39	2.92	3.12	0.333	1
19	40	2.92	3.12	0.333	1

```
Fitting scores
RMSD =    0.242 Hz
Cornilescu Q factor =    0.123
Cornilescu Q factor with experimental errors =    0.123
Cornilescu Q factor,<r3> averaged =    0.122
Cornilescu Q factor <r3> averaged, with experimental errors =    0.122
Chi^2 =    1.056
Number of input rdcs =  18
Number of non equivalent rdcs= 11
n/Chi^2 10.4165028349
```

Na-gel swollen in 1 : 1 D_2O:DMSO-d_6 (*Na*-gel/*m*)

```
Alignment tensor
  6.59120055e-06  -1.20071531e-04  -3.06513521e-05
 -1.20071531e-04   5.77140635e-05  -1.27766001e-04
 -3.06513521e-05  -1.27766001e-04  -6.43052641e-05
GDO (*1.0e6)= 266.216419068
Alignment tensor eigenvalues
 -1.90999401e-04+0.j   5.63862677e-06+0.j   1.85360774e-04+0.j
Alignment tensor eigenvectors
 0.45299316  0.75165149  0.47939257
 0.57058382  0.16873025 -0.80371898
 0.68500460 -0.63761285  0.35244653

 0.45299316  0.57058382  0.68500460
 0.75165149  0.16873025 -0.63761285
```

```
 0.47939257 -0.80371898  0.35244653
Ordered alignment tensor ( |Ax| < |Ay| < |Az| )
Ax = 5.639e-06
Ay = 1.854e-04
Az = -1.910e-04
Principal frame eigenvectors
X'= [ 0.752, 0.169,-0.638]
Y'= [ 0.479,-0.804, 0.352]
Z'= [ 0.453, 0.571, 0.685]
```

Experimental and back calculated RDCs

I	J	*RDC(exp)*	*RDC(calc)*	*weight*	*stderr*
11	28	−3.13	−4.17	1	1
12	29	1.66	2.28	1	1
13	30	1.72	1.92	1	1
14	31	−2.76	−1.52	1	1
9	25	8.69	8.07	1	1
8	24	1.89	1.79	1	1
10	26	−1.69	−2	0.5	1
10	27	−1.69	−2	0.5	1
17	34	5	4.71	0.5	1
17	35	5	4.71	0.5	1
18	36	1.27	1.27	0.5	1
18	37	1.27	1.27	0.5	1
33	45	−1.16	−1.21	0.333	1
33	46	−1.16	−1.21	0.333	1
33	47	−1.16	−1.21	0.333	1
19	38	3.87	4.25	0.333	1
19	39	3.87	4.25	0.333	1
19	40	3.87	4.25	0.333	1

```
Fitting scores
RMSD =    0.457 Hz
Cornilescu Q factor =    0.159
Cornilescu Q factor with experimental errors =    0.159
Cornilescu Q factor,<r3> averaged =    0.158
Cornilescu Q factor <r3> averaged, with experimental errors =    0.158
Chi^2 =    3.755
Number of input rdcs =  18
Number of non equivalent rdcs= 11
n/Chi^2 2.92905038892
```

Na-gel swollen in DMSO-d_6 (*Na*-gel/*d*)

```
Alignment tensor
 -1.03156861e-04  -2.34532453e-05  -1.31882030e-05
 -2.34532453e-05   5.89961165e-05  -9.10263080e-06
```

```
 -1.31882030e-05  -9.10263080e-06   4.41607448e-05
GDO (*1.0e6)= 132.987720436
Alignment tensor eigenvalues
 -1.07835570e-04+0.j   4.28957972e-05+0.j   6.49397731e-05+0.j
Alignment tensor eigenvectors
-0.98515008 -0.13714022  0.10330480
-0.14362582  0.32855296 -0.93350125
-0.09407946  0.93447606  0.34337085

-0.98515008 -0.14362582 -0.09407946
-0.13714022  0.32855296  0.93447606
 0.10330480 -0.93350125  0.34337085
Ordered alignment tensor ( |Ax| < |Ay| < |Az| )
Ax = 4.290e-05
Ay = 6.494e-05
Az = -1.078e-04
Principal frame eigenvectors
X'= [-0.137, 0.329, 0.934]
Y'= [ 0.103,-0.934, 0.343]
Z'= [-0.985,-0.144,-0.094]
```

Experimental and back calculated RDCs

I	J	*RDC(exp)*	*RDC(calc)*	*weight*	*stderr*
11	28	−3.09	−3.52	1	1
12	29	2.09	1.34	1	1
13	30	−0.78	−1.27	1	1
14	31	−2.83	−2.52	1	1
9	25	3.91	3.08	1	1
8	24	1.83	1.56	1	1
10	26	−2.56	−2.92	0.5	1
10	27	−2.56	−2.92	0.5	1
17	34	0.39	1.01	0.5	1
17	35	0.39	1.01	0.5	1
18	36	−0.65	−0.05	0.5	1
18	37	−0.65	−0.05	0.5	1
33	45	−0.62	−0.91	0.333	1
33	46	−0.62	−0.91	0.333	1
33	47	−0.62	−0.91	0.333	1
19	38	−0.07	0.81	0.333	1
19	39	−0.07	0.81	0.333	1
19	40	−0.07	0.81	0.333	1

```
Fitting scores
RMSD =    0.446 Hz
Cornilescu Q factor =    0.271
Cornilescu Q factor with experimental errors =    0.271
Cornilescu Q factor,<r3> averaged =    0.271
```

```
Cornilescu Q factor <r3> averaged, with experimental errors =    0.271
Chi^2 =    3.576
Number of input rdcs =  18
Number of non equivalent rdcs= 11
n/Chi^2 3.07588320766
```

B.2 Salsolidine

B.2.1 DFT-minimized XYZ geometries

1A

C	1.708773	-0.637213	0.040780
C	1.474211	0.753683	-0.079130
C	0.172707	1.219638	-0.150042
C	-0.922494	0.334549	-0.111525
C	-0.692059	-1.032046	-0.017261
C	0.629816	-1.503235	0.068456
O	3.008451	-1.010309	0.112426
C	3.280463	-2.398503	0.234422
O	2.581210	1.533841	-0.114163
C	2.388946	2.935724	-0.232015
C	-2.320698	0.921636	-0.219546
N	-3.338099	-0.130435	0.160933
C	-3.112144	-1.446365	-0.530365
C	-1.822534	-2.039143	0.004848
C	-2.571081	2.142373	0.659924
H	-0.004687	2.280636	-0.244162
H	0.788549	-2.570231	0.149625
H	4.362692	-2.482777	0.285633
H	2.836825	-2.804631	1.146685
H	2.909644	-2.946383	-0.635501
H	3.384799	3.370078	-0.237108
H	1.876674	3.182597	-1.165254
H	1.824302	3.325298	0.618803
H	-2.552367	1.160330	-1.260924
H	-3.311998	-0.271652	1.175268
H	-4.273333	0.224920	-0.053962
H	-3.975878	-2.080302	-0.337925
H	-3.054100	-1.227184	-1.596444
H	-1.981990	-2.415946	1.020766
H	-1.568506	-2.903877	-0.610146
H	-2.299038	1.938895	1.697472
H	-3.617017	2.453019	0.612004
H	-1.978601	2.985428	0.311259

1B

C	1.702003	-0.648302	0.026532
C	1.509148	0.752035	-0.075720
C	0.222985	1.251516	-0.182308
C	-0.891164	0.390595	-0.185909
C	-0.705894	-0.979621	-0.088303
C	0.601872	-1.487519	0.013823
O	2.989590	-1.056024	0.129502
C	3.223808	-2.453564	0.215609
O	2.636950	1.501778	-0.061814
C	2.484407	2.908909	-0.176299
C	-2.261410	1.033873	-0.260899
N	-3.298183	-0.013579	-0.617843
C	-3.186550	-1.272314	0.203714
C	-1.869159	-1.947137	-0.129576
C	-2.675565	1.740598	1.027621
H	0.061777	2.318182	-0.270041
H	0.731495	-2.559082	0.086354
H	4.302460	-2.567501	0.282974
H	2.753457	-2.874114	1.107939
H	2.855022	-2.966505	-0.676084
H	3.490531	3.317766	-0.147771
H	2.008015	3.173817	-1.123324
H	1.903096	3.307631	0.658960
H	-2.290852	1.747728	-1.086598
H	-4.230410	0.393236	-0.505122
H	-3.208921	-0.260419	-1.606572
H	-3.243737	-0.972007	1.247748
H	-4.045632	-1.896753	-0.034683
H	-1.717524	-2.754655	0.588583
H	-1.933929	-2.417697	-1.116527
H	-2.637185	1.072849	1.887861
H	-3.673208	2.177511	0.944806
H	-1.976383	2.554622	1.215183

1C

C	1.706701	-0.613959	-0.032782
C	1.440364	0.775710	-0.006076
C	0.130822	1.227907	-0.102994
C	-0.921349	0.310974	-0.241521
C	-0.660506	-1.054047	-0.269397
C	0.655596	-1.514293	-0.153244
O	3.011300	-0.961031	0.067825
C	3.318623	-2.347288	0.071524
O	2.527588	1.574690	0.117434
C	2.305866	2.975973	0.185434
C	-2.362419	0.723427	-0.419265
N	-3.202870	-0.099045	0.560271
C	-2.944348	-1.603774	0.564577
C	-1.842043	-1.973687	-0.427722
C	-2.685101	2.191221	-0.199901
H	-0.074960	2.287869	-0.081640
H	0.842412	-2.579287	-0.168379
H	4.399221	-2.408233	0.168332
H	2.845943	-2.849311	0.919050
H	3.005395	-2.819290	-0.863114
H	3.288975	3.424329	0.297247
H	1.838114	3.342943	-0.731665
H	1.687384	3.232074	1.049023
H	-2.719628	0.410344	-1.403691
H	-3.032033	0.273219	1.495707
H	-4.190702	0.078475	0.364349
H	-2.651081	-1.864533	1.578446
H	-3.885213	-2.096131	0.329396
H	-1.568599	-3.012921	-0.243639
H	-2.232836	-1.925337	-1.448839
H	-2.351987	2.533370	0.782417
H	-3.756005	2.378353	-0.295345
H	-2.188642	2.789135	-0.963418

B.2.2 DFT energies of conformers

TABLE B.3: Gibbs free energies of the lorcaserin conformations.

Structure	$\Delta G_{298.15K}$ $(kcal/mol)$	$G_{298.15K}$ $(a.u.)$	SCF Energy $(a.u.)$
2A	0.00	−673.04433	−672.94275
2B	0.10	−673.04415	−672.94241
2C	1.90	−673.04124	−672.94138

B.2.3 Mspin formatted RDC input file

```
rdc_data {
#Me (C1)
15 31 5.1
15 32 5.1
15 33 5.1
#C1H
11 24 -19.1
#C3
13 27 3.36
13 28 7.26
#C4
14 29 -18.34
14 30 -11.04
#C5
6 17 -30.39
#C8
3 16  -30.41
}

permutations {
27 28
29 30
}

grid {
256
}
```

B.2.4 RDC fit output parameters

Fit parameters for the ensemble 2A+2B 50% populated for each conformer.

```
Conformational averaged solution
Alignment tensor
A'x= 1.241e-04
A'y= 4.398e-04
A'z=-5.640e-04
Saupe tensor
S'x= 1.862e-04
S'y= 6.597e-04
S'z=-8.459e-04
Alignment tensor eigenvectors
e[x]=( 0.116, 0.308, 0.944)
e[y]=( 0.106,-0.949, 0.297)
e[z]=( 0.988, 0.066,-0.143)

Alignment tensor in laboratory coordinates:
[-5.434e-04,-7.635e-05,1.070e-04]
[-7.635e-05,4.055e-04,-8.254e-05]
[ 1.070e-04,-8.254e-05,1.379e-04]

SVD condition number is 6.364e+00
Axial component Aa = -8.459e-04
Rhombic component Ar = -3.157e-04
rhombicity R = 3.732e-01
Asymmetry parameter etha =5.598e-01
GDO = 1.051e-03

Euler Angles (degrees)
Set 1
(155.3,-81.0,42.4)
Set 2
(-24.7,261.0,-137.6)
```

Fit parameters for the 2A+2B 47 : 53% populated

```
Conformational averaged solution
Alignment tensor
A'x= 3.436e-04
A'y= 4.248e-04
A'z=-7.684e-04
Saupe tensor
S'x= 5.154e-04
S'y= 6.373e-04
S'z=-1.153e-03
Alignment tensor eigenvectors
e[x]=(-0.397, 0.074,-0.915)
e[y]=(-0.049, 0.994, 0.102)
e[z]=( 0.917, 0.085,-0.391)

Alignment tensor in laboratory coordinates:
[-5.905e-04,-9.111e-05,3.977e-04]
[-9.111e-05,4.157e-04,4.534e-05]
[ 3.977e-04,4.534e-05,1.748e-04]

SVD condition number is 6.506e+00
Axial component Aa = -1.153e-03
Rhombic component Ar = -8.126e-05
rhombicity R = 7.050e-02
Asymmetry parameter etha =1.057e-01
GDO = 1.335e-03

Euler Angles (degrees)
Set 1
(167.7,-66.4,-172.9)
Set 2
(-12.3,246.4,7.1)
```

B.3 19-OH-(−)-Eburnamonine

B.3.1 DFT-minimized XYZ geometries

SSS anti

C	-4.816635	-0.692443	-0.542881
C	-4.740298	0.661509	-0.174207
C	-3.531538	1.244343	0.215902
C	-2.399554	0.427301	0.224375
C	-2.452553	-0.948710	-0.146208
C	-3.682149	-1.503779	-0.533300
C	-1.114372	-1.479575	-0.021011
C	-0.330269	-0.449643	0.397389
C	1.119118	-0.492573	0.681079
N	1.706980	-1.712710	0.146517
C	0.874128	-2.881458	0.437069
C	-0.495230	-2.819142	-0.258358
N	-1.066484	0.721152	0.555909
C	-0.466980	1.956503	0.854219
C	1.050913	1.907013	1.030678
C	1.801678	0.816302	0.223519
C	3.277916	0.714846	0.658300
C	3.921561	-0.637475	0.317378
C	3.053833	-1.839561	0.693649
O	-1.116970	2.980323	0.987822
C	1.610076	1.174968	-1.284894
C	2.153152	0.233598	-2.358836
O	2.183530	2.475015	-1.437311
H	-5.777760	-1.114405	-0.840932
H	-5.642147	1.275071	-0.189625
H	-3.469152	2.290460	0.502288
H	-3.749020	-2.553272	-0.823191
H	1.239794	-0.488803	1.795362
H	0.714503	-3.003634	1.531087
H	1.420628	-3.771508	0.092463
H	-1.126191	-3.631302	0.135555
H	-0.379485	-3.007123	-1.338793
H	1.234894	1.756851	2.108005
H	1.418892	2.907195	0.780600
H	3.858195	1.537933	0.220488
H	3.318017	0.851270	1.751751
H	4.882678	-0.721655	0.847832
H	4.154883	-0.694612	-0.751906
H	3.025849	-1.948191	1.802599
H	3.506828	-2.759926	0.295225
H	0.517789	1.239181	-1.444510
H	3.247034	0.283850	-2.413510
H	1.850158	-0.803653	-2.182850
H	1.761681	0.546930	-3.339939
H	1.979204	2.755392	-2.342242

SSS (+)-sync

C	-4.797653	-0.661640	-0.461865
C	-4.706056	0.670276	-0.021810
C	-3.486339	1.228853	0.367514
C	-2.356713	0.410072	0.304253
C	-2.425141	-0.941478	-0.144878
C	-3.666112	-1.473501	-0.528996
C	-1.085583	-1.480006	-0.087256
C	-0.287069	-0.476936	0.377820
C	1.165254	-0.550785	0.648216
N	1.732635	-1.735064	0.014722
C	0.900595	-2.919773	0.237484
C	-0.481099	-2.805958	-0.421983
N	-1.017346	0.681743	0.623456
C	-0.396950	1.925419	0.845887
C	1.125253	1.880995	0.973412
C	1.884047	0.760125	0.236692
C	3.320965	0.592327	0.772642
C	3.966733	-0.697252	0.253595
C	3.092800	-1.926466	0.509654
O	-1.036411	2.951687	0.994688
C	1.867074	0.978372	-1.314707
C	2.756446	2.111756	-1.818075
O	0.566369	1.272081	-1.806739
H	-5.768035	-1.064685	-0.755787
H	-5.605725	1.285833	0.019319
H	-3.412079	2.258922	0.704114
H	-3.743412	-2.505168	-0.874805
H	1.289018	-0.618258	1.759368
H	0.764058	-3.118686	1.323056
H	1.436182	-3.784528	-0.180583
H	-1.110090	-3.637211	-0.067224
H	-0.389839	-2.926361	-1.514414
H	1.330471	1.797542	2.054716
H	1.474025	2.877988	0.680753
H	3.948100	1.463414	0.538457
H	3.269783	0.542115	1.873651
H	4.938643	-0.849993	0.747223
H	4.174658	-0.617988	-0.822800
H	3.093455	-2.153342	1.600563
H	3.525247	-2.802091	0.002350
H	2.220843	0.037587	-1.767820
H	2.493498	3.076498	-1.365354
H	3.820847	1.919157	-1.641599
H	2.605972	2.206682	-2.900995
H	0.054987	0.449278	-1.748612

SSS (-)-sync

C	-4.721946	-0.977242	-0.352348
C	-4.734216	0.392263	-0.037342
C	-3.560224	1.077116	0.287320
C	-2.371409	0.345452	0.286690
C	-2.334576	-1.043606	-0.031100
C	-3.530907	-1.702618	-0.353522
C	-0.957347	-1.470318	0.065841
C	-0.238128	-0.370446	0.418058
C	1.212591	-0.296089	0.676985
N	1.884904	-1.492610	0.162194
C	1.139574	-2.715130	0.491719
C	-0.249702	-2.767165	-0.159856
N	-1.052292	0.749603	0.555402
C	-0.538158	2.043049	0.720234
C	0.988450	2.136847	0.784867
C	1.821940	1.014244	0.131301
C	3.270956	1.027920	0.676269
C	4.037514	-0.264424	0.367244
C	3.243142	-1.524242	0.720474
O	-1.255827	3.023247	0.835051
C	1.837945	1.153382	-1.436874
C	0.528539	1.504862	-2.155207
O	2.423842	0.031627	-2.069689
H	-5.658318	-1.479281	-0.600482
H	-5.679104	0.937244	-0.043518
H	-3.566042	2.135567	0.531515
H	-3.528888	-2.764373	-0.603321
H	1.349457	-0.249152	1.786207
H	1.025240	-2.815902	1.591277
H	1.741360	-3.569220	0.149857
H	-0.804716	-3.614060	0.272730
H	-0.155898	-2.972613	-1.239333
H	1.233542	2.186836	1.859818
H	1.246739	3.123930	0.378972
H	3.823524	1.895289	0.284476
H	3.221165	1.159483	1.770813
H	4.976475	-0.279650	0.942136
H	4.301390	-0.289487	-0.695671
H	3.195663	-1.643321	1.824980
H	3.756580	-2.411892	0.322425
H	2.524242	2.000055	-1.624028
H	-0.217315	0.705426	-2.082642
H	0.078769	2.447624	-1.816535
H	0.773298	1.622959	-3.219332
H	2.126199	-0.746531	-1.540928

SSR anti

C	-4.821792	-0.648968	-0.428712
C	-4.719392	0.686696	-0.004194
C	-3.491210	1.243474	0.362809
C	-2.366875	0.418968	0.292451
C	-2.446260	-0.938269	-0.137071
C	-3.694509	-1.467227	-0.500256
C	-1.109634	-1.484534	-0.084058
C	-0.301211	-0.481341	0.353949

```
C      1.149736   -0.558155    0.616635
N      1.717887   -1.749876   -0.002284
C      0.877263   -2.926955    0.229744
C     -0.508990   -2.811341   -0.420647
N     -1.019079    0.688347    0.586097
C     -0.401222    1.917595    0.861199
C      1.124972    1.882931    0.948323
C      1.877578    0.747760    0.225306
C      3.306709    0.576032    0.790443
C      3.953723   -0.720584    0.294199
C      3.067966   -1.946914    0.521161
O     -1.041333    2.940860    1.044433
C      1.970194    0.993173   -1.312539
C      0.686732    1.309174   -2.079040
O      2.899549    2.068201   -1.469847
H     -5.797558   -1.051149   -0.705725
H     -5.615894    1.306345    0.044221
H     -3.408721    2.275414    0.691899
H     -3.781370   -2.502528   -0.832642
H      1.269715   -0.619647    1.729328
H      0.745690   -3.119719    1.317500
H      1.405425   -3.797476   -0.186103
H     -1.134923   -3.643573   -0.062549
H     -0.424432   -2.928872   -1.513912
H      1.357714    1.842992    2.026088
H      1.475537    2.863087    0.603090
H      3.932315    1.439111    0.534480
H      3.233910    0.541053    1.890416
H      4.908450   -0.878904    0.818774
H      4.201124   -0.635645   -0.773069
H      3.040605   -2.179979    1.610676
H      3.507250   -2.822992    0.020040
H      2.385054    0.067308   -1.746339
H      0.937884    1.447936   -3.143342
H     -0.041484    0.492374   -2.034672
H      0.213109    2.236445   -1.733763
H      2.910546    2.277535   -2.415993

SSR (+)-sync
C     -4.738740   -0.928519   -0.384458
C     -4.731348    0.433824   -0.039397
C     -3.546029    1.095136    0.293748
C     -2.366084    0.349674    0.269456
C     -2.349034   -1.032695   -0.081215
C     -3.556565   -1.668373   -0.409423
C     -0.977763   -1.477800   -0.000805
C     -0.245177   -0.393917    0.371006
C      1.205781   -0.352345    0.642312
N      1.864376   -1.560696    0.163473
C      1.080740   -2.761557    0.467346
C     -0.284537   -2.780507   -0.235401
N     -1.045280    0.726509    0.553901
C     -0.509821    2.007116    0.747336
C      1.015291    2.062102    0.877859
C      1.843883    0.970003    0.153828
C      3.297620    0.940954    0.682076
C      4.025465   -0.388385    0.431873
C      3.185907   -1.618779    0.779528
O     -1.207729    2.999192    0.876921
C      1.798825    1.309351   -1.370145
C      2.457204    0.365062   -2.377256
O      0.435508    1.501354   -1.718910
H     -5.683156   -1.413572   -0.636834
H     -5.669178    0.991147   -0.027717
H     -3.535663    2.148020    0.561235
H     -3.569970   -2.724645   -0.682088
H      1.314826   -0.300007    1.757558
H      0.918030   -2.868996    1.562537
H      1.673717   -3.630514    0.145417
H     -0.867352   -3.627865    0.158393
H     -0.152238   -2.963623   -1.314960
H      1.229101    2.018684    1.959315
H      1.303670    3.069804    0.552258
H      3.882873    1.777425    0.269002
H      3.256039    1.108766    1.770568
H      4.944095   -0.410754    1.038418
H      4.348776   -0.466756   -0.612013
H      3.103718   -1.710716    1.887518
H      3.699642   -2.524410    0.423212
H      2.329852    2.283479   -1.433198
H      2.231581    0.726743   -3.393155
H      3.549642    0.355326   -2.290655
H      2.080016   -0.656727   -2.276045
H      0.422364    1.872062   -2.614096

SSR (-)-sync
C     -4.804143   -0.723581   -0.554974
C     -4.746707    0.634941   -0.199744
C     -3.550418    1.234578    0.202312
C     -2.410648    0.428989    0.236216
C     -2.444986   -0.950171   -0.119755
C     -3.662168   -1.523103   -0.519370
C     -1.103464   -1.465471    0.028452
C     -0.334289   -0.426202    0.453160
C      1.118404   -0.449238    0.732490
N      1.717406   -1.647148    0.136155
C      0.917253   -2.845876    0.418183
C     -0.473322   -2.794839   -0.231028
N     -1.085396    0.740198    0.583948
C     -0.499578    1.995259    0.810250
C      1.015723    1.962480    1.007411
C      1.787366    0.860737    0.249123
C      3.253966    0.774669    0.712158
C      3.927394   -0.534549    0.282440
C      3.093080   -1.768965    0.630328
O     -1.154386    3.022404    0.877457
C      1.607633    1.055128   -1.299735
C      2.254376    2.313769   -1.872579
O      2.096111   -0.032101   -2.055749
H     -5.756121   -1.157956   -0.863998
H     -5.654022    1.239342   -0.236845
H     -3.502933    2.285110    0.474621
H     -3.713624   -2.575911   -0.799914
H      1.263729   -0.480284    1.840525
H      0.803313   -2.990400    1.512826
H      1.473387   -3.712650    0.033858
H     -1.079489   -3.619272    0.176076
H     -0.391210   -2.972928   -1.316244
H      1.182564    1.834088    2.091044
H      1.377491    2.968211    0.766319
H      3.834144    1.634709    0.350690
H      3.266671    0.838269    1.814029
H      4.905473   -0.627563    0.779174
H      4.107151   -0.516371   -0.798605
H      3.093260   -1.928434    1.730804
H      3.542558   -2.663553    0.174693
H      0.515763    1.157659   -1.468236
H      2.058281    2.331846   -2.952570
H      1.839918    3.234256   -1.443438
H      3.343487    2.313350   -1.737897
H      1.859503   -0.836637   -1.533322

SRS anti
C     -4.959575   -0.126154   -0.111731
C     -4.658957    1.240319   -0.248557
C     -3.346230    1.710144   -0.160720
C     -2.342586    0.766599    0.068482
C     -2.622813   -0.621701    0.211861
C     -3.952380   -1.062856    0.117835
C     -1.364779   -1.294156    0.438457
C     -0.395657   -0.330334    0.429595
C      1.069612   -0.592100    0.585999
N      1.373545   -1.948518    0.109251
C      0.512588   -2.904213    0.805807
C     -1.013959   -2.741733    0.572095
N     -0.955468    0.927546    0.207915
C     -0.179501    2.083756    0.066335
C      1.301208    1.874532    0.343898
C      1.917851    0.521600   -0.068068
C      1.849246    0.352552   -1.602412
```

C	2.184351	-1.069136	-2.046865
C	1.288955	-2.082811	-1.345894
O	-0.662348	3.172802	-0.200482
C	3.401711	0.424521	0.402650
C	3.679251	0.663090	1.890115
O	4.138199	1.360253	-0.381055
H	-5.996606	-0.457645	-0.185302
H	-5.465069	1.953266	-0.427392
H	-3.108955	2.764592	-0.269519
H	-4.193850	-2.121199	0.224974
H	1.294848	-0.596599	1.666454
H	0.722269	-2.799473	1.882209
H	0.826711	-3.919407	0.519950
H	-1.553051	-3.211943	1.410772
H	-1.335813	-3.287785	-0.330687
H	1.414798	2.021085	1.430523
H	1.832565	2.702750	-0.138138
H	0.828597	0.592275	-1.939725
H	2.520173	1.078097	-2.077679
H	3.241203	-1.301676	-1.852567
H	2.039975	-1.159255	-3.134155
H	1.604802	-3.108071	-1.594520
H	0.255120	-1.961102	-1.726390
H	3.726412	-0.605826	0.172859
H	3.464419	1.697531	2.187618
H	3.125048	-0.018778	2.547639
H	4.746991	0.482173	2.089201
H	5.066053	1.267831	-0.117284

SRS (+)-sync

C	-4.925318	-0.488649	-0.021213
C	-4.752421	0.893676	-0.209186
C	-3.485858	1.482531	-0.167943
C	-2.395892	0.641360	0.066779
C	-2.547089	-0.759856	0.262300
C	-3.832727	-1.322232	0.214915
C	-1.230085	-1.309692	0.478218
C	-0.348026	-0.265398	0.419555
C	1.141243	-0.403571	0.526349
N	1.519785	-1.730536	0.015986
C	0.787204	-2.757317	0.755337
C	-0.757446	-2.721232	0.616783
N	-1.025912	0.927132	0.165785
C	-0.361758	2.137467	-0.045110
C	1.132372	2.078269	0.209823
C	1.899970	0.774321	-0.130996
C	1.941722	0.600198	-1.665400
C	2.221070	-0.823958	-2.152010
C	1.348991	-1.843919	-1.430929
O	-0.943619	3.166524	-0.353900
C	3.319853	0.971829	0.492573
C	4.385442	-0.092507	0.237851
O	3.118936	1.133298	1.896562
H	-5.929508	-0.914011	-0.059221
H	-5.623144	1.525199	-0.391406
H	-3.348231	2.549822	-0.315859
H	-3.974984	-2.393530	0.363135
H	1.421097	-0.389513	1.592571
H	1.056278	-2.637177	1.816338
H	1.166440	-3.741574	0.441255
H	-1.205379	-3.215419	1.494469
H	-1.089253	-3.308266	-0.256262
H	1.272343	2.297700	1.279850
H	1.565059	2.925452	-0.338435
H	0.962242	0.905702	-2.065760
H	2.670002	1.305908	-2.096169
H	3.274221	-1.096040	-2.016335
H	2.019364	-0.879986	-3.232741
H	1.650657	-2.862557	-1.719411
H	0.295326	-1.715981	-1.750918
H	3.692035	1.927148	0.066435
H	4.035267	-1.089931	0.523818
H	4.700732	-0.111032	-0.811882
H	5.285448	0.150800	0.825879
H	3.996865	1.244622	2.290783

SRS (-)-sync

C	-4.964894	-0.340813	-0.224714
C	-4.743477	1.045396	-0.299747
C	-3.465674	1.589665	-0.148871
C	-2.414271	0.699318	0.078882
C	-2.614791	-0.708078	0.161448
C	-3.910903	-1.224187	0.005119
C	-1.326059	-1.313894	0.404489
C	-0.419225	-0.294045	0.459216
C	1.048947	-0.460812	0.652737
N	1.456440	-1.790289	0.155792
C	0.631910	-2.818427	0.803021
C	-0.887400	-2.740700	0.506267
N	-1.043134	0.935710	0.266203
C	-0.329304	2.135802	0.190566
C	1.163889	2.005464	0.470026
C	1.875595	0.700095	0.051806
C	1.894832	0.571136	-1.484037
C	2.291948	-0.823766	-1.965453
C	1.420596	-1.889394	-1.311290
O	-0.866226	3.208106	-0.037247
C	3.292309	0.660087	0.727251
C	4.281593	1.725485	0.266022
O	3.936742	-0.587087	0.584301
H	-5.976846	-0.729988	-0.347132
H	-5.585298	1.715421	-0.479627
H	-3.290567	2.660068	-0.209191
H	-4.090642	-2.298401	0.064045
H	1.259737	-0.470207	1.739769
H	0.794891	-2.720416	1.887181
H	1.022689	-3.803357	0.509009
H	-1.427590	-3.272025	1.306508
H	-1.135166	-3.276735	-0.425254
H	1.283481	2.155891	1.557756
H	1.624273	2.876154	-0.011030
H	0.887239	0.800450	-1.867571
H	2.560667	1.332016	-1.914512
H	3.349011	-1.016026	-1.746214
H	2.162923	-0.890129	-3.056524
H	1.785735	-2.894137	-1.572414
H	0.389438	-1.802637	-1.701879
H	3.091568	0.851283	1.807186
H	4.519522	1.625106	-0.800301
H	3.922045	2.744098	0.458305
H	5.217090	1.588753	0.824170
H	3.224785	-1.271269	0.590886

SRR anti

C	-4.974527	-0.283820	-0.164218
C	-4.726984	1.095776	-0.275330
C	-3.436903	1.617676	-0.153294
C	-2.400427	0.711786	0.083364
C	-2.627161	-0.687599	0.202111
C	-3.935050	-1.182032	0.073878
C	-1.348322	-1.312029	0.443977
C	-0.415787	-0.311040	0.473356
C	1.060719	-0.525541	0.634163
N	1.396229	-1.849281	0.089632
C	0.579179	-2.868175	0.746846
C	-0.952251	-2.750268	0.534275
N	-1.024663	0.924265	0.253305
C	-0.293458	2.109263	0.136850
C	1.183146	1.958467	0.452960
C	1.897592	0.640745	0.053925
C	1.968033	0.535954	-1.484422
C	2.230548	-0.876713	-2.009931
C	1.299448	-1.898327	-1.369286
O	-0.812104	3.180375	-0.137175
C	3.294860	0.714471	0.752656
C	4.247128	-0.474790	0.636938
O	3.921140	1.900962	0.273831
H	-5.995515	-0.655625	-0.264423
H	-5.557884	1.777874	-0.460967
H	-3.240870	2.682412	-0.242588

```
H     -4.134799   -2.250753    0.161987
H      1.286149   -0.575283    1.716466
H      0.798116   -2.806603    1.824642
H      0.928334   -3.855587    0.409195
H     -1.468176   -3.263150    1.362399
H     -1.265364   -3.276869   -0.383075
H      1.269166    2.098562    1.545038
H      1.689883    2.816047   -0.000427
H      1.007469    0.880745   -1.898716
H      2.729918    1.236980   -1.851399
H      3.270995   -1.176946   -1.838455
H      2.076565   -0.891138   -3.099818
H      1.585371   -2.914826   -1.680678
H      0.267127   -1.726601   -1.734084
H      3.062024    0.834852    1.832722
H      5.072319   -0.348037    1.356362
H      3.745754   -1.425130    0.851495
H      4.688452   -0.530204   -0.365151
H      4.813020    1.910868    0.652762

SRR (+)-sync
C     -4.961682   -0.131748   -0.125483
C     -4.664124    1.235295   -0.262170
C     -3.352154    1.707677   -0.174224
C     -2.346583    0.766377    0.054634
C     -2.623657   -0.622732    0.198303
C     -3.952386   -1.066349    0.104193
C     -1.364144   -1.292499    0.424583
C     -0.396789   -0.326847    0.413591
C      1.065918   -0.577741    0.577426
N      1.378404   -1.940187    0.121611
C      0.515487   -2.888813    0.824887
C     -1.008485   -2.737860    0.569431
N     -0.959033    0.929346    0.191270
C     -0.185600    2.084509    0.041542
C      1.302176    1.887246    0.288793
C      1.925459    0.525607   -0.087965
C      1.897731    0.322760   -1.615103
C      2.241865   -1.107510   -2.026241
C      1.314572   -2.096434   -1.330212
O     -0.672891    3.173541   -0.220586
C      3.367070    0.419453    0.494100
C      4.421319    1.326471   -0.142730
O      3.264623    0.694082    1.892552
H     -5.998012   -0.465565   -0.199002
H     -5.471759    1.946707   -0.440360
H     -3.116987    2.762726   -0.281741
H     -4.191682   -2.125207    0.211328
H      1.313569   -0.549176    1.652017
H      0.711446   -2.761136    1.900998
H      0.838023   -3.907633    0.561395
H     -1.556182   -3.206387    1.403501
H     -1.314196   -3.291381   -0.334607
H      1.466867    2.069759    1.362559
H      1.796752    2.708794   -0.243381
H      0.880499    0.547353   -1.973402
H      2.554974    1.048843   -2.113411
H      3.288347   -1.349380   -1.789087
H      2.130979   -1.217563   -3.115654
H      1.618354   -3.129371   -1.561163
H      0.290780   -1.961351   -1.734192
H      3.675482   -0.631370    0.356701
H      5.382718    1.186415    0.374244
H      4.595080    1.092183   -1.200560
H      4.154984    2.387429   -0.055187
H      4.104551    0.419031    2.288984

SRR (-)-sync
C     -4.911544   -0.513593   -0.019576
C     -4.749345    0.870732   -0.202088
C     -3.488157    1.470121   -0.155822
C     -2.391912    0.637215    0.078268
C     -2.532269   -0.767047    0.268232
C     -3.813081   -1.339474    0.215744
C     -1.210577   -1.306882    0.487392
C     -0.343629   -0.252708    0.432924
C      1.140150   -0.363297    0.542121
N      1.562549   -1.694677    0.061248
C      0.830678   -2.724114    0.809676
C     -0.710509   -2.709503    0.634882
N     -1.024589    0.936916    0.183366
C     -0.362538    2.153183   -0.020759
C      1.136776    2.094852    0.245953
C      1.878600    0.800602   -0.151815
C      1.797948    0.613717   -1.683995
C      2.191476   -0.785646   -2.155228
C      1.414455   -1.853570   -1.393916
O     -0.947276    3.179736   -0.326488
C      3.373085    0.856984    0.326223
C      3.613321    1.378056    1.750810
O      4.025056   -0.390814    0.173981
H     -5.912101   -0.946719   -0.061504
H     -5.624713    1.495641   -0.384263
H     -3.359512    2.539098   -0.299861
H     -3.947302   -2.412388    0.359212
H      1.413752   -0.335074    1.612285
H      1.077115   -2.579369    1.872671
H      1.231420   -3.706344    0.520149
H     -1.165035   -3.211670    1.504294
H     -1.012454   -3.305139   -0.242751
H      1.244512    2.270527    1.327958
H      1.571940    2.967766   -0.257134
H      0.765248    0.814317   -2.011973
H      2.426536    1.373625   -2.172635
H      3.269103   -0.941467   -2.027120
H      1.966857   -0.890140   -3.227692
H      1.796512   -2.854242   -1.646135
H      0.352752   -1.820540   -1.701707
H      3.885737    1.558456   -0.357083
H      4.681906    1.251347    1.967073
H      3.375694    2.444096    1.868933
H      3.060948    0.809450    2.512877
H      3.335573   -1.086043    0.298007
```

B.3.2 DFT energies of conformers

B.3.3 Mspin formatted RDC input file

```
rdc_data {
#H9
6 27 20.79
#H6
12 31 -12.90
12 32 -12.90
#H5
11 29 -14.00
11 30 -14.00
#H3
19 39 6.03
19 40 6.03
#H2
18 37 -7.86
18 38 -7.86
#H21
9 28 -21.57
#H18
22 42 -5.41
22 43 -5.41
22 44 -5.41
#H17
15 33 -8.20
15 34 -8.20
#H15
17 35 16.38
17 36 16.38
#H14
21 41 -38.58
#H12
3 26 20.43
#H11
2 25 10.63
#H10
1 24 19.79
```

TABLE B.4: 19-OH-(−)-eburnamonine structres $\Delta G_{298.15K}$ free energies and Boltzmann-averaged populations.

Structure	$G_{298.15K}$, a.u.	SCF Energy, a.u.	$\Delta G_{298.15K}$, kcal/mol	Boltzmann populations, %
SSS				
anti	−996.80908	−996.63058	3.07	0
(+)-*sync*	−996.81398	−996.63568	0.00	85
(−)-*sync*	−996.81385	−996.63591	0.08	14
SSR				
anti	−996.81050	−996.63206	6.07	0
(+)-*sync*	−996.80332	−996.62611	10.57	0
(−)-*sync*	−996.82017	−996.64249	0.00	100
SRS				
anti	−996.81192	−996.63407	4.53	0
(+)-*sync*	−996.80956	−996.63119	6.01	0
(−)-*sync*	−996.81914	−996.64098	0.00	100
SRR				
anti	−996.80868	−996.63031	4.71	0
(+)-*sync*	−996.81443	−996.63634	1.10	51
(−)-*sync*	−996.81639	−996.63810	0.00	49

```
}
```

B.3.4 RDC fit output parameters

```
Alignment tensor
A'x=-2.312e-04
A'y=-3.551e-04
A'z= 5.864e-04
Saupe tensor
S'x=-3.469e-04
S'y=-5.327e-04
S'z= 8.795e-04
Alignment tensor eigenvectors
e[x]=( 0.804,-0.428, 0.412)
e[y]=( 0.512, 0.851,-0.116)
e[z]=(-0.301, 0.304, 0.904)

Alignment tensor in laboratory coordinates:
[-1.897e-04,-1.290e-04,-2.151e-04]
[-1.290e-04,-2.451e-04,2.372e-04]
[-2.151e-04,2.372e-04,4.348e-04]

SVD condition number is 4.475e+00
Axial component Aa = 8.795e-04
Rhombic component Ar = 1.239e-04
rhombicity R = 1.408e-01
Asimmetry parameter etha =2.112e-01
GDO = 1.027e-03

Euler Angles (degrees)
Set 1
(18.6,17.5,32.5)
Set 2
(-161.4,162.5,-147.5)
```

B.4 10-Epi

B.4.1 DFT-minimized XYZ geometries

RRRSR

C	1.353634	-1.746509	0.099862
C	1.728985	-0.246829	0.045081
C	1.982838	0.522994	1.357862
H	2.901364	0.186419	1.860675
H	1.161393	0.430971	2.086617
C	2.069342	-2.438220	1.270482
O	1.893314	-2.268207	-1.129325
H	1.690400	-2.105102	2.245275
H	3.149963	-2.249833	1.229561
H	1.917246	-3.527083	1.214193
H	1.708248	-3.221743	-1.119688
C	-0.603193	0.872422	-0.692374
C	0.926270	0.758970	-0.827664
H	1.109230	0.564347	-1.897733
C	-1.518615	-0.276519	-1.177925
O	-0.981997	1.133703	0.678657
H	-0.897825	1.772748	-1.257970
C	-2.703526	-0.087561	-0.267982
H	-1.804308	-0.087511	-2.223569
C	-2.249024	0.689451	0.921615
O	-2.829123	0.917377	1.961590
C	-3.973783	-0.492599	-0.417008
C	1.558307	2.076871	-0.378760
H	-4.311130	-1.046116	-1.294650
H	-4.710213	-0.267475	0.355354
C	2.092954	1.930088	0.847685
C	1.508258	3.327041	-1.197436
H	2.554304	2.734554	1.423314
H	1.986867	3.180533	-2.179357
H	2.023542	4.153134	-0.690708
H	0.476707	3.657307	-1.397348
C	-0.156942	-2.060541	0.163934
C	-0.926923	-1.693644	-1.107823
H	-0.597245	-1.603300	1.060311
H	-0.251549	-3.148701	0.320119
H	-1.760822	-2.397089	-1.250045
H	-0.250905	-1.836513	-1.962912
H	2.715696	-0.265014	-0.446101

RRRSS

C	1.323799	-1.682278	0.068726
C	1.868150	-0.273343	-0.273445
C	2.659175	0.418994	0.859488
H	3.716174	0.112083	0.864895
H	2.258382	0.165922	1.854595
C	2.493718	-2.668855	0.177904
O	0.699576	-1.723880	1.348048
H	2.118030	-3.640635	0.524599
H	3.238633	-2.320893	0.903558
H	2.986671	-2.813763	-0.792973
H	0.154039	-0.923115	1.427297
C	-0.611819	0.786037	-0.535324
C	0.891306	0.813601	-0.826665
H	0.910059	0.775127	-1.931759
C	-1.514686	-0.319820	-1.112018
O	-0.931457	0.812735	0.880742
H	-1.006212	1.730297	-0.949064
C	-2.753376	-0.012808	-0.309176
H	-1.651528	-0.150010	-2.189560
C	-2.273938	0.512806	1.009610
O	-2.871093	0.662240	2.049798
C	-4.053067	-0.143351	-0.612170
C	1.537298	2.114310	-0.364840
H	-4.387762	-0.517272	-1.580607
H	-4.816169	0.128170	0.118218
C	2.491990	1.872180	0.550535
C	1.127647	3.455236	-0.888442
H	3.083318	2.649556	1.038553
H	1.792084	4.243407	-0.512207
H	0.102454	3.726248	-0.590927
H	1.157625	3.485876	-1.989704
C	0.354329	-2.184206	-1.043004
C	-1.106728	-1.786968	-0.853644
H	0.378062	-3.284908	-1.036343
H	0.731997	-1.880224	-2.032493
H	-1.390425	-2.074514	0.168731
H	-1.737512	-2.404440	-1.511140
H	2.583069	-0.424895	-1.097575

RRSSR

C	-2.015593	-1.227867	0.242478
C	-1.863744	0.318433	0.287571
C	-2.491607	1.127414	-0.870990
H	-3.588474	1.172626	-0.795312
H	-2.270387	0.699377	-1.864629
C	-3.502781	-1.579608	0.055105
O	-1.566569	-1.797268	1.478789
H	-3.638581	-2.658702	0.206419
H	-3.871500	-1.324785	-0.947053
H	-4.133690	-1.050319	0.786110
H	-2.139553	-1.427943	2.169734
C	0.677178	0.349475	-0.299499
C	-0.469214	0.981671	0.505468
H	-0.175915	0.920140	1.566000
C	1.042934	-1.081596	0.138162
O	1.876038	1.134290	-0.116789
H	0.447722	0.385040	-1.376796
C	2.537483	-1.085399	0.006894
H	0.799915	-1.156214	1.210401
C	2.983593	0.338545	-0.022102
O	4.103115	0.798653	0.029219
C	3.409116	-2.102329	-0.072093
C	-0.752540	2.422822	0.090428
H	3.095788	-3.146446	-0.056545
H	4.476435	-1.893427	-0.152599
C	-1.846333	2.469505	-0.694243
C	0.085216	3.585877	0.517357
H	-2.238333	3.378234	-1.155055
H	0.154931	3.635378	1.615707
H	-0.343727	4.532119	0.162723
H	1.114861	3.507739	0.144699
C	-1.198488	-1.916712	-0.868950
C	0.284239	-2.195483	-0.585002
H	-1.300883	-1.314886	-1.782542
H	-1.673064	-2.882179	-1.098930
H	0.779880	-2.410507	-1.544276
H	0.365751	-3.108731	0.019426
H	-2.438609	0.629595	1.179248

RSRSR

C	1.984919	-1.345966	0.215856
C	1.682833	-0.010029	-0.500285
C	2.895277	0.940282	-0.649807
H	3.404628	0.793423	-1.613859
H	3.666297	0.761214	0.116866
C	2.332815	-1.192969	1.704000
O	3.131153	-1.856642	-0.479075
H	1.481484	-0.866965	2.314427
H	3.153191	-0.481209	1.854270
H	2.658440	-2.164448	2.106843
H	3.357176	-2.695119	-0.045615
C	-0.825306	0.535307	-0.500536
C	0.542038	0.844823	0.116825
H	0.441115	0.659596	1.199464
C	-1.247650	-0.953006	-0.621928
O	-1.846649	1.170697	0.303021
H	-0.880078	0.994401	-1.502152
C	-2.733494	-0.841764	-0.410470
H	-1.027417	-1.313078	-1.637446
C	-3.002921	0.444576	0.294407
O	-4.026410	0.845720	0.803523
C	-3.715644	-1.685891	-0.760222
C	1.015861	2.279878	-0.076793
H	-3.523032	-2.619058	-1.291615
H	-4.750827	-1.445251	-0.515382
C	2.293843	2.299373	-0.496995
C	0.180086	3.477423	0.249860
H	2.866125	3.214499	-0.663846
H	0.756176	4.402684	0.120501

```
H     -0.183512    3.439687    1.287870
H     -0.717355    3.543141   -0.381735
C      0.829345   -2.353603    0.021778
C     -0.597674   -1.928787    0.380107
H      1.073427   -3.258208    0.605166
H      0.853246   -2.660579   -1.036004
H     -0.656910   -1.522542    1.400674
H     -1.222338   -2.834680    0.391131
H      1.379456   -0.283601   -1.524086

10-epi RRSSS
C      1.993742   -1.261098    0.042227
C      1.903452    0.222595   -0.435123
C      2.755929    1.201686    0.429695
H      3.783527    1.284116    0.052387
H      2.844027    0.878883    1.479604
C      3.446749   -1.744723   -0.128951
O      1.614653   -1.403975    1.428539
H      3.735549   -1.764158   -1.186673
H      3.549348   -2.758126    0.273739
H      4.155940   -1.094454    0.397022
H      2.315600   -1.035055    1.987519
C     -0.644184    0.258135    0.161358
C      0.501691    0.932673   -0.606164
H      0.209668    0.912851   -1.667777
C     -1.193724   -1.015598   -0.520035
O     -1.803847    1.146124    0.235649
H     -0.345049    0.047418    1.192478
C     -2.634791   -0.997246   -0.074580
H     -1.186579   -0.816846   -1.606359
C     -2.963806    0.428311    0.238394
O     -4.041026    0.934158    0.452502
C     -3.521657   -1.985898    0.061967
C      0.792473    2.381353   -0.211871
H     -3.273868   -3.021923   -0.153302
H     -4.531382   -1.768156    0.400032
       2.008805    2.500923    0.336016
C     -0.154848    3.513523   -0.492800
H      2.431328    3.441943    0.681942
H     -1.070831    3.436574    0.100182
H     -0.464932    3.510857   -1.547121
H      0.322607    4.476767   -0.281186
C      1.064088   -2.226062   -0.733550
C     -0.397898   -2.292253   -0.263508
H      1.483953   -3.235263   -0.639283
H      1.108433   -1.970972   -1.802041
H     -0.426381   -2.540485    0.802494
H     -0.884588   -3.119176   -0.797524
H      2.349278    0.226834   -1.438107

RSRSS
C      1.991451   -1.322299    0.216253
C      1.708503    0.023910   -0.491170
C      2.909219    1.000076   -0.506180
H      3.548393    0.864591   -1.392575
H      3.550545    0.844950    0.377672
C      3.285127   -1.960668   -0.321120
O      2.148395   -1.008305    1.602458
H      3.244156   -2.112433   -1.408687
H      3.447427   -2.947404    0.140186
H      4.157143   -1.337648   -0.091728
H      2.406999   -1.834171    2.041194
C     -0.827289    0.513604   -0.515179
C      0.531182    0.848229    0.105621
H      0.455324    0.649926    1.187407
C     -1.237039   -0.978468   -0.613589
O     -1.858646    1.151816    0.274958
H     -0.884617    0.958538   -1.523332
C     -2.723388   -0.881854   -0.402466
H     -1.014191   -1.353625   -1.623499
C     -3.006265    0.414393    0.278863
O     -4.035230    0.813660    0.778990
C     -3.696048   -1.744286   -0.733201
C      0.968303    2.295443   -0.089049
H     -3.493776   -2.685674   -1.246288
H     -4.733287   -1.511240   -0.489682
C      2.266356    2.347104   -0.437130
C      0.089097    3.476227    0.178881
H      2.827626    3.273110   -0.577202
H     -0.315267    3.446665    1.201703
H     -0.783098    3.509660   -0.489633
H      0.647948    4.413106    0.056514
C      0.849295   -2.345368    0.028502
C     -0.568182   -1.919262    0.406370
H      1.113433   -3.231003    0.632287
H      0.852245   -2.696031   -1.016755
H     -0.581346   -1.475077    1.411899
H     -1.189261   -2.825533    0.468056
H      1.465511   -0.208741   -1.541226

SRRSR
C      1.156136   -1.773992    0.152798
C      1.292292   -0.284142    0.519931
C      2.713179    0.160734    0.936497
H      3.483125   -0.415758    0.396075
H      2.898249    0.014101    2.011814
C      1.653166   -2.659085    1.308273
O      1.981091   -1.980851   -0.999494
H      1.086461   -2.477477    2.231908
H      2.715295   -2.479021    1.511720
H      1.534671   -3.724587    1.056234
H      1.977837   -2.936678   -1.166893
C     -0.625834    1.029707   -0.659167
C      0.871432    0.726977   -0.577691
H      1.153791    0.338672   -1.572980
C     -1.559156   -0.094627   -1.194709
O     -1.122206    1.415888    0.640227
H     -0.759889    1.914286   -1.300486
C     -2.708378   -0.038099   -0.224193
H     -1.916145    0.213721   -2.190117
C     -2.341785    0.867632    0.903526
O     -2.960328    1.119739    1.915137
C     -3.906927   -0.639169   -0.278338
C      1.750007    1.939025   -0.294380
H     -4.199147   -1.292881   -1.101291
H     -4.630344   -0.475173    0.521139
C      2.755867    1.599694    0.529923
C      1.540638    3.272586   -0.938593
H      3.558453    2.274472    0.833874
H      0.606414    3.747734   -0.600766
H      1.479287    3.189866   -2.036103
H      2.363729    3.958449   -0.701014
C     -0.307067   -2.165132   -0.163699
C     -0.945307   -1.494941   -1.381613
H     -0.929997   -2.015935    0.730913
H     -0.313277   -3.254088   -0.342453
H     -1.757285   -2.140875   -1.749590
H     -0.203543   -1.456295   -2.192785
H      0.634281   -0.103742    1.382463

SRSSS
C     -2.135049   -1.222069    0.180111
C     -1.619028    0.105815   -0.435366
C     -2.707059    1.155701   -0.777347
H     -3.586447    1.105438   -0.112448
H     -3.091937    1.026085   -1.802510
C     -3.001422   -1.036593    1.436237
O     -2.907014   -1.888851   -0.832245
H     -3.273154   -2.020620    1.841488
H     -3.937768   -0.506232    1.216974
H     -2.476947   -0.474642    2.221073
H     -3.637725   -1.290502   -1.054533
C      0.852348    0.363766   -0.239088
C     -0.482790    0.888921    0.292597
H     -0.500731    0.732361    1.386465
C      1.217640   -1.058125    0.213570
O      1.994733    1.149708    0.169948
H      0.824352    0.405305   -1.342294
C      2.685543   -1.051594   -0.105658
H      1.162479   -1.032818    1.320182
```

```
C      3.121322    0.368713    0.088303
O      4.236020    0.829152    0.187853
C      3.524477   -2.024438   -0.487687
C     -0.799953    2.348687   -0.000935
H      3.199579   -3.054752   -0.634068
H      4.575865   -1.795734   -0.664957
C     -2.005552    2.461231   -0.588045
C      0.076217    3.489701    0.412242
H     -2.470720    3.413996   -0.849067
H     -0.427236    4.450112    0.242354
H      1.029365    3.496719   -0.132816
H      0.336486    3.423206    1.479715
C     -1.009071   -2.228702    0.522742
C      0.309160   -2.187828   -0.259513
H     -1.471269   -3.219440    0.405354
H     -0.766421   -2.135823    1.592193
H      0.138258   -2.137057   -1.344998
H      0.825362   -3.142908   -0.078703
H     -1.201995   -0.201891   -1.406839

SSRSR
C     -2.011423   -1.221376    0.031897
C     -1.835371    0.240181   -0.433503
C     -2.701542    1.253146    0.353274
H     -2.802681    0.962800    1.411681
H     -3.721965    1.322534   -0.054505
C     -3.479142   -1.655071   -0.144026
O     -1.689433   -1.251381    1.429401
H     -3.599847   -2.715818    0.127056
H     -3.819729   -1.542983   -1.182367
H     -4.139318   -1.066883    0.504327
H     -1.908639   -2.145300    1.736581
C      0.610445    0.281548    0.521406
C     -0.402168    0.873659   -0.479777
H      0.035477    0.742494   -1.482618
C      1.046121   -1.197527    0.262953
O      1.815716    1.077384    0.474050
H      0.218116    0.388536    1.539098
C      2.537687   -1.089401    0.103096
H      0.818384   -1.776358    1.168938
C      2.932390    0.341408    0.227508
O      4.032820    0.844399    0.143843
C      3.453523   -2.049662   -0.096675
C     -0.690843    2.355910   -0.262340
H      3.193036   -3.106000   -0.175463
H      4.506769   -1.779712   -0.180481
C     -1.943183    2.531005    0.197896
C      0.282102    3.447915   -0.581868
H     -2.376984    3.504298    0.435745
H     -0.192571    4.432974   -0.483468
H      0.655915    3.352793   -1.613625
H      1.165512    3.420766    0.068659
C     -1.119916   -2.221525   -0.740011
C      0.362674   -1.881987   -0.928985
H     -1.561705   -2.401144   -1.732454
H     -1.204470   -3.184409   -0.206723
H      0.502520   -1.267032   -1.828667
H      0.900087   -2.818752   -1.139036
H     -2.188867    0.264645   -1.476102

SSRSS
C      2.268567   -0.916883    0.224896
C      1.719176    0.533556    0.267470
C      2.166353    1.559117   -0.797921
H      2.017510    1.201107   -1.832829
H      3.233106    1.812416   -0.712017
C      3.757328   -0.907601   -0.159361
O      2.141523   -1.490850    1.534319
H      4.326617   -0.206263    0.469908
H      3.918543   -0.620261   -1.206813
H      4.175887   -1.911284   -0.007934
H      2.664634   -0.929063    2.127747
C     -0.687309    0.241580   -0.716351
C      0.197654    0.821193    0.413775
H     -0.186858    0.476841    1.385437
C     -0.987930   -1.278824   -0.727564
O     -1.991336    0.862663   -0.663827
H     -0.262138    0.557303   -1.677443
C     -2.319146   -1.341884   -0.036334
H     -1.178047   -1.518965   -1.790811
C     -2.933942    0.010847   -0.158124
O     -4.058191    0.378799    0.104043
C     -2.935194   -2.353753    0.592790
C      0.208563    2.348272    0.281946
H     -2.487154   -3.342410    0.692631
H     -3.923655   -2.198111    1.026425
C      1.272818    2.719433   -0.458860
C     -0.820308    3.247495    0.887685
H      1.493195    3.745696   -0.758149
H     -0.871754    3.098110    1.978306
H     -1.825194    3.046061    0.495031
H     -0.579272    4.302785    0.703931
C      1.512747   -1.894999   -0.690246
C      0.093356   -2.225828   -0.220583
H      2.098990   -2.825383   -0.707157
H      1.515749   -1.511965   -1.721927
H      0.084999   -2.258181    0.875872
H     -0.167100   -3.239555   -0.559499
H      2.139574    0.945902    1.203693

SSSSR
C      2.351430   -0.726905    0.317524
C      1.801577    0.713588    0.151898
C      1.879016    1.278740   -1.287960
H      1.676672    0.528876   -2.069842
H      2.886194    1.663058   -1.503649
C      2.425803   -1.151658    1.792991
O      3.700985   -0.585495   -0.160842
H      1.447601   -1.286362    2.268132
H      2.987915   -0.403999    2.367722
H      2.958953   -2.111036    1.881544
H      4.110425   -1.461299   -0.072982
C     -0.643825   -0.018037    0.750166
C      0.390005    1.121216    0.691864
H      0.485606    1.480867    1.730174
C     -0.757088   -0.937635   -0.478442
O     -1.963596    0.506347    1.003180
H     -0.416626   -0.645018    1.625986
C     -2.228676   -1.252043   -0.490218
H     -0.532071   -0.347275   -1.382579
C     -2.915132   -0.239052    0.366966
O     -4.100072   -0.039623    0.521395
C     -2.912395   -2.219431   -1.119067
C      0.020170    2.295302   -0.208469
H     -2.430753   -2.968327   -1.748323
H     -3.996091   -2.271976   -1.009138
C      0.837957    2.350568   -1.276738
C     -1.048343    3.289877    0.126428
H      0.771442    3.112211   -2.056610
H     -0.896967    3.698479    1.138052
H     -2.053268    2.849510    0.120983
H     -1.035101    4.128888   -0.581441
C      1.674165   -1.800406   -0.568795
C      0.191179   -2.134733   -0.387536
H      2.241842   -2.736731   -0.427413
H      1.850738   -1.509655   -1.613912
H      0.018122   -2.647851    0.571450
H     -0.074402   -2.864630   -1.166963
H      2.506173    1.326082    0.735274

SSSSS
C      2.296923   -0.781221    0.346479
C      1.837298    0.697284    0.238117
C      2.051597    1.350169   -1.150236
H      1.973587    0.637434   -1.988602
H      3.055141    1.796809   -1.234260
C      3.821909   -0.832099    0.127554
O      2.017462   -1.157912    1.700032
H      4.198654   -1.853461    0.291975
H      4.332459   -0.167300    0.836600
```

```
H      4.108169   -0.541619   -0.891713
H      2.384547   -2.048860    1.817333
C     -0.627213   -0.028831    0.742719
C      0.394211    1.117013    0.694244
H      0.448716    1.488852    1.730613
C     -0.777073   -0.885018   -0.526750
O     -1.940872    0.481921    1.061415
H     -0.346441   -0.682627    1.579885
C     -2.251917   -1.182132   -0.528032
H     -0.554179   -0.257767   -1.407592
C     -2.910287   -0.220317    0.408011
O     -4.091309   -0.028448    0.602663
C     -2.959153   -2.099979   -1.203419
C      0.063558    2.279265   -0.230501
H     -2.498886   -2.810232   -1.890932
H     -4.040649   -2.150313   -1.072392
C      0.973573    2.382805   -1.216954
C     -1.072675    3.227528    0.001051
H      0.952276    3.152883   -1.990983
H     -1.055813    3.614132    1.031684
H     -2.052734    2.751232   -0.128625
H     -1.013154    4.083048   -0.684329
C      1.642028   -1.769472   -0.651362
C      0.155197   -2.095118   -0.502456
H      2.195056   -2.721434   -0.574416
H      1.833984   -1.414985   -1.674943
H     -0.018271   -2.648837    0.432429
H     -0.114280   -2.781735   -1.319188
H      2.501705    1.231045    0.933775

RSSSR Conformer 1
C      1.343791   -1.488338    0.240566
C      0.827617   -0.109715   -0.239910
C      1.865981    1.005053   -0.447586
H      2.499459    0.820868   -1.324899
H      2.546642    1.101147    0.417778
C      1.553142   -1.590929    1.756537
O      2.613374   -1.636165   -0.410600
H      2.010594   -2.560717    2.006487
H      0.617699   -1.521232    2.325747
H      2.231022   -0.803235    2.108379
H      2.933637   -2.521572   -0.174517
C     -1.581044   -0.221079    0.719623
C     -0.280448    0.599333    0.575185
H      0.088967    0.745206    1.607401
C     -1.841557   -1.359902   -0.286129
O     -2.763694    0.602850    0.703050
H     -1.565777   -0.686836    1.719717
C     -3.343801   -1.439686   -0.234620
H     -1.573919   -1.012228   -1.299566
C     -3.837208   -0.117577    0.258895
O     -4.965753    0.320961    0.286479
C     -4.186100   -2.435071   -0.547370
C     -0.249198    2.004254   -0.042401
H     -3.845300   -3.407426   -0.903777
H     -5.261276   -2.281806   -0.448513
C      0.978421    2.207100   -0.562924
C     -1.308388    3.051985    0.096024
H      1.322227    3.163642   -0.961117
H     -1.641549    3.154946    1.139197
H     -2.208012    2.816274   -0.487855
H     -0.926123    4.024932   -0.240148
C      0.428523   -2.613193   -0.307392
C     -1.063561   -2.636709    0.043577
H      0.856975   -3.582269    0.003209
H      0.534447   -2.572943   -1.402919
H     -1.213776   -2.875836    1.108171
H     -1.506893   -3.476405   -0.512364
H      0.406893   -0.272283   -1.245515

RSSSR Conformer 2
C      1.369056   -1.480862    0.394991
C      0.814811   -0.149834   -0.150190
C      1.815758    0.944746   -0.554019
H      2.348608    0.690695   -1.480756
H      2.587081    1.120924    0.216710
C      2.319198   -1.324054    1.590026
O      2.088416   -2.015105   -0.726194
H      2.626684   -2.311611    1.968559
H      1.849804   -0.789197    2.426510
H      3.225982   -0.781317    1.297351
H      2.535355   -2.814129   -0.403913
C     -1.574977   -0.216978    0.797570
C     -0.270965    0.603161    0.658838
H      0.088741    0.783221    1.689679
C     -1.830585   -1.265033   -0.298046
O     -2.763390    0.597513    0.842566
H     -1.548959   -0.733147    1.769556
C     -3.327099   -1.380567   -0.239903
H     -1.608645   -0.776761   -1.263197
C     -3.833330   -0.105332    0.356362
O     -4.966891    0.314225    0.429195
C     -4.151324   -2.364434   -0.628883
C     -0.260224    1.974262   -0.025874
H     -3.787417   -3.293222   -1.069080
H     -5.228974   -2.247148   -0.510469
C      0.917806    2.138657   -0.661370
C     -1.294813    3.041758    0.151062
H      1.227257    3.071074   -1.137594
H     -1.547555    3.185161    1.212058
H     -2.238157    2.798376   -0.354990
H     -0.928981    3.997639   -0.246536
C      0.226017   -2.453428    0.783860
C     -0.980680   -2.548117   -0.161900
H     -0.115755   -2.224149    1.804028
H      0.671816   -3.459630    0.859382
H     -1.623332   -3.359533    0.210869
H     -0.629133   -2.862193   -1.154614
H      0.338047   -0.428842   -1.101667

RSSSR Conformer 1
C      1.343791   -1.488338    0.240566
C      0.827617   -0.109715   -0.239910
C      1.865981    1.005053   -0.447586
H      2.499459    0.820868   -1.324899
H      2.546642    1.101147    0.417778
C      1.553142   -1.590929    1.756537
O      2.613374   -1.636165   -0.410600
H      2.010594   -2.560717    2.006487
H      0.617699   -1.521232    2.325747
H      2.231022   -0.803235    2.108379
H      2.933637   -2.521572   -0.174517
C     -1.581044   -0.221079    0.719623
C     -0.280448    0.599333    0.575185
H      0.088967    0.745206    1.607401
C     -1.841557   -1.359902   -0.286129
O     -2.763694    0.602850    0.703050
H     -1.565777   -0.686836    1.719717
C     -3.343801   -1.439686   -0.234620
H     -1.573919   -1.012228   -1.299566
C     -3.837208   -0.117577    0.258895
O     -4.965753    0.320961    0.286479
C     -4.186100   -2.435071   -0.547370
C     -0.249198    2.004254   -0.042401
H     -3.845300   -3.407426   -0.903777
H     -5.261276   -2.281806   -0.448513
C      0.978421    2.207100   -0.562924
C     -1.308388    3.051985    0.096024
H      1.322227    3.163642   -0.961117
H     -1.641549    3.154946    1.139197
H     -2.208012    2.816274   -0.487855
H     -0.926123    4.024932   -0.240148
C      0.428523   -2.613193   -0.307392
C     -1.063561   -2.636709    0.043577
H      0.856975   -3.582269    0.003209
H      0.534447   -2.572943   -1.402919
H     -1.213776   -2.875836    1.108171
H     -1.506893   -3.476405   -0.512364
H      0.406893   -0.272283   -1.245515
```

RSSSR Conformer 2

C	1.369056	-1.480862	0.394991
C	0.814811	-0.149834	-0.150190
C	1.815758	0.944746	-0.554019
H	2.348608	0.690695	-1.480756
H	2.587081	1.120924	0.216710
C	2.319198	-1.324054	1.590026
O	2.088416	-2.015105	-0.726194
H	2.626684	-2.311611	1.968559
H	1.849804	-0.789197	2.426510
H	3.225982	-0.781317	1.297351
H	2.535355	-2.814129	-0.403913
C	-1.574977	-0.216978	0.797570
C	-0.270965	0.603161	0.658838
H	0.088741	0.783221	1.689679
C	-1.830585	-1.265033	-0.298046
O	-2.763390	0.597513	0.842566
H	-1.548959	-0.733147	1.769556
C	-3.327099	-1.380567	-0.239903
H	-1.608645	-0.776761	-1.263197
C	-3.833330	-0.105332	0.356362
O	-4.966891	0.314225	0.429195
C	-4.151324	-2.364434	-0.628883
C	-0.260224	1.974262	-0.025874
H	-3.787417	-3.293222	-1.069080
H	-5.228974	-2.247148	-0.510469
C	0.917806	2.138657	-0.661370
C	-1.294813	3.041758	0.151062
H	1.227257	3.071074	-1.137594
H	-1.547555	3.185161	1.212058
H	-2.238157	2.798376	-0.354990
H	-0.928981	3.997639	-0.246536
C	0.226017	-2.453428	0.783860
C	-0.980680	-2.548117	-0.161900
H	-0.115755	-2.224149	1.804028
H	0.671816	-3.459630	0.859382
H	-1.623332	-3.359533	0.210869
H	-0.629133	-2.862193	-1.154614
H	0.338047	-0.428842	-1.101667

RSSSS Conformer 1

C	1.988533	-1.262385	0.098486
C	1.523992	0.076837	-0.530492
C	2.556767	1.156205	-0.897575
H	3.179343	0.881603	-1.761693
H	3.241611	1.365347	-0.055042
C	3.414925	-1.629823	-0.343558
O	1.961695	-1.095720	1.518370
H	4.135772	-0.886279	0.016015
H	3.503782	-1.701439	-1.436030
H	3.703313	-2.608646	0.070600
H	2.324580	-1.913323	1.895130
C	-0.803810	0.104160	0.594336
C	0.468848	0.897447	0.251441
H	0.916172	1.163015	1.225746
C	-1.189029	-1.078444	-0.315285
O	-1.971800	0.947910	0.680898
H	-0.656398	-0.306730	1.604707
C	-2.683541	-1.103969	-0.138403
H	-0.990635	-0.810581	-1.368335
C	-3.096145	0.243573	0.362079
O	-4.208612	0.708509	0.483580
C	-3.578667	-2.077710	-0.359608
C	0.458151	2.201897	-0.552822
H	-3.297012	-3.067498	-0.720079
H	-4.636970	-1.888102	-0.176881
C	1.661102	2.333934	-1.148872
C	-0.608103	3.250339	-0.525073
H	1.986733	3.224683	-1.689101
H	-0.882815	3.519276	0.505175
H	-1.535638	2.912926	-1.007014
H	-0.264843	4.157943	-1.039230
C	1.045318	-2.409744	-0.354006
C	-0.431830	-2.358464	0.050317
H	1.456252	-3.360229	0.027970
H	1.116480	-2.481904	-1.451357
H	-0.528473	-2.517243	1.134976
H	-0.925760	-3.218384	-0.426597
H	1.055074	-0.183012	-1.493045

RSSSS Conformer 2

C	2.055844	-1.248924	-0.058449
C	1.468415	0.089165	-0.604304
C	2.478134	1.157715	-1.071772
H	2.790119	1.019991	-2.118940
H	3.400259	1.152444	-0.462778
C	2.579909	-2.115066	-1.212824
O	3.112782	-1.001331	0.875629
H	3.330756	-1.574591	-1.808675
H	1.777622	-2.415258	-1.900325
H	3.052158	-3.024174	-0.816636
H	3.876273	-0.705370	0.356417
C	-0.886704	0.200149	0.493494
C	0.530025	0.830615	0.378126
H	0.972070	0.754412	1.388172
C	-1.111736	-1.119928	-0.267315
O	-1.939990	1.057257	-0.004265
H	-1.112847	0.033142	1.560113
C	-2.610193	-1.140363	-0.337063
H	-0.789521	-0.929110	-1.305149
C	-3.028100	0.296874	-0.346201
O	-4.102021	0.788262	-0.611967
C	-3.486014	-2.154151	-0.388853
C	0.667734	2.293586	-0.027636
H	-3.173402	-3.198549	-0.381425
H	-4.554906	-1.944806	-0.441089
C	1.743134	2.439488	-0.824745
C	-0.145081	3.410365	0.548563
H	2.116849	3.403019	-1.176780
H	-1.194627	3.365057	0.233282
H	0.266408	4.383876	0.252290
H	-0.144429	3.366830	1.649691
C	1.043006	-2.026813	0.812331
C	-0.356003	-2.334545	0.266773
H	0.942518	-1.459941	1.749342
H	1.525284	-2.972471	1.100764
H	-0.933118	-2.786744	1.088188
H	-0.306121	-3.098980	-0.523018
H	0.879992	-0.148390	-1.501740

SRSSR Conformer 1

C	-1.933575	-1.327646	0.237909
C	-1.730184	0.093930	-0.367133
C	-2.922067	1.071295	-0.248240
H	-3.431576	0.990731	0.731423
H	-3.694514	0.886099	-1.009199
C	-3.377696	-1.821856	0.032090
O	-1.644827	-1.337866	1.643454
H	-3.465391	-2.856230	0.391004
H	-3.671684	-1.798270	-1.026573
H	-4.100405	-1.214245	0.593562
H	-2.256026	-0.705224	2.053997
C	0.803852	0.436246	-0.350878
C	-0.534511	0.923495	0.183306
H	-0.490837	0.817809	1.282357
C	1.148990	-1.022842	-0.002309
O	1.894797	1.213883	0.189778
H	0.834026	0.564375	-1.448110
C	2.650065	-0.968222	-0.056059
H	0.871148	-1.177438	1.055314
C	3.028890	0.447785	0.243462
O	4.112896	0.916606	0.510303
C	3.568477	-1.915434	-0.293626
C	-0.939727	2.358063	-0.142229
H	3.304615	-2.950030	-0.514353
H	4.628082	-1.659296	-0.266497
C	-2.266013	2.411196	-0.367125
C	-0.014744	3.532010	-0.089228
H	-2.827091	3.331071	-0.541965
H	-0.562776	4.466864	-0.264482

```
H      0.788269    3.459364   -0.836154
H      0.487534    3.601963    0.886864
C     -1.000761   -2.405842   -0.375236
C      0.408660   -2.039936   -0.863482
H     -1.526645   -2.873742   -1.220411
H     -0.924920   -3.183042    0.399542
H      0.374384   -1.667336   -1.899528
H      0.984076   -2.975947   -0.911975
H     -1.561134   -0.031931   -1.450437

SRSSR Conformer 2
C     -2.143203   -1.192031    0.184415
C     -1.631106    0.123013   -0.445990
C     -2.733614    1.179083   -0.697707
H     -3.550482    1.083309    0.035345
H     -3.187156    1.073497   -1.695835
C     -3.076062   -1.901949   -0.813665
O     -2.887287   -0.817786    1.348690
H     -2.536852   -2.223372   -1.716153
H     -3.904138   -1.251324   -1.119482
H     -3.510490   -2.804099   -0.355632
H     -3.271739   -1.636175    1.700707
C      0.851137    0.344408   -0.251139
C     -0.477724    0.882683    0.281737
H     -0.514343    0.713669    1.372574
C      1.228645   -1.057320    0.250637
O      2.003066    1.143763    0.105876
H      0.808255    0.344007   -1.354921
C      2.688899   -1.067180   -0.097700
H      1.194097   -0.985373    1.355665
C      3.126919    0.360585    0.030974
O      4.244223    0.822314    0.089081
C      3.520835   -2.055732   -0.453811
C     -0.786287    2.349009    0.008191
H      3.194416   -3.091664   -0.549021
H      4.568654   -1.835070   -0.660123
C     -2.016052    2.476624   -0.523011
C      0.110882    3.479365    0.406041
H     -2.486702    3.436915   -0.744147
H     -0.396147    4.443604    0.271176
H      1.042796    3.494539   -0.174678
H      0.410711    3.395598    1.461911
C     -1.013910   -2.168573    0.624487
C      0.309711   -2.203902   -0.153486
H     -1.442000   -3.184031    0.641588
H     -0.781830   -1.925737    1.672461
H      0.150247   -2.218127   -1.242502
H      0.815328   -3.151099    0.088855
H     -1.237063   -0.153890   -1.436802

SRSSR Conformer 3
C     -1.960087   -1.315777    0.054680
C     -1.796791    0.183416   -0.296222
C     -2.923365    1.146906    0.121117
H     -3.234048    0.970021    1.166554
H     -3.821602    1.046130   -0.505656
C     -3.395890   -1.773715   -0.256555
O     -1.710408   -1.456032    1.454976
H     -3.509189   -2.847936   -0.042449
H     -3.653284   -1.624201   -1.314390
H     -4.123019   -1.230365    0.357921
H     -1.943212   -2.369588    1.684768
C      0.751006    0.346821   -0.434833
C     -0.520819    0.915225    0.190641
H     -0.421838    0.772985    1.282803
C      1.145366   -1.016706    0.174563
O      1.880859    1.209093   -0.178822
H      0.650689    0.280146   -1.532123
C      2.643781   -0.959059    0.124024
H      0.856912   -0.954402    1.237842
C      3.018500    0.483823    0.041715
O      4.106051    1.006298    0.147674
C      3.565483   -1.933417    0.146921
C     -0.905796    2.379842   -0.025853
H      3.303157   -2.990059    0.203430
H      4.624386   -1.675177    0.111405
C     -2.249776    2.478006   -0.020036
C      0.048068    3.530735   -0.068198
H     -2.798749    3.420994   -0.051636
H     -0.498834    4.481742   -0.113496
H      0.721329    3.478735   -0.935113
H      0.698800    3.547360    0.818327
C     -0.994064   -2.181305   -0.801771
C      0.489396   -2.269684   -0.405979
H     -1.075653   -1.826399   -1.840525
H     -1.379777   -3.213806   -0.825313
H      1.052584   -2.592786   -1.294992
H      0.617128   -3.072230    0.335133
H     -1.767859    0.233294   -1.398392
```

B.4.2 DFT energies of diastereomers

TABLE B.5: 10-epi-8-deoxycumambrin B diastereomer's basal forms energy differences.

Diastereomer	SCF Energy, kcal/mol
RRRSR	7.6
RRRSS	6.2
RRSSR	3.2
RRSSS	3.0
RSRSR	1.4
RSRSS	0.0
RSSSR	9.8
RSSSS	8.9
SRRSR	0.6
SRRSS	2.7
SRSSR	4.2
SRSSS	5.2
SSRSR	4.3
SSRSS	6.2
SSSSR	5.3
SSSSS	3.3

B.4.3 Mspin formatted RDC input file

Scenario A, short-range RDCs

```
rdc_data {
#Short range RDCs
#C1H1
2 38 -7.49 1.0
#C3H3
26 28 12.02 0.6
#C5H5
13 14 -26.26 1.1
#C6H6
12 17 -32.74 0.8
#C7H7
15 19 -26.85 1.5
#C13H13b
22 24 4.37 0.3
```

```
#C13H13a
22 25 23.37 0.3
#H13bH13a
24 25   10.96 0.3
}
```

Scenario A, short- and long-range RDCs

```
rdc_data {
#Short range RDCs
#C1H1
2 38 -7.49 1.0
#C3H3
26 28 12.02 0.6
#C5H5
13 14 -26.26 1.1
#C6H6
12 17 -32.74 0.8
#C7H7
15 19 -26.85 1.5
#C13H13b
22 24 4.37 0.3
#C13H13a
22 25 23.37 0.3
#H13bH13a
24 25   10.96 0.3
#
#Long range RDCs
#
#H2aC1
4 2 1.9 0.5
#H2aC3
4 26  1.53 0.05
#H3C1
28 2  0.28 0.09
#H3C2
28 3  0.42 0.05
#C5H3
```

```
28 13 0.33 0.03
#H3C15
28 27 0.31 0.01
#
#C1H5
14 2  -1.55 0.12
#H5C2
14 3  0.1 0.1
#H5C3
14 26 -0.06 0.03
#H5C6
14 12 -1.22 0.03
#H5C7
14 15 0.42 0.14
#H5C15
14 27 0.9 0.5
#H6-C5
13 17 -1.0 0.2
#H6C8
17 33 0.62 0.07
#
#H13bC7
24 15 1.14 0.08
}
```

Scenario B, short-range RDCs

```
rdc_data {
#Short range RDCs
#C1H1
2 38 -7.49 1.0
#C2H2a
3 4 -3.79 0.8
#C2H2b
3 5 -3.79 0.8
#C3H3
26 28 12.02 0.6
#C5H5
```

```
13 14 -26.26 1.1
#C6H6
12 17 -32.74 0.8
#C7H7
15 19 -26.85 1.5
#C8H8b
33 36 -6.44 1.2
#C8H8a
33 37 -6.44 1.2
#C9H9b
32 34 -13.12 2.1
#C9H9a
32 35 -13.12 2.1
#C13H13b
22 24 4.37 0.3
#C13H13a
22 25 23.37 0.3
#H13bH13a
24 25   10.96 0.3
#CH3-14
6 8 -5.47 0.3
6 9 -5.47 0.3
6 10 -5.47 0.3
#CH3-15
27 29 0.1 0.4
27 30 0.1 0.4
27 31 0.1 0.4
}
```

Scenario B, short- and long-range RDCs

```
rdc_data {
#Short range RDCs
#
#C1H1
2 38 -7.49 1.0
#C2H2a
3 4 -3.79 0.8
```

```
#C2H2b
3 5 -3.79 0.8
#C3H3
26 28 12.02 0.6
#C5H5
13 14 -26.26 1.1
#C6H6
12 17 -32.74 0.8
#C7H7
15 19 -26.85 1.5
#C8H8b
33 36 -6.44 1.2
#C8H8a
33 37 -6.44 1.2
#C9H9b
32 34 -13.12 2.1
#C9H9a
32 35 -13.12 2.1
#C13H13b
22 24 4.37 0.3
#C13H13a
22 25 23.37 0.3
#H13bH13a
24 25   10.96 0.3
#CH3-14
6 8 -5.47 0.3
6 9 -5.47 0.3
6 10 -5.47 0.3
#CH3-15
27 29 0.1 0.4
27 30 0.1 0.4
27 31 0.1 0.4
#
#Long range RDCs
#
#H2aC1
4 2 1.9 0.5
#H2aC3
4 26  1.53 0.05
```

```
#H3C1
28 2  0.28 0.09
#H3C2
28 3  0.42 0.05
#C5H3
28 13 0.33 0.03
#H3C15
28 27 0.31 0.01
#C1H5
14 2  -1.55 0.12
#H5C2
14 3  0.1 0.1
#H5C3
14 26 -0.06 0.03
#H5C6
14 12 -1.22 0.03
#H5C7
14 15 0.42 0.14
#H5C15
14 27 0.9 0.5
13 17 -1.0 0.2
#H6C8
17 33 0.62 0.07
#H13bC7
24 15 1.14 0.08
}
```

B.4.4 RDC fit output parameters

Scenario A, short-range RDCs

```
Alignment tensor
A'x=-1.109e-04
A'y=-3.815e-04
A'z= 4.924e-04
Saupe tensor
S'x=-1.664e-04
S'y=-5.723e-04
```

```
S'z= 7.387e-04
Alignment tensor eigenvectors
e[x]=( 0.083,-0.899,-0.430)
e[y]=( 0.965, 0.180,-0.189)
e[z]=( 0.247,-0.399, 0.883)

Alignment tensor in laboratory coordinates:
[-3.263e-04,-1.065e-04,1.810e-04]
[-1.065e-04,-2.360e-05,-2.034e-04]
[ 1.810e-04,-2.034e-04,3.499e-04]

SVD condition number is 4.102e+00
Axial component Aa = 7.387e-04
Rhombic component Ar = 2.706e-04
rhombicity R = 3.664e-01
Asimmetry parameter etha =5.496e-01
GDO = 9.151e-04

Euler Angles (degrees)
Set 1
(-24.3,-14.3,85.1)
Set 2
(155.7,194.3,-94.9)
```

Scenario A, short- and long-range RDCs

```
Alignment tensor
A'x=-1.197e-04
A'y=-3.577e-04
A'z= 4.774e-04
Saupe tensor
S'x=-1.796e-04
S'y=-5.365e-04
S'z= 7.161e-04
Alignment tensor eigenvectors
e[x]=( 0.075,-0.906,-0.417)
e[y]=( 0.965, 0.170,-0.197)
e[z]=( 0.250,-0.388, 0.887)
```

```
Alignment tensor in laboratory coordinates:
[-3.042e-04,-9.692e-05,1.777e-04]
[-9.692e-05,-3.694e-05,-1.973e-04]
[ 1.777e-04,-1.973e-04,3.412e-04]

SVD condition number is 2.555e+00
Axial component Aa = 7.161e-04
Rhombic component Ar = 2.380e-04
rhombicity R = 3.323e-01
Asimmetry parameter etha =4.985e-01
GDO = 8.767e-04

Euler Angles (degrees)
Set 1
(-23.6,-14.5,85.6)
Set 2
(156.4,194.5,-94.4)
```

Scenario B, short-range RDCs

```
Alignment tensor
A'x=-1.152e-04
A'y=-3.710e-04
A'z= 4.861e-04
Saupe tensor
S'x=-1.727e-04
S'y=-5.565e-04
S'z= 7.292e-04
Alignment tensor eigenvectors
e[x]=( 0.083,-0.903,-0.421)
e[y]=( 0.964, 0.180,-0.195)
e[z]=( 0.252,-0.390, 0.886)

Alignment tensor in laboratory coordinates:
[-3.149e-04,-1.033e-04,1.821e-04]
[-1.033e-04,-3.197e-05,-1.988e-04]
[ 1.821e-04,-1.988e-04,3.469e-04]
```

```
SVD condition number is 3.729e+00
Axial component Aa = 7.292e-04
Rhombic component Ar = 2.558e-04
rhombicity R = 3.508e-01
Asimmetry parameter etha =5.262e-01
GDO = 8.984e-04

Euler Angles (degrees)
Set 1
(-23.8,-14.6,85.1)
Set 2
(156.2,194.6,-94.9)
```

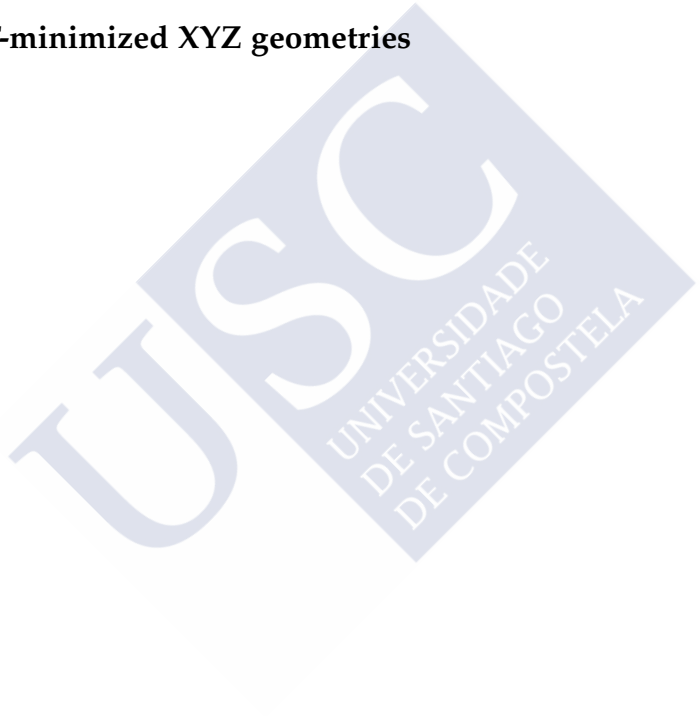
Scenario B, short- and long-range RDCs

```
Alignment tensor
A'x=-1.205e-04
A'y=-3.541e-04
A'z= 4.746e-04
Saupe tensor
S'x=-1.807e-04
S'y=-5.311e-04
S'z= 7.119e-04
Alignment tensor eigenvectors
e[x]=( 0.078,-0.906,-0.416)
e[y]=( 0.965, 0.173,-0.196)
e[z]=( 0.250,-0.386, 0.888)

Alignment tensor in laboratory coordinates:
[-3.010e-04,-9.651e-05,1.762e-04]
[-9.651e-05,-3.862e-05,-1.963e-04]
[ 1.762e-04,-1.963e-04,3.396e-04]

SVD condition number is 2.525e+00
Axial component Aa = 7.119e-04
Rhombic component Ar = 2.336e-04
rhombicity R = 3.281e-01
```

```
Asimmetry parameter etha =4.922e-01
GDO = 8.704e-04

Euler Angles (degrees)
Set 1
(-23.5,-14.5,85.4)
Set 2
(156.5,194.5,-94.6)
```

B.5 Lorcaserin

B.5.1 DFT-minimized XYZ geometries



6A

C	-2.292988	-1.351334	0.293943
C	-2.529233	-0.040954	-0.091102
C	-1.482672	0.854223	-0.269147
C	-0.160201	0.454108	-0.045754
C	0.098421	-0.866437	0.360025
C	-0.977384	-1.746556	0.512454
Cl	-4.178041	0.492959	-0.357984
C	0.926185	1.508355	-0.196245
C	2.085553	1.133978	-1.126087
N	3.066815	0.152797	-0.543712
C	2.552994	-1.254607	-0.391603
C	1.486489	-1.380044	0.690895
C	1.404061	2.043511	1.164317
H	-3.111905	-2.046579	0.419966
H	-1.696856	1.870417	-0.577317
H	-0.782443	-2.767875	0.817753
H	0.467351	2.350799	-0.720816
H	2.662045	2.030107	-1.357144
H	1.729247	0.693119	-2.056741
H	3.893502	0.129271	-1.148934
H	3.405214	0.496220	0.360001
H	2.182129	-1.556285	-1.370824
H	3.417526	-1.866855	-0.136364
H	1.413139	-2.447382	0.904259
H	1.854229	-0.928332	1.617711
H	1.804171	1.261806	1.812252
H	0.563727	2.488788	1.696235
H	2.166772	2.815315	1.036928

6B

C	-2.264250	-1.367885	0.176390
C	-2.429098	0.003411	0.281005
C	-1.366968	0.880684	0.091733
C	-0.097489	0.386531	-0.221905
C	0.087118	-1.002338	-0.335809
C	-0.997676	-1.856404	-0.131616
Cl	-4.017972	0.649269	0.653595
C	1.077622	1.327558	-0.466434
C	2.205822	1.144423	0.564624
N	3.136711	0.002080	0.248020
C	2.544551	-1.381823	0.321960
C	1.430762	-1.596158	-0.701469
C	0.690866	2.810650	-0.480094
H	-3.098872	-2.038522	0.329677
H	-1.543628	1.942511	0.185098
H	-0.853616	-2.926646	-0.222287
H	1.484057	1.095089	-1.459061
H	1.812191	0.962543	1.564652
H	2.836365	2.033151	0.589830
H	3.539329	0.143986	-0.683354
H	3.933007	0.045310	0.892168
H	3.370061	-2.069560	0.139768
H	2.189271	-1.515694	1.343120
H	1.759849	-1.239378	-1.684057
H	1.320706	-2.676200	-0.805592
H	0.360327	3.147407	0.505323
H	-0.106038	2.993561	-1.200664
H	1.547808	3.417025	-0.772732

6C

C	-2.315478	-1.369420	0.226035
C	-2.526081	-0.005596	0.077479
C	-1.465772	0.862660	-0.124331
C	-0.148339	0.384258	-0.180318
C	0.085290	-0.991737	-0.029437
C	-1.010880	-1.840795	0.172256
Cl	-4.164013	0.622496	0.135476
C	0.930667	1.438914	-0.411046
C	2.197892	0.940195	-1.104242
N	3.157751	0.223494	-0.188840
C	2.565317	-0.886904	0.651382
C	1.459779	-1.639226	-0.077449
C	1.233868	2.261662	0.853033
H	-3.146527	-2.043768	0.382736
H	-1.658625	1.921949	-0.249210
H	-0.833746	-2.903426	0.291365
H	0.511420	2.135851	-1.143069
H	2.755523	1.782843	-1.513353
H	1.965977	0.257669	-1.919525
H	3.915063	-0.158649	-0.763302
H	3.612343	0.900612	0.429925
H	3.392336	-1.547682	0.906933
H	2.189052	-0.434119	1.565737
H	1.759268	-1.857726	-1.107683
H	1.389625	-2.612649	0.409588
H	1.596952	1.646270	1.678047
H	0.326866	2.752056	1.205836
H	1.970131	3.040993	0.642874

6D

C	-2.301900	-1.348920	0.332692
C	-2.482256	0.008153	0.122084
C	-1.404732	0.854742	-0.094072
C	-0.091997	0.361338	-0.107095
C	0.107993	-1.018091	0.098806
C	-1.003451	-1.840723	0.313911
Cl	-4.109199	0.670066	0.133727
C	1.066100	1.337985	-0.356372
C	2.297837	1.008873	0.502457
N	3.227714	0.032530	-0.173481
C	2.540322	-1.158392	-0.792514
C	1.462013	-1.714461	0.126654
C	0.691557	2.806900	-0.106795
H	-3.146008	-2.003667	0.501579
H	-1.599640	1.905325	-0.245802
H	-0.844451	-2.901238	0.471536
H	1.352849	1.268583	-1.412479
H	2.012660	0.571676	1.458142
H	2.892166	1.899992	0.698106
H	3.772637	0.515040	-0.893974
H	3.919321	-0.292205	0.509264
H	2.129752	-0.837526	-1.747715
H	3.317261	-1.897628	-0.982325
H	1.308645	-2.748972	-0.183060
H	1.841534	-1.778008	1.152839
H	0.333236	2.955919	0.914093
H	-0.073813	3.145255	-0.803117
H	1.562878	3.442989	-0.262557

6E

C	-2.188641	-1.370988	0.138923
C	-2.362317	-0.001351	0.271953
C	-1.306677	0.886634	0.087727
C	-0.046330	0.400672	-0.263172
C	0.143762	-0.987205	-0.423268
C	-0.928243	-1.853478	-0.206075
Cl	-3.944258	0.627084	0.670910
C	1.138745	1.319954	-0.509621
C	2.197333	1.189240	0.598476
N	2.360782	-0.220671	1.103880
C	2.663604	-1.278262	0.073084
C	1.483303	-1.514384	-0.895915
C	0.788938	2.802755	-0.646902
H	-3.018619	-2.046650	0.295548
H	-1.481905	1.947575	0.201590
H	-0.783588	-2.920754	-0.326032
H	1.600463	1.007579	-1.450131
H	1.917579	1.784832	1.468706
H	3.175190	1.517191	0.246748
H	3.110592	-0.227240	1.800503
H	1.502086	-0.490786	1.590602
H	3.566016	-0.963237	-0.448476
H	2.892464	-2.179132	0.639850
H	1.720551	-1.067894	-1.863074
H	1.410609	-2.588020	-1.064972
H	0.365751	3.203898	0.276345
H	0.074937	2.951853	-1.457208
H	1.685195	3.379146	-0.879958

```
6F
C     -2.207766   -1.356361    0.271149
C     -2.443468   -0.046940   -0.126875
C     -1.399853    0.853530   -0.285372
C     -0.082357    0.467484   -0.008621
C      0.170245   -0.848411    0.413117
C     -0.898375   -1.742886    0.530492
Cl    -4.084668    0.474343   -0.433187
C      0.979154    1.559007   -0.111010
C      2.412266    1.119142   -0.474631
N      2.441819   -0.150450   -1.286633
C      2.503566   -1.404146   -0.440911
C      1.569662   -1.302448    0.763324
C      1.039778    2.378135    1.189161
H     -3.026565   -2.055438    0.374465
H     -1.610771    1.865822   -0.610444
H     -0.701238   -2.760590    0.846813
H      0.647962    2.228509   -0.909670
H      3.030931    0.940310    0.403262
H      2.894309    1.895487   -1.066364
H      1.610184   -0.186096   -1.879896
H      3.249272   -0.145155   -1.915258
H      2.237985   -2.233547   -1.094712
H      3.539724   -1.519271   -0.124624
H      1.537923   -2.294616    1.214964
H      2.002740   -0.643018    1.517531
H      1.306140    1.740659    2.034783
H      0.068757    2.829048    1.393365
H      1.780633    3.176038    1.110950
```

B.5.2 DFT energies of conformers

TABLE B.6: Gibbs free energies of the lorcaserin conformations.

Structure	$\Delta G_{298.15K}$ $(kcal/mol)$	$G_{298.15K}$ $(a.u.)$	SCF Energy $(a.u.)$
6A	0.00	−942.69175	−942.90306
6B	0.07	−942.69164	−942.90368
6C	2.15	−942.68833	−942.89763
6D	3.19	−942.68667	−942.90025
6E	3.34	−942.68643	−942.89730
6F	3.84	−942.68563	−942.89862

B.5.3 RdcFit formatted RDC input file

```
#RDC from F1-HSQC experiments. Using HP HCl neutralized gel
rdc_data {
8 17 8.6
1 14 -13.9
3 15 -30.1
#
#C10-H1
13 17 -1.23
#
#C4-H5beta
11 24 5.27
#
#C10-H2alpha
#13 18 0.33
#
#C5-H4beta
#12 22 1.82
#
#equivalent
9 19 2.1
9 18 2.1
#equivalent
11 22 2.9
11 23 2.9
```

```
#equivalent
12 24 -1.3
12 25 -1.3
#methyl
13 26 3.0
13 27 3.0
13 28 3.0
}
standard_error {
1.5
}
penalty_function {
q_normalized
}
gridpoints {
16
}
bootstrapping {
false
}
bootstrappingpoints {
128
}
optimize_populations {
true
}
populations {
0.5
0.5
}
adaptative_grid {
false
}
print_penalty_histogram {
true
}
print_populations_histogram {
true
}
```

B.5.4 RDC fit output parameters

Short-range RDCs, 6A+6B ensemble

```
Alignment tensor
[[  1.24303408e-04  -7.61450412e-05  -2.86554317e-04]
 [ -7.61450412e-05   4.44268599e-04   1.44296997e-04]
 [ -2.86554317e-04   1.44296997e-04  -5.68572007e-04]]
GDO (*1.0e6)= 868.081900815
Alignment tensor eigenvalues
[-0.00068261+0.j  0.00017604+0.j  0.00050657+0.j]
Alignment tensor eigenvectors
[[ 0.32475233  0.88152757  0.34270261]
 [-0.09850674  0.39189633 -0.91472055]
 [ 0.94065528 -0.26329911 -0.21410565]]

[[ 0.32475233 -0.09850674  0.94065528]
 [ 0.88152757  0.39189633 -0.26329911]
 [ 0.34270261 -0.91472055 -0.21410565]]
Ordered alignment tensor ( |Ax| < |Ay| < |Az| )
Ax = 1.760e-04
Ay = 5.066e-04
Az = -6.826e-04
Principal frame eigenvectors
X'= [ 0.882, 0.392,-0.263]
Y'= [ 0.343,-0.915,-0.214]
Z'= [ 0.325,-0.099, 0.941]
#####
Experimental and back calculated RDCs
 I    J    RDC(exp) RDC(calc) weight   std err
   8   17     8.59     8.22    1.000     1.50
   1   14   -13.95   -14.27    1.000     1.50
   3   15   -30.07   -29.76    1.000     1.50
   9   19     2.09     2.94    0.500     1.50
   9   18     2.09     2.94    0.500     1.50
  11   22     2.87     3.24    0.500     1.50
```

```
  11    23     2.87     3.24    0.500     1.50
  12    24    -1.32    -1.78    0.500     1.50
  12    25    -1.32    -1.78    0.500     1.50
  13    26     3.03     3.83    0.333     1.50
  13    27     3.03     3.83    0.333     1.50
  13    28     3.03     3.83    0.333     1.50
####
Fitting scores
RMSD =    0.412 Hz
Cornilescu Q factor =    0.041
Cornilescu Q factor with experimental errors =    0.041
Cornilescu Q factor,<r3> averaged =    0.042
Cornilescu Q factor <r3> averaged, with experimental errors =    0.042
Chi^2 =    0.907
Number of rdcs in input =  12
Number of rdcs for fit =  12
Number of non equivalent rdcs= 7
n/Chi^2 7.72011269981
Number of parameters k = 6
AIC = Chi^2 +2k =   12.907
```

Short-range + $^2D_{\mathrm{C10\text{–}H1}}$ RDCs, 6A+6B ensemble

```
Alignment tensor
[[  1.26134073e-04  -6.61684067e-05  -3.11205088e-04]
 [ -6.61684067e-05   4.56798297e-04   2.71719407e-04]
 [ -3.11205088e-04   2.71719407e-04  -5.82932370e-04]]
GDO (*1.0e6)= 956.295581652
Alignment tensor eigenvalues
[-0.00074546+0.j  0.00017156+0.j  0.00057390+0.j]
Alignment tensor eigenvectors
[[ 0.31698523  0.88510728  0.34074253]
 [-0.19244747  0.41182159 -0.89071148]
 [ 0.92870035 -0.21676734 -0.30087802]]

[[ 0.31698523 -0.19244747  0.92870035]
 [ 0.88510728  0.41182159 -0.21676734]
 [ 0.34074253 -0.89071148 -0.30087802]]
```

```
Ordered alignment tensor ( |Ax| < |Ay| < |Az| )
Ax = 1.716e-04
Ay = 5.739e-04
Az = -7.455e-04
Principal frame eigenvectors
X'= [ 0.885, 0.412,-0.217]
Y'= [ 0.341,-0.891,-0.301]
Z'= [ 0.317,-0.192, 0.929]
#####
Experimental and back calculated RDCs
 I    J    RDC(exp) RDC(calc) weight   std err
   8   17     8.59     8.17    1.000     1.50
   1   14   -13.95   -14.20    1.000     1.50
   3   15   -30.07   -29.82    1.000     1.50
  13   17    -1.23    -1.21    1.000     1.50
   9   19     2.09     2.65    0.500     1.50
   9   18     2.09     2.65    0.500     1.50
  11   22     2.87     3.21    0.500     1.50
  11   23     2.87     3.21    0.500     1.50
  12   24    -1.32    -1.45    0.500     1.50
  12   25    -1.32    -1.45    0.500     1.50
  13   26     3.03     4.36    0.333     1.50
  13   27     3.03     4.36    0.333     1.50
  13   28     3.03     4.36    0.333     1.50
####
Fitting scores
RMSD =    0.441 Hz
Cornilescu Q factor =    0.046
Cornilescu Q factor with experimental errors =    0.046
Cornilescu Q factor,<r3> averaged =    0.045
Cornilescu Q factor <r3> averaged, with experimental errors =    0.045
Chi^2 =    1.122
Number of rdcs in input =  13
Number of rdcs for fit =  13
Number of non equivalent rdcs= 8
n/Chi^2 7.12810000149
Number of parameters k = 6
AIC = Chi^2 +2k =   13.122
```

Short-range + $^2D_{\text{C10–H1}}$ + $^2D_{\text{C4–H4}\beta}$ RDCs, 6A+6B ensemble

```
Alignment tensor
[[  1.30550446e-04  -7.10159973e-05  -3.95775884e-04]
 [ -7.10159973e-05   4.93497771e-04   4.52361672e-04]
 [ -3.95775884e-04   4.52361672e-04  -6.24048217e-04]]
GDO (*1.0e6)= 1175.85945411
Alignment tensor eigenvalues
[-0.00090504+0.j  0.00017504+0.j  0.00073000+0.j]
Alignment tensor eigenvectors
[[ 0.32658019  0.87517754  0.35694489]
 [-0.27582681  0.4494605  -0.84964983]
 [ 0.90402707 -0.17902383 -0.38818232]]

[[ 0.32658019 -0.27582681  0.90402707]
 [ 0.87517754  0.4494605  -0.17902383]
 [ 0.35694489 -0.84964983 -0.38818232]]
Ordered alignment tensor ( |Ax| < |Ay| < |Az| )
Ax = 1.750e-04
Ay = 7.300e-04
Az = -9.050e-04
Principal frame eigenvectors
X'= [ 0.875, 0.449,-0.179]
Y'= [ 0.357,-0.850,-0.388]
Z'= [ 0.327,-0.276, 0.904]
#####
Experimental and back calculated RDCs
 I    J    RDC(exp) RDC(calc) weight   std err
   8   17     8.59     7.99    1.000     1.50
   1   14   -13.95   -14.14    1.000     1.50
   3   15   -30.07   -29.91    1.000     1.50
  13   17    -1.23    -1.21    1.000     1.50
  11   24     5.25     5.25    1.000     1.50
   9   19     2.09     2.50    0.500     1.50
   9   18     2.09     2.50    0.500     1.50
  11   22     2.87     3.56    0.500     1.50
  11   23     2.87     3.56    0.500     1.50
  12   24    -1.32    -1.09    0.500     1.50
  12   25    -1.32    -1.09    0.500     1.50
```

```
  13   26     3.03     4.33    0.333     1.50
  13   27     3.03     4.33    0.333     1.50
  13   28     3.03     4.33    0.333     1.50
####
Fitting scores
RMSD =    0.447 Hz
Cornilescu Q factor =    0.048
Cornilescu Q factor with experimental errors =    0.048
Cornilescu Q factor,<r3> averaged =    0.032
Cornilescu Q factor <r3> averaged, with experimental errors =    0.032
Chi^2 =    1.243
Number of rdcs in input =  14
Number of rdcs for fit =  14
Number of non equivalent rdcs= 9
n/Chi^2 7.23808462261
Number of parameters k = 6
AIC = Chi^2 +2k =   13.243
```

Short-range + $^2D_{\mathrm{C10\text{–}H1}}$ + $^2D_{\mathrm{C4\text{–}H4}\beta}$ RDCs, 6A+6B+6C ensemble

```
Alignment tensor
[[  1.26485005e-04  -7.01617094e-05  -4.06124036e-04]
 [ -7.01617094e-05   4.92208677e-04   4.58553408e-04]
 [ -4.06124036e-04   4.58553408e-04  -6.18693682e-04]]
GDO (*1.0e6)= 1183.76981004
Alignment tensor eigenvalues
[-0.00091064+0.j  0.00017478+0.j  0.00073586+0.j]
Alignment tensor eigenvectors
[[ 0.33393338  0.87111364  0.36006878]
 [-0.27773009  0.45596847 -0.84555234]
 [ 0.90075218 -0.18235622 -0.39419756]]

[[ 0.33393338 -0.27773009  0.90075218]
 [ 0.87111364  0.45596847 -0.18235622]
 [ 0.36006878 -0.84555234 -0.39419756]]
Ordered alignment tensor ( |Ax| < |Ay| < |Az| )
Ax = 1.748e-04
Ay = 7.359e-04
```

```
Az = -9.106e-04
Principal frame eigenvectors
X'= [ 0.871, 0.456,-0.182]
Y'= [ 0.360,-0.846,-0.394]
Z'= [ 0.334,-0.278, 0.901]
#####
Experimental and back calculated RDCs
 I    J    RDC(exp) RDC(calc) weight   std err
   8   17     8.59     8.30    1.000     1.50
   1   14   -13.95   -14.07    1.000     1.50
   3   15   -30.07   -30.01    1.000     1.50
  13   17    -1.23    -1.20    1.000     1.50
  11   24     5.25     5.25    1.000     1.50
   9   19     2.09     2.10    0.500     1.50
   9   18     2.09     2.10    0.500     1.50
  11   22     2.87     2.95    0.500     1.50
  11   23     2.87     2.95    0.500     1.50
  12   24    -1.32    -1.02    0.500     1.50
  12   25    -1.32    -1.02    0.500     1.50
  13   26     3.03     4.41    0.333     1.50
  13   27     3.03     4.41    0.333     1.50
  13   28     3.03     4.41    0.333     1.50
####
Fitting scores
RMSD =    0.387 Hz
Cornilescu Q factor =    0.041
Cornilescu Q factor with experimental errors =    0.041
Cornilescu Q factor,<r3> averaged =    0.028
Cornilescu Q factor <r3> averaged, with experimental errors =    0.028
Chi^2 =    0.934
Number of rdcs in input =  14
Number of rdcs for fit =  14
Number of non equivalent rdcs= 9
n/Chi^2 9.63893075305
Number of parameters k = 7
AIC = Chi^2 +2k =   14.934
```

Short-range + $^2D_{\mathrm{C10–H1}}$ + $^2D_{\mathrm{C4–H4}\beta}$ RDCs, 6A+6B+6C+6D ensemble

```
Alignment tensor
[[  1.26477279e-04  -7.01526139e-05  -4.06097500e-04]
 [ -7.01526139e-05   4.92210507e-04   4.58553466e-04]
 [ -4.06097500e-04   4.58553466e-04  -6.18687786e-04]]
GDO (*1.0e6)= 1183.74742283
Alignment tensor eigenvalues
[-0.00091062+0.j  0.00017477+0.j  0.00073585+0.j]
Alignment tensor eigenvectors
[[ 0.33392159  0.87113029  0.36003944]
 [-0.27773689  0.45593635 -0.84556742]
 [ 0.90075446 -0.18235698 -0.394192  ]]

[[ 0.33392159 -0.27773689  0.90075446]
 [ 0.87113029  0.45593635 -0.18235698]
 [ 0.36003944 -0.84556742 -0.394192  ]]
Ordered alignment tensor ( |Ax| < |Ay| < |Az| )
Ax = 1.748e-04
Ay = 7.359e-04
Az = -9.106e-04
Principal frame eigenvectors
X'= [ 0.871, 0.456,-0.182]
Y'= [ 0.360,-0.846,-0.394]
Z'= [ 0.334,-0.278, 0.901]
#####
Experimental and back calculated RDCs
 I    J    RDC(exp) RDC(calc) weight   std err
   8   17     8.59     8.30    1.000     1.50
   1   14   -13.95   -14.07    1.000     1.50
   3   15   -30.07   -30.01    1.000     1.50
  13   17    -1.23    -1.20    1.000     1.50
  11   24     5.25     5.25    1.000     1.50
   9   19     2.09     2.09    0.500     1.50
   9   18     2.09     2.09    0.500     1.50
  11   22     2.87     2.95    0.500     1.50
  11   23     2.87     2.95    0.500     1.50
  12   24    -1.32    -1.02    0.500     1.50
  12   25    -1.32    -1.02    0.500     1.50
  13   26     3.03     4.41    0.333     1.50
  13   27     3.03     4.41    0.333     1.50
```

```
  13   28     3.03     4.41    0.333     1.50
####
Fitting scores
RMSD =    0.387 Hz
Cornilescu Q factor =    0.041
Cornilescu Q factor with experimental errors =    0.041
Cornilescu Q factor,<r3> averaged =    0.028
Cornilescu Q factor <r3> averaged, with experimental errors =    0.028
Chi^2 =    0.934
Number of rdcs in input =  14
Number of rdcs for fit =  14
Number of non equivalent rdcs= 9
n/Chi^2 9.63987811504
Number of parameters k = 8
AIC = Chi^2 +2k =   16.934
```

Appendix C

Appendix C: NMR experiments pulse programs in Bruker format

All pulse sequences used along this thesis were the standard ones from Bruker library. The only non-standar pulse sequences used were the following (in Bruker TS3.x syntax).

C.1 1D steady-state NOESY with zero-quantum coherence filter

```
#include <Avance.incl>
#include <Grad.incl>

"d0=3u"
"d3=d2-3u-50u-300u-p11-100u-d5"

1 ze

2 50u BLKGRAD
  d1 pl2:f1

3 p1 ph1
  d0

  p1 ph2
  3u pl0:f1
  50u UNBLKGRAD
```

```
  300u gron0
  p11:sp1:f1 ph4
  d5 gron5
  100u groff
  d3 pl2:f1
  p1 ph3

  go=2 ph31

  50u BLKGRAD
  d1 wr #0 if #0 id0 ip1 zd
  lo to 3 times td1

exit

ph1 =0 2
ph2 =0
ph3 =0
ph4 =0
ph31=0 2


;p1  : f1 channel -  90 degree high power pulse
;p11 : duration of sweep
;d0  : incremented delay
;d1  : relaxation delay
;d2  : mixing time
;d5  : homospoil duration
;pl0 : zero power (120 dB)
;pl1 : zero power (120 dB)
;pl2 : high power
;sp1 : power for sweep
;gpz0: gradient strength for ZQ suppression
;gpz5: homospoil gradient strength
;in0 : 1/(2 * SW) = DW
;nd0 : 2
;NS  : 2 * n
;DS  : 8
;td1 : number of t1 increments
;MC2 : TPPI
```

C.2 HSQC-TOCSY(DIPSI2)

```
;hsqcdietgpsisp
;avance-version (12/01/11)
;HSQC-TOCSY
;2D H-1/X correlation via double inept transfer
;   using sensitivity improvement and DIPSI2
;   for homonuclear Hartman-Hahn mixing
;phase sensitive using Echo/Antiecho-TPPI gradient selection
;with decoupling during acquisition - using f2 (and f3)
;using trim pulses in inept transfer
;
;$CLASS=HighRes
;$DIM=2D
;$TYPE=
;$SUBTYPE=
;$COMMENT=


#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>


"p2=p1*2"
"p4=p3*2"
"d11=30m"
"d4=1s/(cnst2*4)"

#   ifdef LABEL_CN
"p22=p21*2"
#   else
#   endif /*LABEL_CN*/


"d0=3u"

"in0=inf1/2"


"FACTOR1=(d9/(p6*115.112))/2"
"l1=FACTOR1*2"
```

```
"DELTA1=p16+d16-p1*0.78+de+8u"
"DELTA2=d4-p14/2"

#   ifdef LABEL_CN
"DELTA=p16+d16+larger(p2,p22)+d0*2"
#   else
"DELTA=p16+d16+p2+d0*2"
#   endif /*LABEL_CN*/


"acqt0=0"
baseopt_echo


1 ze

#   ifdef LABEL_CN
  d11 pl12:f2 pl16:f3
2 d1 do:f2 do:f3
  10u pl0:f2 pl3:f3
#   else
  d11 pl12:f2
2 d1 do:f2
  10u pl0:f2
#   endif /*LABEL_CN*/

3 (p1 ph1)
  DELTA2
  4u
  (center (p2 ph1) (p14:sp3 ph6):f2 )
  4u
  DELTA2 pl2:f2 UNBLKGRAD
  p28 ph1
  4u
  (p1 ph2) (p3 ph3):f2
  d0

#   ifdef LABEL_CN
  (center (p2 ph7) (p22 ph1):f3 )
#   else
  (p2 ph7)
#   endif /*LABEL_CN*/
```

```
  d0
  p16:gp1*EA
  d16
  (p4 ph4):f2
  DELTA
  (center (p1 ph1) (p3 ph4):f2 )
  d24
  (center (p2 ph1) (p4 ph1):f2 )
  d24
  (center (p1 ph2) (p3 ph5):f2 )
  DELTA2 pl0:f2
  (center (p2 ph1) (p14:sp3 ph1):f2 )
  DELTA2 pl10:f1

;begin DIPSI2
4 p6*3.556 ph22
  p6*4.556 ph24
  p6*3.222 ph22
  p6*3.167 ph24
  p6*0.333 ph22
  p6*2.722 ph24
  p6*4.167 ph22
  p6*2.944 ph24
  p6*4.111 ph22

  p6*3.556 ph24
  p6*4.556 ph22
  p6*3.222 ph24
  p6*3.167 ph22
  p6*0.333 ph24
  p6*2.722 ph22
  p6*4.167 ph24
  p6*2.944 ph22
  p6*4.111 ph24

  p6*3.556 ph24
  p6*4.556 ph22
  p6*3.222 ph24
  p6*3.167 ph22
  p6*0.333 ph24
  p6*2.722 ph22
  p6*4.167 ph24
  p6*2.944 ph22
  p6*4.111 ph24
```

```
  p6*3.556 ph22
  p6*4.556 ph24
  p6*3.222 ph22
  p6*3.167 ph24
  p6*0.333 ph22
  p6*2.722 ph24
  p6*4.167 ph22
  p6*2.944 ph24
  p6*4.111 ph22
  lo to 4 times l1
;end DIPSI2

  4u pl1:f1
  (p1 ph1)
  DELTA1
  (p2 ph1)
  4u
  p16:gp2

#   ifdef LABEL_CN
  d16 pl12:f2 pl16:f3
  4u BLKGRAD
  go=2 ph31 cpd2:f2 cpd3:f3
  d1 do:f2 do:f3 mc #0 to 2
#   else
  d16 pl12:f2
  4u BLKGRAD
  go=2 ph31 cpd2:f2
  d1 do:f2 mc #0 to 2
#   endif /*LABEL_CN*/

     F1EA(calgrad(EA) & calph(ph5, +180), caldel(d0, +in0) & calph(ph3, +180) &
     calph(ph6, +180) & calph(ph31, +180))
exit


ph1=0
ph2=1
ph3=0 2
ph4=0 0 2 2
ph5=1 1 3 3
ph6=0
ph7=0 0 2 2
```

```
ph22=3
ph24=1
ph31=0 2 2 0


;pl0 : 0W
;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl10: f1 channel - power level for TOCSY-spinlock
;pl12: f2 channel - power level for CPD/BB decoupling
;pl16: f3 channel - power level for CPD/BB decoupling
;sp3: f2 channel - shaped pulse 180 degree
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel -  90 degree high power pulse
;p4 : f2 channel - 180 degree high power pulse
;p6 : f1 channel -  90 degree low power pulse
;p14: f2 channel - 180 degree shaped pulse for inversion
;p16: homospoil/gradient pulse                    [1 msec]
;p22: f3 channel - 180 degree high power pulse
;p28: f1 channel - trim pulse                     [1 msec]
;d0 : incremented delay (2D)                      [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d4 : 1/(4J(XH))
;d9 : TOCSY mixing time
;d11: delay for disk I/O                          [30 msec]
;d16: delay for homospoil/gradient recovery
;d24: 1/(8J)XH for all multiplicities
;     1/(4J)XH for XH
;cnst2: = J(XH)
;l1: loop for DIPSI cycle: ((p6*115.112) * l1) = mixing time
;inf1: 1/SW(X) = 2 * DW(X)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;ns: 1 * n
;ds: >= 16
;td1: number of experiments
;FnMODE: echo-antiecho
;cpd2: decoupling according to sequence defined by cpdprg2
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence
```

```
;use gradient ratio: gp 1 : gp 2
;   80 : 20.1       for C-13
;   80 :  8.1       for N-15

;for z-only gradients:
;gpz1: 80%
;gpz2: 20.1% for C-13, 8.1% for N-15

;use gradient files:
;gpnam1: SMSQ10.100
;gpnam2: SMSQ10.100



                                        ;preprocessor-flags-start
;LABEL_CN: for C-13 and N-15 labeled samples start experiment with
;option -DLABEL_CN (eda: ZGOPTNS)
                                        ;preprocessor-flags-end



;$Id: hsqcdietgpsisp,v 1.8.2.1.4.1 2012/01/31 17:56:32 ber Exp $
```

C.3 SJS-HSQC

```
; author: Jinfa Ying October 1, 2010
#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"
"p4=p3*2"
"d4=1s/(cnst2*4)"  ;cnst2 set to 155 Hz
"d11=30m"

"d0=3u"
"d29=0.85*p6/2 - p14/2"
"d30=0.15*p6/2"
"d28=0.85*p6/2 - p14/2"
"d3=d1*0.25"
"in0=inf1/2"
```

```
"p26=p16*0.251"
"DELTA1=p26+d16+8u"
"DELTA2=d4-larger(p2,p14)/2"
"DELTA=p16*0.2+50u-d0*2-p3*1.26-p1*2"
"DELTA4=d4*2-p14/2"

1 ze
  d11 pl12:f2
2 d1 do:f2
3 (p1 ph1)
  DELTA4 pl0:f2
  4u
  (center (p2 ph1) (p14:sp3 ph6):f2 )
;500us hyperbolic secant for sp3
  4u
  DELTA4
  (p1 ph11)
  d3
  (p1 ph1)
  DELTA2 pl0:f2
  4u
  (center (p2 ph1) (p14:sp3 ph6):f2 )
;500us hyperbolic secant for sp3
  4u
  DELTA2 pl2:f2
  (p1 ph2)
  3u pl2:f2
  3u pl0:f1

  (d0*20 p6:sp10 ph8):f1 (d29 p3 ph3 3u pl0 d0*20 d30 p14:sp3 ph3 3u
  d30 d0*20 d29):f2
;6.5-25ms center-lobe sinc pulse for sp10

  50u UNBLKGRAD
  p16*0.6:gp1*EA
;for older Bruker spectrometers, assign a new pulse name to p16*0.6
  6u pl0:f2
  (p14:sp3 ph4):f2
;500us hyperbolic secant for sp3
  6u
  p16*0.4:gp3*EA
;for older Bruker spectrometers, assign a new pulse name to p16*0.4
  DELTA pl2:f2
  d29
```

```
  d0 pl1:f1
  (p1*2 ph0):f1
  d0 pl2:f2
  (center (p1 ph1) (p3 ph4):f2 )
  d24
  (center (p2 ph1) (p4 ph1):f2 )
  d24
  (center (p1 ph2) (p3 ph5):f2 )
  DELTA2 pl0:f2
  (center (p2 ph1) (p14:sp3 ph1):f2 )
;500us hyperbolic secant for sp3
  DELTA2
  (p1 ph1)
  DELTA1
  (p2 ph1)
  4u
  p26:gp2
  d16 pl12:f2
  4u BLKGRAD
  go=2 ph31 cpd2:f2
  d1 do:f2 mc #0 to 2
     F1EA(igrad EA & ip5*2, id0 & ip3*2 & ip6*2 & ip13*2 & ip31*2)
exit

ph0=0
ph1=0
ph2=1
ph3=0 2
ph13=0
ph4=0 0 2 2
ph14=1 1 3 3
ph5=1 1 3 3
ph6=0
ph7=0 0 2 2
ph8=0 0 0 0 2 2 2 2
ph11=2
ph31=0 2 2 0
```

C.4 SJS-HSQC from Bruker's library

Written by Wolfang Bermell

```
;hsqcetgpsisp_sjs2.2.t2.be
;avance-version (15/05/16)
;HSQC
;2D H-1/X correlation via double inept transfer
;   using sensitivity improvement
;phase sensitive using Echo/Antiecho-TPPI gradient selection
;with decoupling during acquisition
;using trim pulses in inept transfer
;using shaped pulses for all 180degree pulses on f2 - channel
;with gradients in back-inept
;
;P. Trigo-Mourino, A. Navarro-Vazquez, J. Ying, R.R. Gil & A. Bax,
;   Angew. Chem. Int. Ed. 50, 7576-7580 (2011)
;(A.G. Palmer III, J. Cavanagh, P.E. Wright & M. Rance, J. Magn.
;   Reson. 93, 151-170 (1991) )
;(L.E. Kay, P. Keifer & T. Saarinen, J. Am. Chem. Soc. 114,
;   10663-5 (1992) )
;(J. Schleucher, M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky,
;   S.J. Glaser, O.W. Sorensen & C. Griesinger, J. Biomol. NMR 4,
;   301-306 (1994) )
;
;$CLASS=HighRes
;$DIM=2D
;$TYPE=
;$SUBTYPE=
;$COMMENT=


#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>


"p2=p1*2"
"d2=1s/(cnst2*2)"
"d4=1s/(cnst2*4)"
"d11=30m"


"d0=3u"
"d20=3u"

"in0=inf1/2"
```

```
"in20=in0*cnst16"


"DELTA=p2+d0*2+p3*4/PI+4u"
"DELTA1=p16+d16-p1*0.78+de+8u"
"DELTA2=d4-larger(p2,p14)/2-4u"
"DELTA3=d24-cnst17*p24/2-p19-d16-4u"
"DELTA4=d4-larger(p2,p14)/2-p16-d16-4u"
"DELTA5=d2-larger(p2,p14)/2"

"TAU=p1*2+d2*2+d7"
"TAU1=p12*cnst18/2-p14/2"
"TAU2=p12*(1-cnst18)/2"


#   ifdef CALC_SPOFFS
"spoff2=bf1*(cnst21/1000000)-o1"
#   else
#   endif /*CALC_SPOFFS*/


"acqt0=0"
baseopt_echo


1 ze
  d11 pl12:f2
2 d11 do:f2
3 d1 pl1:f1

#   ifdef BIRD_FLAG
  (p1 ph1)
  DELTA5
  (center (p2 ph1) (p14:sp3 ph1):f2 )
  DELTA5
  (p1 ph8)
  d7
#   else
  TAU
#   endif /*BIRD_FLAG*/

  (p1 ph1)
  DELTA2
  4u
```

```
(center (p2 ph1) (p14:sp3 ph6):f2 )
4u
DELTA2 p12:f2 UNBLKGRAD
(p1 ph2)

(d20 p12:sp2 ph7) (TAU1 p3 ph3 d20 TAU2 p14:sp3 ph3 TAU2 d20 TAU1):f2

p16:gp1*EA*-1
d16 pl1:f1
DELTA
(p14:sp3 ph4):f2
p16:gp1*EA
d16
TAU1
d0
(p2 ph1)
d0
4u pl2:f2
(center (p1 ph1) (p3 ph4):f2 )
4u
p19:gp3
d16
DELTA3
(center (p2 ph1) (p24:sp7 ph1):f2 )
4u
DELTA3 pl2:f2
p19:gp3
d16
(center (p1 ph2) (p3 ph5):f2 )
4u
p16:gp4
d16
DELTA4
(center (p2 ph1) (p14:sp3 ph1):f2 )
4u
DELTA4
p16:gp4
d16
(p1 ph1)
DELTA1
(p2 ph1)
4u
p16:gp2
d16 pl12:f2
```

```
  4u BLKGRAD
  go=2 ph31 cpd2:f2
  d11 do:f2 mc #0 to 2
     F1EA(calgrad(EA) & calph(ph5, +180), caldel(d0, +in0) & caldel(d20, +in20) &
     calph(ph3, +180) & calph(ph6, +180) & calph(ph31, +180))
exit


ph1=0
ph2=1
ph3=0 2
ph4=0 0 2 2
ph5=3 3 1 1
ph6=0
ph7=0 0 0 0 2 2 2 2
ph8=2
ph31=2 0 0 2


;pl0 : 0W
;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl12: f2 channel - power level for CPD/BB decoupling
;sp3: f2 channel - shaped pulse (180degree inversion)
;spnam3: Crp60,0.5,20.1
;sp7: f2 channel - shaped pulse (180degree refocussing)
;spnam7: Crp60comp.4
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel -  90 degree high power pulse
;p14: f2 channel - 180 degree shaped pulse for inversion
;     = 500usec for Crp60,0.5,20.1
;p16: homospoil/gradient pulse                         [1 msec]
;p19: gradient pulse 2                                 [500 usec]
;p22: f3 channel - 180 degree high power pulse
;p24: f2 channel - 180 degree shaped pulse for refocussing
;     = 2msec for Crp60comp.4
;p28: f1 channel - trim pulse
;d0 : incremented delay (2D)                           [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d4 : 1/(4J)XH
;d7 : ca. 0.25*d1
;d11: delay for disk I/O                               [30 msec]
```

```
;d16: delay for homospoil/gradient recovery
;d24: 1/(8J)XH for all multiplicities
;     1/(4J)XH for XH
;cnst2: = J(XH)
;cnst17: = -0.5 for Crp60comp.4

;cnst18 = 0.85

;inf1: 1/SW(X) = 2 * DW(X)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;NS: 1 * n
;DS: >= 16
;td1: number of experiments
;FnMODE: echo-antiecho
;cpd2: decoupling according to sequence defined by cpdprg2
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence


;for z-only gradients:
;gpz1: 80%
;gpz2: 40.2% for C-13
;gpz3: 11%
;gpz4: -5%

;use gradient files:
;gpnam1: SMSQ10.100
;gpnam2: SMSQ10.100
;gpnam3: SMSQ10.100
;gpnam4: SMSQ10.100


;cnst17: Factor to compensate for coupling evolution during a pulse
;       (usually +1). A positive factor indicates that coupling
;       evolution continues during the pulse, whereas a negative
;       factor is necessary if the coupling is (partially) refocussed.



;$Id: $
```

C.5 CLIP-HSQC

```
;hsqcetgpsp.2
;avance-version (07/04/04)
;HSQC
;2D H-1/X correlation via double inept transfer
;phase sensitive using Echo/Antiecho-TPPI gradient selection
;with decoupling during acquisition
;using trim pulses in inept transfer
;using shaped pulses for inversion and refocussing on f2 - channel
;
;$CLASS=HighRes
;$DIM=2D
;$TYPE=
;$SUBTYPE=
;$COMMENT=


#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>


"p2=p1*2"
"d4=1s/(cnst2*4)"
"d11=30m"

#   ifdef LABEL_CN
"p22=p21*2"
#   else
#   endif /*LABEL_CN*/


"d0=3u"

"in0=inf1/2"


"DELTA1=d4-p16-larger(p2,p14)/2-8u-p3"
"DELTA2=d4-larger(p2,p14)/2"

#   ifdef LABEL_CN
"DELTA=p16+d16+larger(p2,p22)+d0*2"
```

```
#   else
"DELTA=p16+d16+p2+d0*2"
#   endif /*LABEL_CN*/


1 ze
  d11 pl12:f2
2 d1 do:f2
3 (p1 ph1)
  DELTA2 pl0:f2
  4u
  (center (p2 ph1) (p14:sp3 ph6):f2 )
  4u
  DELTA2 pl2:f2 UNBLKGRAD
  p28 ph1
  4u
  (p1 ph2) (p3 ph3):f2
  d0

#   ifdef LABEL_CN
  (center (p2 ph5) (p22 ph1):f3 )
#   else
  (p2 ph5)
#   endif /*LABEL_CN*/

  d0
  p16:gp1*EA
  d16 pl0:f2
  4u
  (p24:sp7 ph4):f2
  4u
  DELTA pl2:f2
  (ralign (p1 ph1) (p3 ph4):f2 )
  DELTA2 pl0:f2
  (center (p2 ph1) (p14:sp3 ph1):f2 )
  4u
  p16:gp2
  DELTA1 pl2:f2
  4u BLKGRAD
  (p3 ph7):f2
  go=2 ph31 ;cpd2:f2
  d1 do:f2 mc #0 to 2
     F1EA(igrad EA, id0 & ip3*2 & ip6*2 & ip31*2)
exit
```

```
ph1=0
ph2=1
ph3=0 2
ph4=0 0 2 2
ph5=0 0 0 0 2 2 2 2
ph6=0
ph7=0 2
ph31=2 0 0 2


;pl0 : 120dB
;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;pl3 : f3 channel - power level for pulse (default)
;pl12: f2 channel - power level for CPD/BB decoupling
;sp3: f2 channel - shaped pulse 180 degree for inversion
;sp7: f2 channel - shaped pulse 180 degree for refocussing
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel -  90 degree high power pulse
;p14: f2 channel - 180 degree shaped pulse for inversion
;p16: homospoil/gradient pulse
;p22: f3 channel - 180 degree high power pulse
;p24: f2 channel - 180 degree shaped pulse for refocussing
;p28: f1 channel - trim pulse
;d0 : incremented delay (2D)                          [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d4 : 1/(4J)XH
;d11: delay for disk I/O                              [30 msec]
;d16: delay for homospoil/gradient recovery
;cnst2: = J(XH)
;inf1: 1/SW(X) = 2 * DW(X)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;NS: 1 * n
;DS: >= 16
;td1: number of experiments
;FnMODE: echo-antiecho
;cpd2: decoupling according to sequence defined by cpdprg2
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence
```

```
;use gradient ratio: gp 1 : gp 2
;   80 : 20.1    for C-13
;   80 :  8.1    for N-15

;for z-only gradients:
;gpz1: 80%
;gpz2: 20.1% for C-13, 8.1% for N-15

;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.100



                                        ;preprocessor-flags-start
;LABEL_CN: for C-13 and N-15 labeled samples start experiment with
;             option -DLABEL_CN (eda: ZGOPTNS)
                                        ;preprocessor-flags-end



;$Id: hsqcetgpsp.2,v 1.5 2007/04/11 13:34:30 ber Exp $
```

C.6 F_1-coupled HSQC

```
;hsqcetgpsisp2.2
;avance-version (07/04/04)
;HSQC
;2D H-1/X correlation via double inept transfer
;   using sensitivity improvement
;phase sensitive using Echo/Antiecho-TPPI gradient selection
;with decoupling during acquisition
;using trim pulses in inept transfer
;using shaped pulses for all 180degree pulses on f2 - channel
;with gradients in back-inept
;
;A.G. Palmer III, J. Cavanagh, P.E. Wright & M. Rance, J. Magn.
;   Reson. 93, 151-170 (1991)
;L.E. Kay, P. Keifer & T. Saarinen, J. Am. Chem. Soc. 114,
;   10663-5 (1992)
;J. Schleucher, M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky,
;   S.J. Glaser, O.W. Sorensen & C. Griesinger, J. Biomol. NMR 4,
```

```
;   301-306 (1994)
;
;$CLASS=HighRes
;$DIM=2D
;$TYPE=
;$SUBTYPE=
;$COMMENT=


#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>


"p2=p1*2"
"d4=1s/(cnst2*4)"
"d11=30m"


"d0=3u"

"in0=inf1/2"


"DELTA=p16+d16+d0*2-4u"
"DELTA1=p16+d16+8u"
"DELTA2=d4-larger(p2,p14)/2-4u"
"DELTA3=d24-cnst17*p24/2-p19-d16-4u"
"DELTA4=d4-larger(p2,p14)/2-p16-d16-4u"


1 ze
  d11 pl12:f2
2 d1 do:f2
3 (p1 ph1)
  DELTA2 pl0:f2
  4u
  (center (p2 ph1) (p14:sp3 ph6):f2 )
  4u
  DELTA2 pl2:f2 UNBLKGRAD
  p28 ph1
  4u
  (p1 ph2) (p3 ph3):f2
  d0
```

```
  d0
  p16:gp1*EA
  d16 pl0:f2
  (p24:sp7 ph8:r):f2
  4u
  DELTA pl2:f2
  (center (p1 ph1) (p3 ph4):f2 )
  4u
  p19:gp3
  d16
  DELTA3 pl0:f2
  (center (p2 ph1) (p24:sp7 ph9:r):f2 )
  4u
  DELTA3 pl2:f2
  p19:gp3
  d16
  (center (p1 ph2) (p3 ph5):f2 )
  4u
  p16:gp4
  d16
  DELTA4 pl0:f2
  (center (p2 ph1) (p14:sp3 ph1):f2 )
  4u
  DELTA4
  p16:gp4
  d16
  (p1 ph1)
  DELTA1
  (p2 ph1)
  4u
  p16:gp2
  d16 pl12:f2
  4u BLKGRAD
  go=2 ph31 cpd2:f2
  d1 do:f2 mc #0 to 2
     F1EA(igrad EA & ip5*2, id0 & ip3*2 & ip6*2 & ip31*2)
exit


ph1=0
ph2=1
ph3=0 2
ph4=0 0 2 2
ph5=1 1 3 3
```

```
ph6=0
ph7=0 0 2 2
ph8=0 0 2 2
ph9=0
ph31=0 2 2 0
```

C.7 F_1-coupled *J*-scaled HSQC

```
; author: Jinfa Ying October 1, 2010
;hsqcetgpsp
;avance-version (07/04/04)
;HSQC
;2D H-1/X correlation via double inept transfer
;phase sensitive using Echo/Antiecho-TPPI gradient selection
;with decoupling during acquisition
;using trim pulses in inept transfer
;using shaped pulses for inversion on f2 - channel
;
;$CLASS=HighRes
;$DIM=2D
;$TYPE=
;$SUBTYPE=
;$COMMENT=


;$OWNER=deyp01
#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>


"p2=p1*2"
"p4=p3*2"
"d4=1s/(cnst2*4)"
"d11=30m"

"d3=d1*0.25"

"d0=3u"
"d10=p2-p3"
;"d10=p3-p2"
"d8=d10"
```

```
"in0=inf1/2"
"in10=in0*cnst10"

"DELTA1=d4-p26-larger(p2,p14)/2-8u"
"DELTA2=d4-larger(p2,p14)/2"

;DELTA, when p2 longer than p3
"DELTA=p16-p17+d16+d0*2-d8*2-10u-p3*1.26"
;DELTA, when p2 shorter than p3
;"DELTA=p16-p17+d16+d0*2-10u-p3*1.26"

;"DELTA=p16*0.2+d16+d0*2-10u-p3*1.26"
"DELTA4=d4*2-p14/2"

1 ze
  d11 pl12:f2
2 d1 do:f2
3 (p1 ph1)
   DELTA4 pl0:f2
   4u
  (center (p2 ph1) (p14:sp3 ph6):f2 )
   4u DELTA4
  (p1 ph11)
   d3
  (p1 ph1)
   DELTA2 pl0:f2
   4u
  (center (p2 ph1) (p14:sp3 ph6):f2 )
   4u
   DELTA2 pl2:f2 UNBLKGRAD
  (p1 ph2)
   5u
  (p3 ph3 d10):f2 (d10 p2 ph8 d8):f1
   3u pl0:f2
  (p14:sp3 ph1):f2
   3u
   d10
   d0
   d0
   p16:gp1*EA
   d16
  (p14:sp3 ph4):f2
   10u
```

```
   p17:gp3*EA
   DELTA p12:f2
  (ralign (p1 ph1) (p3 ph4):f2 )
   DELTA2 p10:f2
  (center (p2 ph1) (p14:sp3 ph1):f2 )
   4u
   p26:gp2
   DELTA1 p112:f2
   4u BLKGRAD
  go=2 ph31 cpd2:f2
  d1 do:f2 mc #0 to 2
     F1EA(igrad EA, id0 & id10 & ip3*2 & ip6*2 & ip31*2)
exit


ph1=0
ph2=1
ph3=0 2
ph4=0 0 0 0 2 2 2 2
ph5=0 0 2 2
ph6=0
ph8=0
ph11=2
ph31=0 2 0 2 2 0 2 0
```

Bibliography

[1] F. Bloch, W. W. Hansen, M. Packard, Nuclear Induction, *Phys. Rev.* **1946**, *69*, 127–127, DOI 10.1103/PhysRev.69.127.

[2] J. T. Arnold, S. S. Dharmatti, M. E. Packard, Chemical Effects on Nuclear Induction Signals from Organic Compounds, *The Journal of Chemical Physics* **1951**, *19*(4), 507–507, DOI 10.1063/1.1748264.

[3] W. D. Knight, Nuclear Magnetic Resonance Shift in Metals, *Phys. Rev.* **1949**, *76*, 1259–1260, DOI 10.1103/PhysRev.76.1259.2.

[4] W. G. Proctor, F. C. Yu, The Dependence of a Nuclear Magnetic Resonance Frequency upon Chemical Compound, *Phys. Rev.* **1950**, *77*, 717–717, DOI 10.1103/PhysRev.77.717.

[5] W. C. Dickinson, Dependence of the F^{19} Nuclear Resonance Position on Chemical Compound, *Phys. Rev.* **1950**, *77*, 736–737, DOI 10.1103/PhysRev.77.736.2.

[6] W. D. Aue, E. Bartholdi, R. R. Ernst, Two-dimensional spectroscopy. Application to nuclear magnetic resonance, *J. Chem. Phys.* **1976**, *64*(5), 2229–2246, DOI 10.1063/1.432450.

[7] R. R. Ernst, Nuclear Magnetic Resonance Fourier Transform Spectroscopy (Nobel Lecture), *Angew. Chem. Int. Ed. Eng.* **1992**, *31*(7), 805–823, DOI 10.1002/anie.199208053.

[8] K. Wüthrich, NMR Studies of Structure and Function of Biological Macromolecules (Nobel Lecture), *Angew. Chem. Int. Ed. Eng.* **2003**, *42*(29), 3340–3363, DOI 10.1002/anie.200300595.

[9] M. Karplus, Vicinal Proton Coupling in Nuclear Magnetic Resonance, *J. Am. Chem. Soc.* **1963**, *85*(18), 2870–2871, DOI 10.1021/ja00901a059.

[10] M. Karplus, Comments on the Signs of Proton Coupling Constants, *J. Am. Chem. Soc.* **1962**, *84*(12), 2458–2460, DOI 10.1021/ja00871a044.

[11] M. Karplus, Theory of Proton Coupling Constants in Unsaturated Molecules, *J. Am. Chem. Soc.* **1960**, *82*(16), 4431–4432, DOI 10.1021/ja01501a080.

[12] C. Altona, *Vicinal Coupling Constants and Conformation of Biomolecules*, John Wiley Sons, Ltd, **2007**, DOI 10.1002/9780470034590.emrstm0587.

[13] F. A. L. Anet, A. J. R. Bourn, Nuclear Magnetic Resonance Spectral Assignments from Nuclear Overhauser Effects, *J. Am. Chem. Soc.* **1965**, *87*(22), 5250–5251, DOI 10.1021/ja00950a048.

[14] M. C. Woods, H.-C. Chiang, Y. Nakadaira, K. Nakanishi, Nuclear Overhauser effect, a unique method of defining the relative stereochemistry and conformation of taxane derivatives, *J. Am. Chem. Soc.* **1968**, *90*(2), 522–523, DOI 10.1021/ja01004a074.

[15] M. C. Woods, I. Miura, Y. Nakadaira, A. Terahara, M. Maruyama, K. Nakanishi, The ginkgolides. V. Some aspects of their NMR spectra, *Tetrahedron Lett.* **1967**, *8*(4), 321 – 326, DOI 10.1016/S0040-4039(00)71542-5.

[16] R. R. Ernst, G. Bodenhausen, A. Wokaun, *Principles of nuclear magnetic resonance in one and two dimensions*, Clarendon Press Oxford, **1987**.

[17] J. Cavanagh, *Protein NMR spectroscopy: principles and practice*, Academic Press, **2007**.

[18] M. H. Levitt, *Spin Dynamics: Basics of Nuclear Magnetic Resonance*, Wiley, **2008**.

[19] J. Keeler, *Understanding NMR spectroscopy*, Wiley, **2010**.

[20] A. Abragam, *Principles of nuclear magnetism*, Bd. 32, Oxford University Press, USA, **1983**.

[21] S. A. Smith, W. E. Palke, J. T. Gerig, The Hamiltonians of NMR. part I, *Conc. Magn. Reson.* **1992**, *4*(2), 107–144, DOI 10.1002/cmr.1820040202.

[22] S. A. Smith, W. E. Palke, J. T. Gerig, The hamiltonians of NMR. Part II, *Conc. Magn. Reson.* **1992**, *4*(3), 181–204, DOI 10.1002/cmr.1820040302.

[23] S. A. Smith, W. E. Palke, J. T. Gerig, The Hamiltonians of NMR. Part III, *Conc. Magn. Reson.* **1993**, *5*(2), 151–177, DOI 10.1002/cmr.1820050204.

[24] S. A. Smith, W. E. Palke, J. T. Gerig, The hamiltonians of NMR. Part IV: NMR relaxation, *Conc. Magn. Reson.* **1994**, *6*(2), 137–162, DOI 10.1002/cmr.1820060205.

[25] L. J. Mueller, Tensors and rotations in NMR, *Concepts Magn. Reson.* **2011**, *38A*(5), 221–235, DOI 10.1002/cmr.a.20224.

[26] F. Kramer, M. V. Deshmukh, H. Kessler, S. J. Glaser, Residual dipolar coupling constants: An elementary derivation of key equations, *Concepts Magn. Reson.* **2004**, *21A*(1), 10–21, DOI 10.1002/cmr.a.20003.

[27] M. Blackledge, Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings, *Progr. Nuc. Magn. Reson. Spectrosc.* **2005**, *46*(1), 23–62, DOI 10.1016/j.pnmrs.2004.11.002.

[28] N. F. Ramsey, Magnetic Shielding of Nuclei in Molecules, *Phys. Rev.* **1950**, *78*, 699–703, DOI 10.1103/PhysRev.78.699.

[29] N. F. Ramsey, Dependence of Magnetic Shielding of Nuclei upon Molecular Orientation, *Phys. Rev.* **1951**, *83*, 540–541, DOI 10.1103/PhysRev.83.540.

[30] J. Vaara, J. Jokisaari, R. E. Wasylishen, D. L. Bryce, Spin–spin coupling tensors as determined by experiment and computational chemistry, *Progr. Nucl. Magn. Reson. Spectrosc.* **2002**, *41*(3), 233–304, DOI 10.1016/S0079-6565(02)00050-X.

[31] A. A. Bothner-By, *Computer Analysis of High–Resolution NMR Spectra*, John Wiley Sons, Ltd, **2007**, DOI 10.1002/9780470034590.emrhp0026.

[32] C. A. G. Haasnoot, F. A. A. M. de Leeuw, C. Altona, The relationship between proton-proton NMR coupling constants and substituent electronegativities – I : An empirical generalization of the karplus equation, *Tetrahedron* **1980**, *36*(19), 2783–2792, DOI 10.1016/0040-4020(80)80155-4.

[33] C. M. Thiele, Residual Dipolar Couplings (RDCs) in Organic Structure Determination, *Eur. J. Org. Chem.* **2008**, *2008*(34), 5673–5685, DOI 10.1002/ejoc.200800686.

[34] N. Tjandra, A. Bax, Direct Measurement of Distances and Angles in Biomolecules by NMR in a Dilute Liquid Crystalline Medium, *Science* **1997**, *278*(5340), 1111–1114, DOI 10.1126/science.278.5340.1111.

[35] D. Catalano, L. di Bari, C. A. Veracini, G. N. Shilstone, C. Zannoni, A maximum-entropy analysis of the problem of the rotameric distribution for substituted biphenyls studied by 1H nuclear magnetic resonance spectroscopy in nematic liquid crystals, *J. Chem. Phys.* **1991**, *94*(5), 3928, DOI 10.1063/1.460669.

[36] A. Saupe, Recent Results in the Field of Liquid Crystals, *Angew. Chem. Int. Ed. Engl.* **1968**, *7*(2), 97–112, DOI 10.1002/anie.196800971.

[37] J. W. Emsley, J. C. Lindon, *NMR Spectroscopy Using Liquid Crystal Solvents*, Pergamon Press, **1975**.

[38] A. A. Bothner-By, P. J. Domaille, C. Gayathri, Ultra-high field NMR spectroscopy: observation of proton-proton dipolar coupling in paramagnetic bis[tolyltris(pyrazolyl)borato]cobalt(II), *J. Am. Chem. Soc.* **1981**, *103*(18), 5602–5603, DOI 10.1021/ja00408a069.

[39] H. C. Kung, K. Y. Wang, I. Goljer, P. H. Bolton, Magnetic Alignment of Duplex and Quadruplex DNAs, *J. Magn. Reson. B* **1995**, *109*(3), 323 – 325, DOI 10.1006/jmrb.1995.9987.

[40] A. Saupe, Recent Results in the Field of Liquid Crystals, *Angew. Chem. Int. Ed. Engl.* **1968**, *7*(2), 97–112, DOI 10.1002/anie.196800971.

[41] J. A. Losonczi, M. Andrec, M. W. F. Fischer, J. H. Prestegard, Order Matrix Analysis of Residual Dipolar Couplings Using Singular Value Decomposition, *J. Magn. Reson.* **1999**, *138*(2), 334–342, DOI 10.1006/jmre.1999.1754.

[42] MestReNova Suite, Mestrelab Research http://mestrelab.com/, **2013**.

[43] A. Navarro-Vázquez, MSpin-RDC. A program for the use of residual dipolar couplings for structure elucidation of small molecules, *Magn. Reson. Chem.* **2012**, *50*, S73–S79, DOI 10.1002/mrc.3905.

[44] G. Cornilescu, J. L. Marquardt, M. Ottiger, A. Bax, Validation of Protein Structure from Anisotropic Carbonyl Chemical Shifts in a Dilute Liquid Crystalline Phase, *J. Am. Chem. Soc.* **1998**, *120*(27), 6836–6837, DOI 10.1021/ja9812610.

[45] G. Kummerlöwe, B. Crone, M. Kretschmer, S. F. Kirsch, B. Luy, Residual Dipolar Couplings as a Powerful Tool for Constitutional Analysis: The Unexpected Formation of Tricyclic Compounds, *Angew. Chem. Int. Ed. Engl.* **2011**, *50*(11), 2643–2645, DOI 10.1002/anie.201007305.

[46] D. Intelmann, G. Kummerlöwe, G. Haseleu, N. Desmer, K. Schulze, R. Fröhlich, O. Frank, B. Luy, T. Hofmann, Structures of storage-induced transformation products of the beer's bitter principles, revealed by sophisticated NMR spectroscopic and LC-MS techniques., *Chem. Eur. J.* **2009**, *15*(47), 13047–13058, DOI 10.1002/chem.200902058.

[47] C. M. Thiele, V. Schmidts, B. Böttcher, I. Louzao, R. Berger, A. Maliniak, B. Stevensson, On the Treatment of Conformational Flexibility when Using Residual Dipolar Couplings for Structure Determination, *Angew. Chem. Int. Ed. Engl.* **2009**, *48*(36), 6708–6712, DOI 10.1002/anie.200902398.

[48] G. Kontaxis, G. M. Clore, A. Bax, Evaluation of Cross-Correlation Effects and Measurement of One-Bond Couplings in Proteins with Short Transverse Relaxation Times, *J.Magn. Reson.* **2000**, *143*(1), 184–196, DOI 10.1006/jmre.1999.1979.

[49] G. Kummerlöwe, S. Schmitt, B. Luy, Cross-Fitting of Residual Dipolar Couplings, *Open Spectrosc. J.* **2010**, *4*, 16–27, DOI 10.2174/1874383801004010016.

[50] B. Yu, H. van Ingen, S. Vivekanandan, C. Rademacher, S. E. Norris, D. Freedberg, More accurate $^{1}J_{CH}$ coupling measurement in the presence of $^{3}J_{HH}$ strong coupling in natural abundance, *J. Magn. Reson.* **2012**, *215*, 10–22, DOI 10.1016/j.jmr.2011.09.037.

[51] V. M. Sánchez-Pedregal, R. Santamaría-Fernández, A. Navarro-Vázquez, Residual Dipolar Couplings of Freely Rotating Groups in Small Molecules. Stereochemical Assignment and Side-Chain Conformation of 8-Phenylmenthol, *Org. Lett.* **2009**, *11*(7), 1471–1474, DOI 10.1021/ol8029565.

[52] J. R. Tolman, J. H. Prestegard, Measurement of amide ^{15}N–^{1}H one-bond couplings in proteins using accordion heteronuclear-shift-correlation experiments., *J. Magn. Reson. B* **1996**, *112*(3), 269–274, DOI 10.1006/jmrb.1996.0141.

[53] A. Enthart, J. C. Freudenberger, J. Furrer, H. Kessler, B. Luy, The CLIP/CLAP-HSQC: Pure absorptive spectra for the measurement of one-bond couplings, *J. Magn. Reson.* **2008**, *192*(2), 314–322, DOI 10.1016/j.jmr.2008.03.009.

[54] P. Trigo-Mouriño, R. Santamaría-Fernández, V. M. Sánchez-Pedregal, A. Navarro-Vázquez, Conformational analysis of an isoquinolinium hydrochloride in water using residual dipolar couplings., *J. Org. Chem.* **2010**, *75*(9), 3101–3104, DOI 10.1021/jo902502h.

[55] D. J. Newman, G. M. Cragg, Natural Products as Sources of New Drugs over the Last 25 Years, *J. Nat. Prod.* **2007**, *70*(3), 461–477, DOI 10.1021/np068054v.

[56] P. Thepchatri, T. Eliseo, D. O. Cicero, D. Myles, J. P. Snyder, Relationship Among Ligand Conformations in Solution, in the Solid State, and at the Hsp90 Binding Site: Geldanamycin and Radicicol, *J. Am. Chem. Soc.* **2007**, *129*(11), 3127–3134, DOI 10.1021/ja064863p.

[57] J. Jiménez-Barbero, F. Amat-Guerri, J. P. Snyder, The Solid State, Solution and Tubulin-Bound Conformations of Agents that Promote Microtubule Stabilization, *Curr. Med. Chem. Anti-Cancer Agents* **2002**, *2*(1), 91–122, DOI 10.2174/1568011023354416.

[58] C. M. Thiele, Simultaneous Assignment of All Diastereotopic Protons in Strychnine Using RDCs: PELG as Alignment Medium for Organic Molecules, *J. Org. Chem.* **2004**, *69*(22), 7403–7413, DOI 10.1021/jo049867w.

[59] C. M. Thiele, Use of RDCs in rigid organic compounds and some practical considerations concerning alignment media, *Concepts Magn. Reson.* **2007**, *30A*(2), 65–80, DOI 10.1002/cmr.a.20075.

[60] R. R. Gil, C. Gayathri, N. V. Tsarevsky, K. Matyjaszewski, Stretched Poly(methyl methacrylate) Gel Aligns Small Organic Molecules in Chloroform. Stereochemical Analysis and Diastereotopic Proton NMR Assignment in Ludartin Using Residual Dipolar Couplings and ^3J Coupling Constant Analysis, *J. Org. Chem.* **2008**, *73*(3), 840–848, DOI 10.1021/jo701871g.

[61] M. E. García, S. Pagola, A. Navarro-Vázquez, D. D. Phillips, C. Gayathri, H. Krakauer, P. W. Stephens, V. E. Nicotra, R. R. Gil, Stereochemistry Determination by Powder X-Ray Diffraction Analysis and NMR Spectroscopy Residual Dipolar Couplings, *Angew. Chem. Int. Ed. Engl.* **2009**, *48*(31), 5670–5674, DOI 10.1002/anie.200900760.

[62] J. W. Emsley, G. R. Luckhurst, C. P. Stockley, A theory of orientational ordering in uniaxial liquid crystals composed of molecules with alkyl chains, *Proc. R. Soc. A* **1982**, *381*(1780), 117–138, DOI 10.1098/rspa.1982.0061.

[63] B. Stevensson, C. Landersjö, G. Widmalm, A. Maliniak, Conformational Distribution Function of a Disaccharide in a Liquid Crystalline Phase Determined Using NMR Spectroscopy, *J. Am. Chem. Soc.* **2002**, *124*(21), 5946–5947, DOI 10.1021/ja025751a.

[64] B. Stevensson, D. Sandström, A. Maliniak, Conformational distribution functions extracted from residual dipolar couplings: A hybrid model based on maximum entropy and molecular field theory, *J. Chem. Phys.* **2003**, *119*(5), 2738, DOI 10.1063/1.1586914.

[65] J. Klages, C. Neubauer, M. Coles, H. Kessler, B. Luy, Structure Refinement of Cyclosporin A in Chloroform by Using RDCs Measured in a Stretched PDMS-Gel, *ChemBioChem* **2005**, *6*(9), 1672–1678, DOI 10.1002/cbic.200500146.

[66] U. M. Reinscheid, J. Farjon, M. Radzom, P. Haberz, A. Zeeck, M. Blackledge, C. Griesinger, Effect of the solvent on the conformation of a depsipeptide: NMR-derived solution structure of hormaomycin in DMSO from residual dipolar couplings in a novel DMSO-compatible alignment medium., *ChemBioChem* **2006**, *7*(2), 287–296, DOI 10.1002/cbic.200500277.

[67] C. Farès, J. Hassfeld, D. Menche, T. Carlomagno, Simultaneous Determination of the Conformation and Relative Configuration of Archazolide$_A$ by Using Nuclear

Overhauser Effects, *J* Couplings, and Residual Dipolar Couplings, *Angew. Chem. Int. Ed. Engl.* **2008**, *47*(20), 3722–3726, DOI 10.1002/anie.200800225.

[68] A. Schuetz, T. Murakami, N. Takada, J. Junker, M. Hashimoto, C. Griesinger, RDC-Enhanced NMR Spectroscopy in Structure Elucidation of Sucro-Neolambertellin, *Angew. Chem. Int. Ed. Engl.* **2008**, *47*(11), 2032–2034, DOI 10.1002/anie.200705037.

[69] C. M. Thiele, A. Marx, R. Berger, J. Fischer, M. Biel, A. Giannis, Determination of the Relative Configuration of a Five-Membered Lactone from Residual Dipolar Couplings, *Angew. Chem. Int. Ed. Engl.* **2006**, *45*(27), 4455–4460, DOI 10.1002/anie.200503247.

[70] E. E. Burnell, C. A. de Lange, Effects of interaction between molecular internal motion and reorientation on NMR of anisotropic liquids, *J. Magn. Reson.* **1980**, *39*(3), 461–480, DOI 10.1016/0022-2364(80)90033-5.

[71] C. Eckart, Some studies concerning rotating axes and polyatomic molecules, *Phys. Rev.* **1935**, *47*(7), 552, DOI 10.1103/PhysRev.47.552.

[72] K. N. Kudin, A. Y. Dymarsky, Eckart axis conditions and the minimization of the root-mean-square deviation: Two closely related problems, *J. Chem. Phys.* **2005**, *122*(22), 224105–224102, DOI 10.1063/1.1929739.

[73] W. Kabsch, A solution for the best rotation to relate two sets of vectors, *Acta Crystallogr. Sect. A* **1976**, *32*(5), 922–923, DOI 10.1107/S0567739476001873.

[74] D. I. Freedberg, An Alternative Method for Pucker Determination in Carbohydrates from Residual Dipolar Couplings: A Solution NMR Study of the Fructofuranosyl Ring of Sucrose, *J. Am. Chem. Soc.* **2002**, *124*(10), 2358–2362, DOI 10.1021/ja010981v.

[75] A. Schuetz, J. Junker, A. Leonov, O. F. Lange, T. F. Molinski, C. Griesinger, Stereochemistry of Sagittamide A from Residual Dipolar Coupling Enhanced NMR, *J. Am. Chem. Soc.* **2007**, *129*(49), 15114–15115, DOI 10.1021/ja075876l.

[76] H. Kovacs, D. Moskau, M. Spraul, Cryogenically cooled probes—a leap in NMR technology, *Progr. Nucl. Magn. Reson. Spectrosc.* **2005**, *46*(2-3), 131–155, DOI 10.1016/j.pnmrs.2005.03.001.

[77] J. Furrer, M. John, H. Kessler, B. Luy, J-Spectroscopy in the presence of residual dipolar couplings: determination of one-bond coupling constants and scalable resolution, *J. Biomol. NMR* **2007**, *37*(3), 231–243, DOI 10.1007/s10858-006-9130-x.

[78] L. Verdier, P. Sakhaii, M. Zweckstetter, C. Griesinger, Measurement of long range H,C couplings in natural products in orienting media: a tool for structure elucidation of natural products, *J.Magn. Reson.* **2003**, *163*(2), 353–359, DOI 10.1016/S1090-7807(03)00063-6.

[79] K. Kobzar, B. Luy, Analyses, extensions and comparison of three experimental schemes for measuring ($^{n}J_{CH}$+D_{CH})-couplings at natural abundance, *J. Magn. Reson.* **2007**, *186*(1), 131–141, DOI 10.1016/j.jmr.2007.02.005.

[80] M. Panar, W. D. Phillips, Magnetic ordering of poly(γ-benzyl-L-glutamate) solutions, *J. Am. Chem. Soc.* **1968**, *90*(14), 3880–3882, DOI 10.1021/ja01016a061.

[81] P. Lesot, Y. Gounelle, D. Merlet, A. Loewenstein, J. Courtieu, Measurement and Analysis of the Molecular Ordering Tensors of Two Enantiomers Oriented in a Polypeptide Liquid Crystalline System, *J. Phys. Chem.* **1995**, *99*(40), 14871–14875, DOI 10.1021/j100040a042.

[82] C. M. Thiele, S. Berger, Probing the Diastereotopicity of Methylene Protons in Strychnine Using Residual Dipolar Couplings, *Org. Lett.* **2003**, *5*(5), 705–708, DOI 10.1021/ol0275163.

[83] A. Marx, B. Böttcher, C. M. Thiele, Enhancing the Orienting Properties of Poly(γ-benzyl-L-glutamate) by means of Additives, *Chem. Eur. J.* **2010**, *16*(5), 1656–1663, DOI 10.1002/chem.200902287.

[84] B. Bendiak, Sensitive through-space dipolar correlations between nuclei of small organic molecules by partial alignment in a deuterated liquid solvent, *J. Am. Chem. Soc.* **2002**, *124*(50), 14862–14863, DOI 10.1021/ja028412q.

[85] M. Martín-Pastor, A. Canales, F. Corzana, J. L. Asensio, J. Jiménez-Barbero, Limited flexibility of lactose detected from residual dipolar couplings using molecular dynamics simulations and steric alignment methods., *J. Am. Chem. Soc.* **2005**, *127*(10), 3589–3595, DOI 10.1021/ja043445m.

[86] M. Martín-Pastor, C. A. Bush, The use of NMR residual dipolar couplings in aqueous dilute liquid crystalline medium for conformational studies of complex oligosaccharides., *Carbohydr. Res.* **2000**, *323*(1-4), 147–155, DOI 10.1016/S0008-6215(99)00237-2.

[87] R. Tycko, F. J. Blanco, Y. Ishii, Alignment of biopolymers in strained gels: a new way to create detectable dipole-dipole couplings in high-resolution biomolecular NMR, *J. Am. Chem. Soc.* **2000**, *122*(38), 9340–9341, DOI 10.1021/ja002133q.

[88] H. J. Sass, G. Musco, S. J. Stahl, P. T. Wingfield, S. Grzesiek, Solution NMR of proteins within polyacrylamide gels: diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes, *J. Biomol. NMR* **2000**, *18*(4), 303–309, DOI 10.1023/A:1026703605147.

[89] B. Deloche, E. T. Samulski, Short-range nematic-like orientational order in strained elastomers: a deuterium magnetic resonance study, *Macromolecules* **1981**, *14*(3), 575–581, DOI 10.1021/ma50004a024.

[90] B. Luy, K. Kobzar, H. Kessler, An easy and scalable method for the partial alignment of organic molecules for measuring residual dipolar couplings, *Angew. Chem. Int. Ed. Engl.* **2004**, *43*(9), 1092–1094, DOI 10.1002/anie.200352860.

[91] J. C. Freudenberger, P. Spiteller, R. Bauer, H. Kessler, B. Luy, Stretched Poly(dimethylsiloxane) Gels as NMR Alignment Media for Apolar and Weakly Polar Organic Solvents: An Ideal Tool for Measuring RDCs at Low Molecular Concentrations, *J. Am. Chem. Soc.* **2004**, *126*(45), 14690–14691, DOI 10.1021/ja046155e.

[92] J. C. Freudenberger, S. Knör, K. Kobzar, D. Heckmann, T. Pulat, H. Kessler, B. Luy, Stretched Poly(vinyl acetate) Gels as NMR Alignment Media for the Measurement of Residual Dipolar Couplings in Polar Organic Solvents, *Angew. Chem. Int. Ed. Engl.* **2005**, *44*(3), 423–426, DOI 10.1002/anie.200461241.

[93] G. Kummerlöwe, J. Auernheimer, A. Lendlein, B. Luy, Stretched Poly(acrylonitrile) as a Scalable Alignment Medium for DMSO, *J. Am. Chem. Soc.* **2007**, *129*(19), 6080–6081, DOI 10.1021/ja071248s.

[94] P. Haberz, J. Farjon, C. Griesinger, A DMSO-compatible orienting medium: towards the investigation of the stereochemistry of natural products, *Angew. Chem. Int. Ed. Engl.* **2005**, *44*(3), 427–429, DOI 10.1002/anie.200461267.

[95] G. Kummerlöwe, E. F. McCord, S. F. Cheatham, S. Niss, R. W. Schnell, B. Luy, Tunable alignment for all polymer gel/solvent combinations for the measurement of anisotropic NMR parameters., *Chem. Eur. J.* **2010**, *16*(24), 7087–7089, DOI 10.1002/chem.201000108.

[96] C. Gayathri, N. V. Tsarevsky, R. R. Gil, Residual Dipolar Couplings (RDCs) Analysis of Small Molecules Made Easy: Fast and Tuneable Alignment by Reversible Compression/Relaxation of Reusable PMMA Gels, *Chem. Eur. J.* **2010**, *16*(12), 3622–3626, DOI 10.1002/chem.200903378.

[97] K. Kobzar, H. Kessler, B. Luy, Stretched Gelatin Gels as Chiral Alignment Media for the Discrimination of Enantiomers by NMR Spectroscopy, *Angew. Chem. Int. Ed. Engl.* **2005**, *44*(20), 3145–3147, DOI 10.1002/anie.200462736.

[98] P. Trigo-Mouriño, A. Navarro-Vázquez, V. M. Sánchez-Pedregal, Influence of solvent and salt concentration on the alignment properties of acrylamide copolymer gels for the measurement of RDC, *Magn. Reson. Chem.* **2012**, *50*, S29–S37, DOI 10.1002/mrc.3889.

[99] J. H. Prestegard, C. M. Bougault, A. I. Kishore, Residual dipolar couplings in structure determination of biomolecules, *Chem. Rev.* **2004**, *104*(8), 3519–3540, DOI 10.1021/cr030419i.

[100] J. Yan, A. D. Kline, H. Mo, M. J. Shapiro, E. R. Zartler, A novel method for the determination of stereochemistry in six-membered chairlike rings using residual dipolar couplings., *J. Org. Chem.* **2003**, *68*(5), 1786–1795, DOI 10.1021/jo020670i.

[101] C. Gayathri, M. C. de la Fuente, B. Luy, R. R. Gil, A. Navarro-Vázquez, Probing heterocycle conformation with residual dipolar couplings, *Chem. Comm.* **2010**, *46*(32), 5879, DOI 10.1039/c0cc01271h.

[102] B. Luy, K. Kobzar, S. Knör, J. Furrer, D. Heckmann, H. Kessler, Orientational Properties of Stretched Polystyrene Gels in Organic Solvents and the Suppression of Their Residual ^{1}H NMR Signals, *J. Am. Chem. Soc.* **2005**, *127*(17), 6459–6465, DOI 10.1021/ja043344o.

[103] A.-C. Pöppler, H. Keil, D. Stalke, M. John, ^{7}Li Residual Quadrupolar Couplings as a Powerful Tool To Identify the Degree of Organolithium Aggregation, *Angew. Chem. Int. Ed. Eng.* **2012**, 7843–7846, DOI 10.1002/anie.201202116.

[104] M. D. Carr, E. M. Purcell, Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments, *Phys. Rev* **1954**, *94*(3), 630–639, DOI 10.1103/PhysRev.94.630.

[105] S. Meiboom, D. Gill, Modified Spin-Echo Method for Measuring Nuclear Relaxation Times, *Rev. Sci. Instrum.* **1958**, *29*(8), 688, DOI 10.1063/1.1716296.

[106] M. Schmidt, H. Sun, P. Rogne, G. K. E. Scriba, C. Griesinger, L. T. Kuhn, U. M. Reinscheid, Determining the absolute configuration of (+)-mefloquine HCl, the side-effect-reducing enantiomer of the antimalaria drug Lariam., *J. Am. Chem. Soc.* **2012**, *134*(6), 3080–3083, DOI 10.1021/ja209050k.

[107] H. M. Ge, H. Sun, N. Jiang, Y. H. Qin, H. Dou, T. Yan, Y. Y. Hou, C. Griesinger, R. X. Tan, Relative and absolute configuration of vatiparol (1 mg): a novel anti-inflammatory polyphenol., *Chem. Eur. J.* **2012**, *18*(17), 5213–5221, DOI 10.1002/chem.201104078.

[108] X. Liu, Z. Tong, F. Gao, Novel shrinking–swelling hysteresis in DMSO/THF mixtures of copolymer gels containing sulphonate groups, *Polym. Int.* **1998**, *47*(2), 215–220, DOI 10.1002/(SICI)1097-0126(1998100)47:2<215::AID-PI58>3.0.CO;2-H.

[109] M. Schmidt, H. Sun, A. Leonov, C. Griesinger, U. M. Reinscheid, Chiral discrimination of amines by anisotropic NMR parameters using chiral polyacrylamide-based gels, *Magn. Reson. Chem.* **2012**, *50*, S38–S44, DOI 10.1002/mrc.3886.

[110] F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, W. C. Still, Macromodel—an integrated software system for modeling organic and bioorganic molecules using molecular mechanics, *J. Comp. Chem.* **1990**, *11*(4), 440–467, DOI 10.1002/jcc.540110405.

[111] Serenasoft, PCModel, http://serenasoft.com, **2012**.

[112] N. O'Boyle, M. Banck, C. James, C. Morley, T. Vandermeersch, G. Hutchison, Open Babel: An open chemical toolbox, *J. Cheminf.* **2011**, *3*(1), 33, DOI 10.1186/1758-2946-3-33.

[113] The Open Babel Package, version 2.3.1, http://openbabel.org, **2013**.

[114] E. Jones, T. Oliphant, P. Peterson, Others, SciPy: Open Source Scientific Tools for Python, http://www.scipy.org, **2001**.

[115] SciPy: Open Source Scientific Tools for Python, http://www.scipy.org/, **2013**.

[116] NumPy: Open Numerical Python, http://numpy.scipy.org/, **2013**.

[117] M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Comput. J.* **1964**, *7*(2), 155–162, DOI 10.1093/comjnl/7.2.155.

[118] J. D. Hunter, Matplotlib: A 2D graphics environment, *Comput. Sci. Eng.* **2007**, *9*(3), 90–95, DOI 10.1109/MCSE.2007.55.

[119] R. Koradi, M. Billeter, K. Wüthrich, MOLMOL: A program for display and analysis of macromolecular structures, *J. Mol. Graph.* **1996**, *14*(1), 51 – 55, DOI 10.1016/0263-7855(96)00009-4.

[120] X. Liu, Z. Tong, O. Hu, Swelling equilibria of hydrogels with sulfonate groups in water and in aqueous salt solutions, *Macromolecules* **1995**, *28*(11), 3813–3817, DOI 10.1021/ma00115a010.

[121] X. Liu, Z. Tong, X. Cao, O. Hu, Volume phase transition of polyelectrolyte gels in dimethyl sulfoxide/tetrahydrofuran mixtures, *Polymer* **1996**, *37*(26), 5947–5949.

[122] R. W. Seidel, B. R. Bakalska, T. Kolev, D. Vassilev, H. Mayer-Figge, M. Spiteller, W. S. Sheldrick, B. B. Koleva, N-methylcodeinium iodide—Crystal structure and spectroscopic elucidation, *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2009**, *73*(1), 61–66, DOI 10.1016/j.saa.2009.01.017.

[123] T. A. Halgren, Merck molecular force field I: Basis. form, scope, parametrization and performance of MMFF94, *J. Comput. Chem.* **1996**, *17*, 490–519, DOI 10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P.

[124] J. Tomasi, B. Mennucci, R. Cammi, Quantum Mechanical Continuum Solvation Models, *Chem. Rev.* **2005**, *105*(8), 2999–3094, DOI 10.1021/cr9904009.

[125] A. L. Mackay, Quaternion transformation of molecular orientation, *Acta Cryst. A* **1984**, *40*(2), 165–166, DOI 10.1107/S0108767384000362.

[126] J. H. Prestegard, H. M. Al-Hashimi, J. R. Tolman, NMR structures of biomolecules using field oriented media and residual dipolar couplings., *Q. Rev. Biophys.* **2000**, *33*(4), 371–424, DOI 10.1017/S0033583500003656.

[127] M. Zweckstetter, A. Bax, Characterization of molecular alignment in aqueous suspensions of Pf1 bacteriophage, *J. Biomol. NMR* **2001**, *20*, 365–377, DOI 10.1023/A:1011263920003.

[128] J. Sass, F. Cordier, A. Hoffmann, M. Rogowski, A. Cousin, J. G. Omichinski, H. Löwen, S. Grzesiek, Purple Membrane Induced Alignment of Biological Macromolecules in the Magnetic Field, *J. Am. Chem. Soc.* **1999**, *121*(10), 2047–2055, DOI 10.1021/ja983887w.

[129] S. Meier, D. Häussinger, S. Grzesiek, Charged acrylamide copolymer gels as media for weak alignment, *J. Biomol. NMR* **2002**, *24*(4), 351–356, DOI 10.1023/a:1021609207024.

[130] I. Kolossváry, W. C. Guida, Low Mode Search. An Efficient, Automated Computational Method for Conformational Analysis: Application to Cyclic and Acyclic Alkanes and Cyclic Peptides, *J. Am. Chem. Soc.* **1996**, *118*(21), 5011–5019, DOI 10.1021/ja952478m.

[131] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Ds, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, , D. J. Fox, Gaussian 09, Revision A.1, Gaussian, Inc., Wallingford CT, **2009**.

[132] Y. Zhao, N. E. Schultz, D. G. Truhlar, Design of Density Functionals by Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions, *J. Chem. Theor. Comp.* **2006**, *2*(2), 364–382, DOI 10.1021/ct0502763.

[133] Proskurina, Orekhov, *Bull. Chim. Soc. Fr.* **1937**, *4*, 1265.

[134] J. W. Langston, P. Ballard, J. W. Tetrud, I. Irwin, Chronic Parkinsonism in humans due to a product of meperidine-analog synthesis., *Science* **1983**, *219*(4587), 979–980, DOI 10.1126/science.6823561.

[135] L. Antkiewicz-Michaluk, J. Wardas, J. Michaluk, I. Romaska, A. Bojarski, J. Vetulani, Protective effect of 1-methyl-1,2,3,4-tetrahydroisoquinoline against dopaminergic neurodegeneration in the extrapyramidal structures produced by intracerebral injection of rotenone., *Int. J. Neuropsychopharmacol.* **2004**, *7*(2), 155–163, DOI 10.1017/S1461145703004036.

[136] Y. Tasaki, Y. Makino, S. Ohta, M. Hirobe, 1-Methyl- 1,2,3,4-Tetrahydroisoquinoline, Decreasing in 1-Methyl-4-Phenyl-1,2,3,6-Tetrahydropyridine-Treated Mouse, Prevents Parkinsonism-Like Behavior Abnormalities, *J. Neurochem.* **1991**, *57*(6), 1940–1943, DOI 10.1111/j.1471-4159.1991.tb06407.x.

[137] T. Yamakawa, Y. Kotake, M. Fujitani, H. Shintani, Y. Makino, S. Ohta, Regional distribution of parkinsonism-preventing endogenous tetrahydroisoquinoline derivatives and an endogenous parkinsonism-preventing substance-synthesizing enzyme in monkey brain, *Neurosci. Lett.* **1999**, *276*(1), 68–70, DOI 10.1016/S0304-3940(99)00786-7.

[138] L. Antkiewicz-Michaluk, A. Wasik, A. Romańska, Iand Bojarski, J. Michaluk, Both Stereoselective (R)- and (S)-1-Methyl-1,2,3,4-tetrahydroisoquinoline Enantiomers Protect Striatal Terminals Against Rotenone-Induced Suppression of Dopamine Release, *Neurotox. Res.* **2011**, *20*, 134–149, DOI 10.1007/s12640-010-9228-5.

[139] K. Iwasa, M. Kamigauchi, K. Saiki, N. Takao, W. Wiegrebe, O-Methylating enzymes of dopamine and dopamine derived tetrahydroisoquinoline, salsolinol, *Phytochemistry* **1993**, *32*(6), 1443–1448, DOI 10.1016/0031-9422(93)85156-L.

[140] J. Adachi, Y. Mizoi, T. Fukunaga, M. Kogame, I. Ninomiya, T. Naito, Effect of acetaldehyde on urinary salsolinol in healthy man after ethanol intake., *Alcohol* **1986**, *3*(3), 215–220, DOI 10.1016/0741-8329(86)90047-9.

[141] T. Niwa, N. Takeda, H. Yoshizumi, A. Tatematsu, M. Yoshida, P. Dostert, M. Naoi, T. Nagatsu, Presence of tetrahydroisoquinoline-related compounds, possible MPTP-like neurotoxins, in parkinsonian brain., *Adv. Neurol.* **1993**, *60*, 234–237.

[142] M. Naoi, P. Dostert, M. Yoshida, T. Nagatsu, N-methylated tetrahydroisoquinolines as dopaminergic neurotoxins., *Adv. Neurol.* **1993**, *60*, 212–217.

[143] G. F. Oxenkrug, 1-methyl and 1-benzyl derivatives of tetrahydroisoquinoline inhibit 5-hydroxytryptamine uptake by human blood platelets., *Med. Biol.* **1978**, *56*(5), 286–287.

[144] K. Sanft, H. Z. Thomas, *Naturforsch* **1989**, (44), 173–175.

[145] R. T, A. O, J. G. van A, Role of acetaldehyde in tobacco smoke addiction, *Eur. Neuropsychopharmacol.* **2007**, *17*(10), 627 – 636, DOI 10.1016/j.euroneuro.2007.02.013.

[146] J. J. Luszczki, L. Antkiewicz-Michaluk, S. J. Czuczwar, 1-Methyl-1,2,3,4-tetrahydroisoquinoline enhances the anticonvulsant action of carbamazepine and valproate in the mouse maximal electroshock seizure model., *Neuropharmacology* **2006**, *50*(2), 133–142, DOI 10.1016/j.neuropharm.2005.07.016.

[147] M. Karimov, M. G. Levkovich, V. B. Leont'ev, A. S. Sadykov, K. A. Aslanov, T. K. Yunusov, A. A. Sadykov, Study of the conformational states of amino-alcohol derivatives of some alkaloids and bases and their protonated forms and quaternary salts, *Chem. Nat. Compd.* **1974**, *10*(4), 490–495, DOI 10.1007/BF00563816.

[148] P. S. Charifson, J. P. Bowen, S. D. Wyrick, A. J. Hoffman, M. Cory, A. T. McPhail, R. B. Mailman, Conformational analysis and molecular modeling of 1-phenyl-,

4-phenyl-, and 1-benzyl-1, 2, 3, 4-tetrahydroisoquinolines as D1 dopamine receptor ligands, *J. Med. Chem.* **1989**, *32*(9), 2050–2058, DOI 10.1021/jm00129a006.

[149] E. M. Olefirowicz, E. L. Eliel, Conformational Analysis. 50. C-Methyl-1, 2, 3, 4-tetrahydroisoquinolines, *J. Org. Chem.* **1997**, *62*(26), 9154–9158, DOI 10.1021/jo971257r.

[150] N. L. Allinger, Y. H. Yuh, J. H. Lii, Molecular mechanics. The MM3 force field for hydrocarbons. 1, *J. Am. Chem. Soc.* **1989**, *111*(23), 8551–8566, DOI 10.1021/ja00205a001.

[151] J. H. Lii, N. L. Allinger, Molecular mechanics. The MM3 force field for hydrocarbons. 3. The van der Waals' potentials and crystal data for aliphatic and aromatic hydrocarbons, *J. Am. Chem. Soc.* **1989**, *111*(23), 8576–8582, DOI 10.1021/ja00205a003.

[152] M. W. Wong, K. B. Wiberg, M. J. Frisch, Solvent effects. 2. Medium effect on the structure, energy, charge density, and vibrational frequencies of sulfamic acid, *J. Am. Chem. Soc.* **1992**, *114*(2), 523–529, DOI 10.1021/ja00028a019.

[153] Y. Zhang, A. N. Wu, X. Xu, Y. J. Yan, OPBE: A promising densitiy functional for the calculation of nuclear shielding constants, *Chem. Phys. Lett.* **2006**, *421*, 383–388, DOI 10.1016/j.cplett.2006.01.095.

[154] F. Jensen, Basis Set Convergence of Nuclear Magnetic Shielding Constants Calculated by Density Functional Methods, *J. Chem. Theory Comput.* **2008**, *4*(5), 719–727, DOI 10.1021/ct800013z.

[155] SpinWorks 3.1.8 by K Marat, http://www.columbia.edu/cu/chemistry/groups/nmr/SpinWorks.html, **2012**.

[156] A. R. Quirt, J. S. Martin, NMR spectra of symmetric molecules. I. The spin Hamiltonian for twofold symmetry, *J. Magn. Reson.* **1971**, *5*(3), 318–327, DOI 10.1016/0022-2364(71)90082-5.

[157] R. S. Prosser, J. A. Losonczi, I. V. Shiyanovskaya, Use of a Novel Aqueous Liquid Crystalline Medium for High-Resolution NMR of Macromolecules in Solution, *J. Am. Chem. Soc.* **1998**, *120*(42), 11010–11011, DOI 10.1021/ja982671r.

[158] L. Skora, M.-K. Cho, H.-Y. Kim, S. Becker, C. O. Fernandez, M. Blackledge, M. Zweckstetter, Charge-Induced Molecular Alignment of Intrinsically Disordered Proteins, *Angew. Chem. Int. Ed. Engl.* **2006**, *45*(42), 7012–7015, DOI 10.1002/anie.200602317.

[159] C. Landersjö, C. Höög, A. Maliniak, G. Widmalm, NMR Investigation of a Tetrasaccharide Using Residual Dipolar Couplings in Dilute Liquid Crystalline Media – Effect of the Environment, *J. Phys. Chem. B* **2000**, *104*(23), 5618–5624, DOI 10.1021/jp000061j.

[160] M. Martín-Pastor, C. A. Bush, Refined structure of a flexible heptasaccharide using ^{13}C–^{1}H and ^{1}H–^{1}H NMR residual dipolar couplings in concert with NOE and long range scalar coupling constants, *J. Biomol. NMR* **2001**, *19*, 125–139, DOI 10.1023/A:1008327926009.

[161] K. Lycknert, A. Maliniak, G. Widmalm, Analysis of Oligosaccharide Conformation by NMR Spectroscopy Utilizing 1H,1H and 1H,13C Residual Dipolar Couplings in a Dilute Liquid Crystalline Phase, *J. Phys. Chem. A* **2001**, *105*(21), 5119–5122, DOI 10.1021/jp0103068.

[162] K. Ruan, K. Briggman, J. Tolman, De novo determination of internuclear vector orientations from residual dipolar couplings measured in three independent alignment media, *J. Biomol. NMR* **2008**, *41*, 61–76.

[163] V. Wong, D. A. Case, A. Szabo, Influence of the coupling of interdomain and overall motions on NMR relaxation, *Proc. Natl. Acad. Sci. USA* **2009**, *106*(27), 11016–11021, DOI 10.1073/pnas.0809994106.

[164] K. Wolinski, J. F. Hinton, P. Pulay, Efficient implementation of the gauge-independent atomic orbital method for NMR chemical shift calculations, *J. Am. Chem. Soc.* **1990**, *112*(23), 8251–8260, DOI 10.1021/ja00179a005.

[165] A. J. Cohen, N. C. Handy, Assessment of exchange correlation functionals, *Chem. Phys. Lett.* **2000**, *316*(1–2), 160 – 166, DOI 10.1016/S0009-2614(99)01273-7.

[166] J. P. Perdew, K. Burke, M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868, DOI 10.1103/PhysRevLett.77.3865.

[167] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E.

Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Ds, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 03, Revision E.01, Gaussian, Inc., Wallingford CT, **2004**.

[168] J. Men, W. I. Taylor, A uniform numbering system for indole alkaloids, *Cell. Mol. Life Sci.* **1965**, *21*(9), 508–510, DOI 10.1007/BF0213896.

[169] T. Adachi, M. Saito, J. Sasaki, Y. Karasawa, H. Araki, K. Hanada, S. Omura, Microbial hydroxylation of (-)-eburnamonine by *Mucor circinelloides* and *Streptomyces violens*., *Chem. Pharm. Bull.* **1993**, *41*(3), 611–613, DOI 10.1248/cpb.41.611.

[170] L. Allorge, New Species and Combinations from American *Apocynaceae-Tabernaemontanoideae*, *Lett. Botaniq.* **1983**, *130*, 339–35.

[171] J. Garnier, G. Croquelois, P. Kaminski, G. Lewin, C. Miet, J. Poisson, C. Moretti, Alkaloids from *Bonafusia macrocalyx*, *Plantes Medicinales et Phytotherapie* **1984**, *18*, 144–149.

[172] J. Garnier, G. Croquelois, P. Kaminski, G. Lewin, C. Miet, J. Poisson, C. Moretti, Alkaloids from *Bonafusia macrocalyx*, *J. Nat. Prod.* **1984**, *47*(6), 1055–1056, DOI 10.1021/np50036a035.

[173] J. Garnier, J. Mahuteau, Bonafousioside, a new glucolignan isolated from *Bonafousia macrocalyx*, *J. Nat. Prod.* **1988**, DOI 10.1021/np50057a003.

[174] Á. Vas, B. Gulyás, Eburnamine derivatives and the brain, *Med. Res. Rev.* **2005**, *25*(6), 737–757, DOI 10.1002/med.20043.

[175] C. Y. Gan, Y. Y. Low, T. Etoh, M. Hayashi, K. Komiyama, T. S. Kam, Leuconicines, (+) and (−)-Eburnamaline, Biologically Active Strychnan and Eburnan Alkaloids from *Leuconotis*, *J. Nat. Prod.* **2009**, *72*(12), 2098–2103, DOI 10.1021/np900576b.

[176] A. G. Petrovic, A. Navarro-Vázquez, J. L. Alonso-Gómez, From relative to absolute configuration of complex natural products. Interplay between NMR, ECD, VCD and ORD assisted by ab initio calculations, *Curr. Org. Chem.* **2010**, *14*, 1612–1628.

[177] M. W. Lodewyk, M. R. Siebert, D. J. Tantillo, Computational Prediction of 1H and 13C Chemical Shifts: A Useful Tool for Natural Product, Mechanistic, and Synthetic Organic Chemistry, *Chem. Rev.* **2012**, *112*(3), 1839–1862, DOI 10.1021/cr200106v.

[178] G. Bifulco, P. Dambruoso, L. Gomez-Paloma, R. Riccio, Determination of relative configuration in organic compounds by NMR spectroscopy and computational methods., *Chem. Rev.* **2007**, *107*(9), 3744–3779, DOI 10.1021/cr030733c.

[179] R. Ditchfield, Self-consistent perturbation theory of diamagnetism, *Mol. Phys.* **1974**, *27*(4), 789–807, DOI 10.1080/00268977400100711.

[180] G. Barone, D. Duca, A. S, L. Gomez-Paloma, R. Riccio, G. Bifulco, Determination of the Relative Stereochemistry of Flexible Organic Compounds by Ab Initio Methods: Conformational Analysis and Boltzmann-Averaged GIAO 13C NMR Chemical Shifts, *Chem. Eur. J.* **2002**, *8*(14), 3240–3245, DOI 10.1002/1521-3765(20020715)8:14<3240::AID-CHEM3240>3.0.CO;2-G.

[181] G. Bifulco, L. Gomez-Paloma, R. Riccio, Configurational analysis of the natural product passifloricin A by quantum mechanical ^{13}C NMR GIAO chemical shift calculations, *tetrahedron lett.* **2003**, *44*(38), 7137–7141, DOI 10.1016/S0040-4039(03)01810-0.

[182] P. Trigo-Mouriño, A. Navarro-Vázquez, J. Ying, R. R. Gil, A. Bax, Structural discrimination in small molecules by accurate measurement of long-range proton-carbon NMR residual dipolar couplings., *Angew. Chem. Int. Ed. Engl.* **2011**, *50*(33), 7576–7580, DOI 10.1002/anie.201101739.

[183] R. R. Gil, Constitutional, configurational, and conformational analysis of small organic molecules on the basis of NMR residual dipolar couplings, *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 7222–7224, DOI 10.1002/anie.201101561.

[184] S. G. Smith, J. M. Goodman, Assigning stereochemistry to single diastereoisomers by GIAO NMR calculation: the DP4 probability., *J. Am. Chem. Soc.* **2010**, *132*(37), 12946–12959, DOI 10.1021/ja105035r.

[185] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields, *J. Phys. Chem.* **1994**, *98*(45), 11623–11627, DOI 10.1021/j100096a001.

[186] J. M. Goodman, DP4 JAVA web applet, http://www.jmg.ch.cam.ac.uk/tools/nmr/DP4, **2013**.

[187] U. M. Reinscheid, M. Köck, C. Cychon, V. Schmidts, C. M. Thiele, C. Griesinger, The Absolute Configuration of Dibromopalau'amine, *Eur. J. Org. Chem.* **2010**, *2010*(36), 6900–6903, DOI 10.1002/ejoc.201001392.

[188] C. Pérez-Balado, H. Sun, C. Griesinger, A. R. de Lera, A. Navarro-Vázquez, Residual Dipolar Coupling Enhanced NMR Spectroscopy and Chiroptics: A Powerful Combination for the Complete Elucidation of Symmetrical Small Molecules, *Chem. Eur. J.* **2011**, *17*(43), 11983–11986, DOI 10.1002/chem.201101385.

[189] C. Adamo, V. Barone, Toward reliable density functional methods without adjustable parameters: The PBE0 model, *J. Chem. Phys.* **1999**, *110*, 6158–6170, DOI 10.1063/1.478522.

[190] N. Matsumori, M. Murata, K. Tachibana, Conformational analysis of natural products using long-range carbon-proton coupling constants: Three-dimensional structure of okadaic acid in solution, *Tetrahedron* **1995**, *51*(45), 12229–12238, DOI 10.1016/0040-4020(95)00790-F.

[191] M. Murata, S. Matsuoka, N. Matsumori, G. K. Paul, K. Tachibana, Absolute Configuration of Amphidinol 3, the First Complete Structure Determination from Amphidinol Homologues: Application of a New Configuration Analysis Based on Carbon–Hydrogen Spin-Coupling Constants, *J. Am. Chem. Soc.* **1999**, *121*(4), 870–871, DOI 10.1021/ja983655x.

[192] T. Rundlöf, A. Kjellberg, C. Damberg, T. Nishida, G. Widmalm, Long-range proton–carbon coupling constants in conformational analysis of oligosaccharides, *Magn. Reson. Chem.* **1998**, *36*(11), 839–847, DOI 10.1002/(SICI)1097-458X(1998110)36:11<839::AID-OMR371>3.0.CO;2-9.

[193] S. X. Huang, X. J. Wang, Y. Yan, J. D. Wang, J. Zhang, C. X. Liu, W. S. Xiang, B. Shen, Neaumycin: A New Macrolide from Streptomycessp. NEAU-x211, *Org. Lett.* **2012**, *14*(5), 1254–1257, DOI 10.1021/ol300074d.

[194] A. Gutiérrez-Cepeda, J. J. Fernández, L. V. Gil, M. López-Rodríguez, M. Norte, M. L. Souto, Nonterpenoid C 15Acetogenins from Laurencia marilzae, *J. Nat. Prod.* **2011**, *74*(3), 441–448, DOI 10.1021/np100866g.

[195] J. Peng, A. R. Place, W. Yoshida, C. Anklin, M. T. Hamann, Structure and Absolute Configuration of Karlotoxin-2, an Ichthyotoxin from the Marine *Dinoflagellate Karlodinium veneficum*, *J. Am. Chem. Soc.* **2010**, *132*(10), 3277–3279, DOI 10.1021/ja9091853.

[196] A. Plaza, R. Garcia, G. Bifulco, J. P. Martinez, S. Hüttel, F. Sasse, A. Meyerhans, M. Stadler, R. Müller, Aetheramides A and B, Potent HIV-Inhibitory Depsipeptides from a Myxobacterium of the New Genus "*Aetherobacter*", *Org. Lett.* **2012**, *14*(11), 2854–2857, DOI 10.1021/ol3011002.

[197] J. J. Poza, C. Jiménez, J. Rodríguez, *J*-Based Analysis and DFT-NMR Assignments of Natural Complex Molecules: Application to 3β,7-Dihydroxy-5,6-epoxycholestanes, *Eur. J. Org. Chem.* **2008**, *2008*(23), 3960–3969, DOI 10.1002/ejoc.200800358.

[198] B. L. Marquez, W. H. Gerwick, R. Thomas Williamson, Survey of NMR experiments for the determination of $^{n}J_{CH}$ heteronuclear coupling constants in small molecules, *Magn. Reson. Chem.* **2001**, *39*(9), 499–530, DOI 10.1002/mrc.902.

[199] J. J. Titman, D. Neuhaus, J. Keeler, Measurement of long-range heteronuclear coupling constants, *J.Magn. Reson.* **1989**, *85*(1), 111–131.

[200] A. Bax, R. Freeman, Long-Range Proton C-13 NMR Spin Coupling-Constants, *J. Am. Chem. Soc.* **1982**, *104*(4), 1099–1100, DOI 10.1021/ja00368a033.

[201] L. Kay, P. Keifer, T. Saarinen, Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, *J. Am. Chem. Soc.* **1992**, *114*(26), 10663–10665, DOI 10.1021/ja00052a088.

[202] A. Bax, S. Subramanian, Sensitivity-Enhanced Two-Dimensional Heteronuclear Shift Correlation NMR Spectroscopy, *J. Magn. Reson.* **1986**, *67*, 565, DOI 10.1021/ja905991s.

[203] J. Schleucher, M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky, S. J. Glaser, O. W. So, C. Griesinger, A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients, *J. Biomol. NMR* **1994**, *4*(2), 301–306, DOI 10.1007/BF00175254.

[204] M. H. Levitt, R. Freeman, Compensation for pulse imperfections in NMR spin-echo experiments, *J. Magn. Reson.* **1981**, *43*(1), 65–80, DOI 10.1016/0022-2364(81)90082-2.

[205] H. Schwalbe, P. Schmieder, C. Griesinger, *Coupling Constants Determined by E.COSY*, Wiley, **2007**, DOI 0.1002/9780470034590.emrstm0098.

[206] V. E. Sosa, J. C. Oberti, R. R. Gil, E. A. Rúveda, V. L. Goedken, W. Herz, 10-Epideoxycumambrin b and other constituents of *Stevia yaconensis* var. subeglandulosa, *Phytochemistry* **2001**, *28*(7), 1925–1929, DOI 10.1016/S0031-9422(00)97888-9.

[207] J. G. Blanco, R. R. Gil, C. I. Álvarez, L. C. Patrito, S. Genti-Raimondi, A. Flury, A novel activity for a group of sesquiterpene lactones: inhibition of aromatase, *FEBS Lett.* **1997**, *409*(3), 396–400, DOI 0.1016/S0014-5793(97)00560-7.

[208] V. E. Sosa, J. C. Oberti, R. R. Gil, E. A. Rúveda, V. L. Goedken, A. B. Gutiérrez, W. Herz, 10-Epideoxycumambrin b and other constituents of *Stevia yaconensis* var. subeglandulosa, *Phytochemistry* **1989**, *28*(7), 1925 – 1929, DOI 10.1016/S0031-9422(00)97888-9.

[209] P. M. Dewick, The biosynthesis of C5–C25 terpenoid compounds, *Nat. Prod. Rep.* **2002**, *19*, 181–222.

[210] K. E. Kövér, K. Fehér, Measurement of one-bond heteronuclear dipolar coupling contributions for amine and diastereotopic methylene protons, *J.Magn. Reson.* **2004**, *168*(2), 307–313, DOI 10.1016/j.jmr.2004.03.007.

[211] J. L. Marshall, *Methods in stereochemical analysis*, Weinheim, **1983**.

[212] TopSpin 3.2, Bruker Corporation http://www.bruker.com/products/magnetic-resonance/nmr/sw/nmr-software/topspin/overview.html, **2013**.

[213] J. Weinstock, J. P. Hieble, J. W. Wilson III, The chemistry and pharmacology of 3-benzazepine derivatives, *Drugs Future* **1985**, *10*(8), 645–697.

[214] P. A. Donets, E. V. Van der Eycken, Efficient Synthesis of the 3-Benzazepine Framework via Intramolecular Heck Reductive Cyclization, *Org. Lett.* **2007**, *9*(16), 3017–3020, DOI 10.1021/ol071079g.

[215] C. Wang, Y. Jiang, J. Ma, H. Wu, D. Wacker, V. Katritch, G. W. Han, W. Liu, X.-P. Huang, E. Vardy, J. D. McCorvy, X. Gao, E. X. Zhou, K. Melcher, C. Zhang, F. Bai, H. Yang, L. Yang, H. Jiang, B. L. Roth, V. Cherezov, R. C. Stevens, H. E. Xu, Structural Basis for Molecular Recognition at Serotonin Receptors, *Science* **2013**, DOI 10.1126/science.1232807.

[216] P. V. Fish, A. D. Brown, E. Evrard, L. R. Roberts, 7-Sulfonamido-3-benzazepines as potent and selective 5-HT_{2C} receptor agonists: hit-to-lead optimization., *Bioorg. Med. Chem. Lett.* **2009**, *19*(7), 1871–1875, DOI 10.1016/j.bmcl.2009.02.071.

[217] C. Kaiser, P. A. Dandridge, E. Garvey, R. A. Hahn, H. M. Sarau, P. E. Setler, L. S. Bass, J. Clardy, Absolute stereochemistry and dopaminergic activity of enantiomers of 2, 3, 4, 5-tetrahydro-7, 8-dihydroxy-1-phenyl-1H-3-benzazepine, *J. Med. Chem.* **1982**, *25*(6), 697–703, DOI 10.1021/jm00348a017.

[218] I. Pettersson, T. Liljefors, K. Bogeso, Conformational analysis and structure-activity relationships of selective dopamine D-1 receptor agonists and antagonists of the benzazepine series, *J. Med. Chem.* **1990**, *33*(8), 2197–2204, DOI 10.1021/jm00170a025.

[219] J. G. Berger, W. K. Chang, J. W. Clader, D. Hou, R. E. Chipkin, A. T. McPhail, Synthesis and receptor affinities of some conformationally restricted analogues of the dopamine D1 selective ligand (5R)-8-chloro-2,3,4,5-tetrahydro-3-methyl-5-phenyl- 1H-3-benzazepin-7-ol., *J. Med. Chem.* **1989**, *32*(8), 1913–1921, DOI 10.1021/jm00128a038.

[220] S. E. Snyder, F. A. Aviles-Garay, R. Chakraborti, D. E. Nichols, V. J. Watts, R. B. Mailman, Synthesis and Evaluation of 6, 7-Dihydroxy-2, 3, 4, 8, 9, 13b-hexahydro-1H-benzo [6, 7] cyclohepta [1, 2, 3-ef][3] benzazepine, 6, 7-Dihydroxy-1, 2, 3, 4, 8, 12b-hexahydroanthr [10, 4a, 4-cd] azepine, and 10-(Aminomethyl)-9, 10-dihydro-1, 2-dihydroxyanthracene as Conformationally Restricted Analogs of. beta.-Phenyldopamine, *J. Med. Chem.* **1995**, *38*(13), 2395–2409, DOI 10.1021/jm00013a015.

[221] I. Alkorta, H. O. Villar, R. E. Cachau, Conformational analysis of 2, 3, 6, 7-tetrahydroazepines with implications for D1-selective benzazepines, *J Comp. Chem.* **2004**, *14*(5), 571–578, DOI 10.1002/jcc.540140510.

[222] H. E. Bays, Lorcaserin and adiposopathy: 5-HT_{2C} agonism as a treatment for 'sick fat' and metabolic disease., *Expert Rev. Cardiovasc. Ther.* **2009**, *7*(11), 1429–1445, DOI 10.1586/erc.09.123.

[223] K. M. Flegal, M. D. Carroll, C. L. Ogden, L. R. Curtin, Prevalence and trends in obesity among us adults, 1999-2008, *JAMA: J. Am. Med. Assoc.* **2010**, *303*(3), 235–241, DOI 10.1001/jama.2009.2014.

[224] NHLBI, Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults–The Evidence Report. National Institutes of Health., *Obes. Res.* **1998**, *6 Suppl 2*, 51S–209S.

[225] G. A. Bray, In defense of a body mass index of 25 as the cut-off point for defining overweight., *Obes. Res.* **1998**, *6*(6), 461–462, DOI 10.1002/j.1550-8528.1998.tb00379.x.

[226] B. M. Smith, J. M. Smith, J. H. Tsai, J. A. Schultz, C. A. Gilson, S. A. Estrada, R. R. Chen, D. M. Park, E. B. Prieto, C. S. Gallardo, D. Sengupta, W. J. Thomsen, H. R. Saldana, K. T. Whelan, F. Menzaghi, R. R. Webb, N. R. A. Beeley, Discovery and SAR of new benzazepines as potent and selective 5-HT(2C) receptor agonists for the treatment of obesity., *Bioorg. Med. Chem. Lett.* **2005**, *15*(5), 1467–1470, DOI 10.1016/j.bmcl.2004.12.080.

[227] B. M. Smith, J. M. Smith, J. H. Tsai, J. A. Schultz, C. A. Gilson, S. A. Estrada, R. R. Chen, D. M. Park, E. B. Prieto, C. S. Gallardo, D. Sengupta, P. I. Dosa, J. A. Covel,

A. Ren, R. R. Webb, N. R. A. Beeley, M. Martin, M. Morgan, S. Espitia, H. R. Saldana, C. Bjenning, K. T. Whelan, A. J. Grottick, F. Menzaghi, W. J. Thomsen, Discovery and structure-activity relationship of (1R)-8-chloro-2,3,4,5-tetrahydro-1-methyl-1H-3-benzazepine (Lorcaserin), a selective serotonin 5-HT2C receptor agonist for the treatment of obesity., *J. Med. Chem.* **2008**, *51*(2), 305–313, DOI 10.1021/jm0709034.

[228] W. J. Thomsen, A. J. Grottick, F. Menzaghi, H. Reyes-Saldana, S. Espitia, D. Yuskin, K. Whelan, M. Martin, M. Morgan, W. Chen, H. Al-Shamma, B. Smith, D. Chalmers, D. Behan, Lorcaserin, a novel selective human 5-Hydroxytryptamine$_{2C}$ agonist: in vitro and in vivo pharmacological characterization., *J. Pharmacol. Exp. Ther.* **2008**, *325*(2), 577–587, DOI 10.1124/jpet.107.133348.

[229] S. R. Smith, N. J. Weissman, C. M. Anderson, M. Sanchez, E. Chuang, S. Stubbe, H. Bays, W. R. Shanahan, Behavioral Modification and Lorcaserin for Overweight and Obesity Management (BLOOM) Study Group, Multicenter, placebo-controlled trial of lorcaserin for weight management., *N. Engl. J. Med.* **2010**, *363*(3), 245–256, DOI 10.1056/NEJMoa0909809.

[230] A. G. Powell, C. M. Apovian, L. J. Aronne, New drug targets for the treatment of obesity., *Clin. Pharmacol. Ther.* **2011**, *90*(1), 40–51, DOI 10.1038/clpt.2011.82.

[231] Lorcaserin Information, Arena Pharmaceuticals, http://www.arenapharm.com/lorcaserin.aspx, **2013**.

[232] Lorcaserin FDA Aproval Press Release, Arena Pharmaceuticals and Eisai Pharmaceuticals, http://invest.arenapharm.com/releasedetail.cfm?ReleaseID=687182, **2013**.

[233] P. A. Sargent, A. L. Sharpley, C. Williams, E. M. Goodall, P. J. Cowen, 5-HT2C receptor activation decreases appetite and body weight in obese subjects., *Psychopharmacology* **1997**, *133*(3), 309–312, DOI 10.1007/s002130050407.

[234] S. P. Vickers, C. T. Dourish, G. A. Kennett, Evidence that hypophagia induced by d-fenfluramine and d-norfenfluramine in the rat is mediated by 5-HT2C receptors., *Neuropharmacology* **2001**, *41*(2), 200–209, DOI 10.1016/S0028-3908(01)00063-6.

[235] S. P. Vickers, N. Easton, L. J. Webster, A. Wyatt, M. J. Bickerdike, C. T. Dourish, G. A. Kennett, Oral administration of the 5-HT2Creceptor agonist, mCPP, reduces body weight gain in rats over 28 days as a result of maintained hypophagia., *Psychopharmacology* **2003**, *167*(3), 274–280, DOI 10.1007/s00213-002-1378-6.

[236] T. Carlomagno, NMR in natural products: understanding conformation, configuration and receptor interactions., *Nat. Prod. Rep.* **2012**, *29*(5), 536–554, DOI 10.1039/c2np00098a.

[237] E. P. Mazzola, A. Parkinson, E. J. Kennelly, B. Coxon, L. S. Einbond, D. I. Freedberg, Utility of coupled-HSQC experiments in the intact structural elucidation of three complex saponins from *Blighia sapida*, *Carbohydr. Res.* **2011**, *346*(6), 759 – 768, DOI 10.1016/j.carres.2011.02.019.

[238] F. Jensen, The Basis Set Convergence of Spin-Spin Coupling Constants Calculated by Density Functional Methods, *J. Chem. Theor. Comp.* **2006**, *2*(5), 1360–1369, DOI 10.1021/ct600166u.

[239] H. Akaike, A new look at the statistical model identification, *IEEE Trans. Autom. Control* **1974**, *19*(6), 716–723, DOI 10.1109/TAC.1974.1100705.

[240] M. Erdélyi, E. d'Auvergne, A. Navarro-Vázquez, A. Leonov, C. Griesinger, Dynamics of the Glycosidic Bond: Conformational Space of Lactose, *Chem. Eur. J.* **2011**, *17*(34), 9368–9376, DOI 10.1002/chem.201100854.

[241] M. J. Thrippleton, J. Keeler, Elimination of Zero-Quantum Interference in Two-Dimensional NMR Spectra, *Angew. Chem. Int. Ed. Engl.* **2003**, *42*(33), 3938–3941, DOI 10.1002/anie.200351947.