KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

Politechnika Wrocławska

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

ROZWÓJ POTENCJAŁU I OFERTY DYDAKTYCZNEJ POLITECHNIKI WROCŁAWSKIEJ

Wrocław University of Technology

Medicinal Chemistry

Michał Jewgiński, Rafał Latajka, Roman Gancarz

SPECTROSCOPIC METHODS IN MEDICINAL CHEMISTRY

Wrocław 2011

Projekt współfinansowany ze środków Unii Europejskiej w ramach
Europejskiego Funduszu Społecznego

Wrocław University of Technology

Medicinal Chemistry

Michał Jewgiński, Rafał Latajka, Roman Gancarz

SPECTROSCOPIC METHODS IN MEDICINAL CHEMISTRY

Wrocław 2011

Reviewer: Paweł Kafarski

ISBN 978-83-62098-39-2

Published by PRINTPAP Łódź, www.printpap.pl

Table of contents:

1. NMR SPECTROSCOPY 5
1.1 Chemical shift 6
1.2 Integration 9
1.3 Spin spin coupling constant 10
1.4 Nuclei equivalence- nonequivalence 13
Chemical equivalence 13
Magnetic equivalence 15
1.5 Spin system labeling 15
1.6 NMR one dimensional techniques 16
Homodecoupling 16
Broadband decoupling 17
DEPT (Distortionless Enhancement by Polarisation Transfer) 17
1.7 Two dimensional methods in NMR 18
HHCOSY 19
TOCSY (Total correlation spectroscopy) 21
HMQC 22
NOESY 23
Principal correlations established by NMR 24
1.8 Other methods 24
1D methods 24
2D methods 24
1.9 Other problems for further individual studies 25
1.10 Examples of NMR techniques in hydrogen bonding and dynamic processes studies 25
Coupling through hydrogen bond 25
Spin saturation transfer 26
2. Mass spectrometry 29
2.1 Introduction 29
2.2 Ionization methods 31
2.3 Analyzers 31
2.4 Principles of fragmentation process 33
2.5 Methods of ionization of macromolecules 34

2.6	Methodology employed in investigations of peptides and proteins	38
3.	Circular dichroism spectroscopy	41
4.	Computational techniques in spectroscopic methods useful in medicinal chemistry	47
4.1	Theoretical simulation of the spectra's of real compounds.	47
4.1.1	Theoretical calculation of the IR spectra.	49
4.1.2	Theoretical calculation of the NMR spectra.	55
4.1.3	Theoretical calculation of the VCD spectra.	69
4.2	Application of spectroscopic data to structural and conformational calculations	70

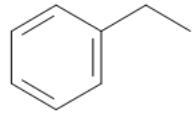
1. NMR SPECTROSCOPY

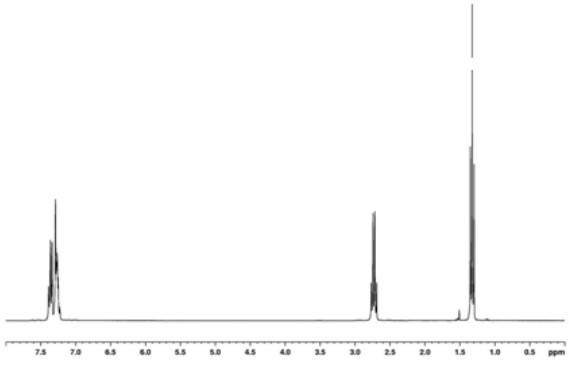
Since the first experiment in 1940 in the laboratory of physicists I. Rabi, F. Bloch and E. M. Purcell), NMR become one of the most powerful and most general structural tools in chemistry.

The intension of this study guide is to provide the basic information on NMR without going into theoretical background. The reader is suggested to expand his knowledge by studying the recommended books.

Every spectrum like the one proton spectrum shown in the figure below may be characterized by some major parameters: chemical shift, coupling constants and relaxation times T1 and T2. The first two are directly observed on the spectrum shown below while the evaluation of the other two needs some special attention.

This chapter will focus on the description of those four parameters.

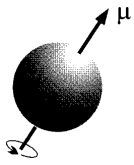




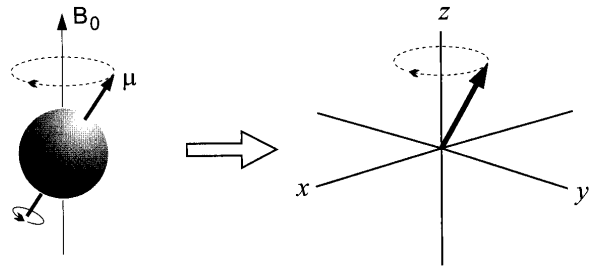
Aromatic **CH_2** **CH_3**

1.1 Chemical shift

The nuclei of all atoms are characterised by nuclear spin quantum number I. When I=0, the nuclei has no nuclear spin and is NMR silent. Otherwise it possesses the nuclear spin and is observed by NMR. The spinning nuclei possesses then angular magnetic moment μ.



When the nuclei with I≠0 is placed in an external magnetic field, the microscopic magnetic moments align themselves in the field in a defined discrete orientation due to the fact that energy states are quantized. For I=1/2 we have two states, (+1/2 and -1/2) for I=1 three states (-1, 0, 1) and so on. The effect of the external magnetic field on the magnetic moment is such that its torgue is imposed and the magnetic field precesses about the direction of the external field B_o with the defined velocity called Larmor precession ν_o.



$$\omega_0 = \gamma * B_0$$

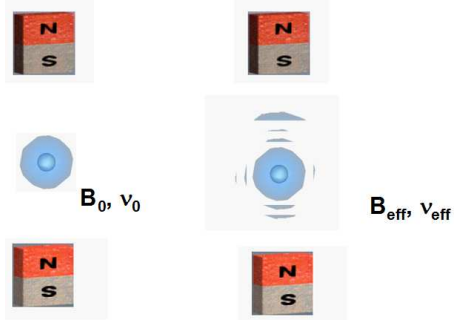
When the electromagnetic radiation with the frequency that matches the Larmor frequency is applied, nuclear magnetic resonance occurs and the nucleus changes its spin state. For the resonance to be satisfied the energy involved is given by the equation

$$\Delta E = h\nu = \frac{h\gamma B_0}{2\pi}$$

When nuclei is surrounded by the other nuclei, as in the case of each chemical molecule, the electronic and also magnetic environment about the nucleus is changed to the effective magnetic environment - B_{eff}. The effect is measured by the screening or shielding constant σ.Thus;

$$B_{eff}=B_o(1-\sigma)$$

and the resonance is then satisfied at ν_{eff} .



The difference $\Delta(\nu_o-\nu_{eff})$ is called a chemical shift. A chemical shift measured in Hz is proportional to the strength of the applied magnetic field. When calculated as:

$$\delta = \frac{\text{difference between a resonance frequency and that of a reference substance}}{\text{frequency of the exciting radiation}}$$

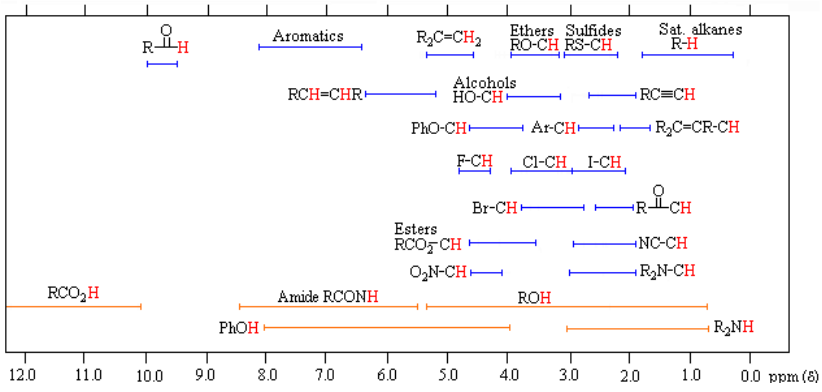
the chemical shift is independent upon applied magnetic field and falls within the range 0-15 (1H), 1-250 (13C).

Due to scaling factor 10^6 the units are given name ppm (parts per million).

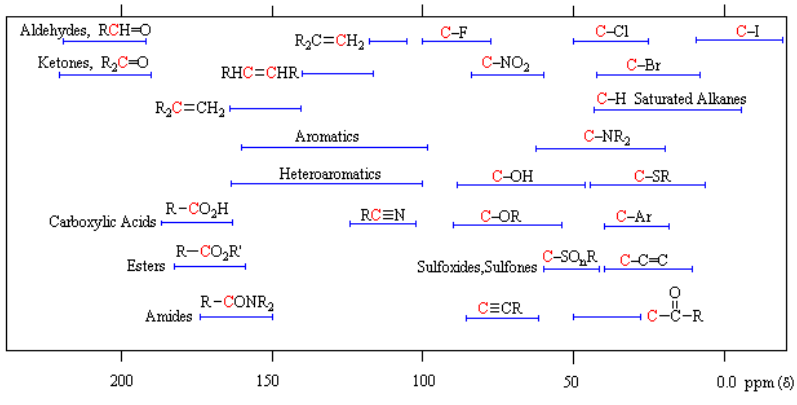
Symmetrically equivalent nuclei i.e.homotopic in all cases and enantiotopic (these terms will be defined later) nuclei in an achiral environment have the same chemical shift. Diasterotopic and heterotopic nuclei are characterized by different chemical shifts.

The chemical shift value yields the information about the environment of the resonating nucleus and thus provides the information about the chemical structure. The typical chemical shifts in ^1H and ^{13}C NMR are shown in the diagrams below.

Proton ^{1}H chemical shift ranges

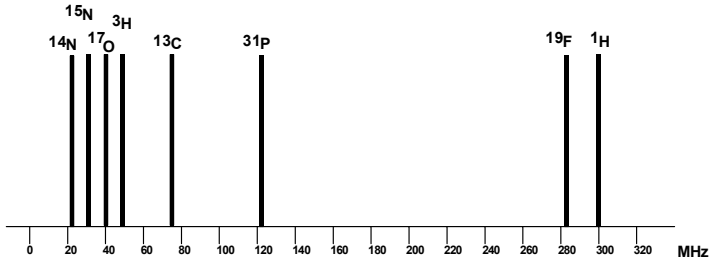


Carbon ^{13}C chemical shift ranges



http://www.cem.msu.edu/

The resonance frequency ranges for the particular nuclei are given in the diagram below. The ranges for the particular nuclei are very narrow with the respect of their position in the whole frequency scale.

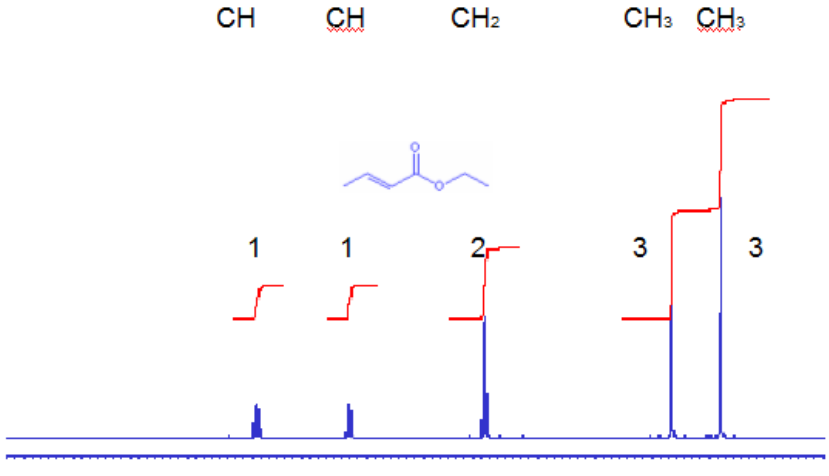


Resonance frequencies for main isotopes characteristic for 300 MHz instrument

The conclusion from the diagram is such that when the frequency window is open for a given nucleus no signals from other nuclei will be observed.

1.2 Integration

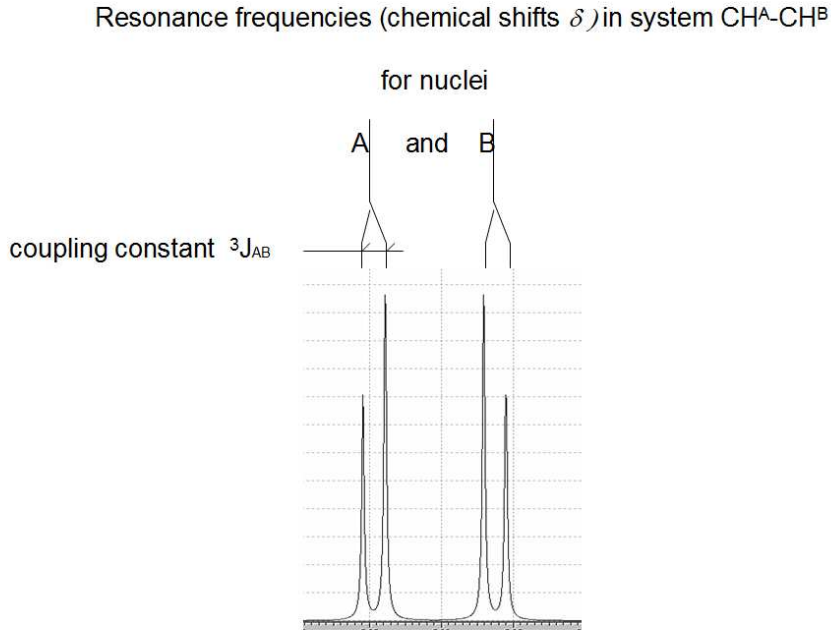
The useful in force doctrine states that the area of an NMR resonance signal proportional to the relative number of nuclei giving rise to that signal. In fact it has some limitations. Mainly it is accurate within 10% and appropriate acquisition parameters have to be provided (the time between consecutive impulses has to be at least 6-10 times longer than the longest relaxation time for the observed nuclei).



1.3 Spin spin coupling constant

Nuclei themselves are small magnets. Depending on the spin quantum number they exist in several states of different energy. As stated above, for spin quantum number I=1/2 there are two states +1/2 and -1/2 whereas for spin quantum number I=1 there are three states -1, 0, and 1. The effective magnetic field of the neighboring nuclei and hence the energy of the nearby nuclei are split in several states. As the consequence several resonance frequencies are observed. This is known as spin-spin coupling The most important is *scalar coupling* which occurs between two nuclei through chemical bonds, and can usually be seen up to three bonds away. The magnitude of the interaction between nuclei is measured as coupling constant and for example when two nuclei A and B are interfering, it is written as $^{n}J_{AB}$ where n indicates the number of bonds between the interacting nuclei.

For the system CH^{A}-CH^{B} of the two nonequivalent nuclei the observed effect is seen in the figure below.



The coupling constant is independent of the magnetic field strength therefore it is quoted in Hertz (frequency) contrary, to chemical shifts which are quoted in ppm.

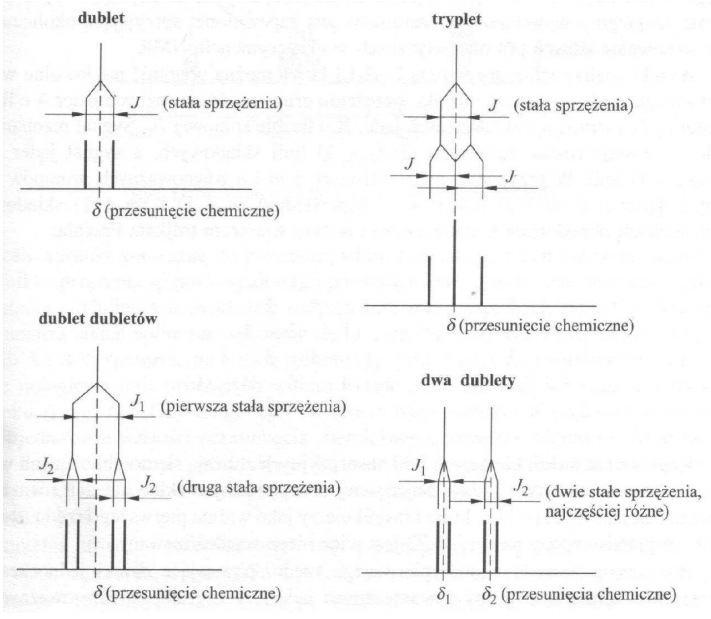
In the figure above every line is split in two frequencies due to to states +1/2 and -1/2 of the neighboring proton. The result is a doublet. In general the two lines are of different intensities and the resonance line lies in the center of gravity. In the ideal situation, namely when the difference in chemical shift is ten times greater than the coupling constant, the lines are of almost the same intensity.

If the observed nucleus or identical nuclei are interacting with more than one equivalent nucleus, the corresponding multiplets are observed (they can be easily calculated from Pascal triangle).

No. of neighboring nuclei	**multiplet**	**relative intensities of the multiplet lines**									
No nuclei	singlet					1					
One nucleus	dublet				1			1			
Two nuclei	triplet			1		2			1		
Three nuclei	quartet		1		3			3		1	
Four nuclei	quintet	1		4		6			4		1
.....											

If the interaction is with more than one but nonequivalent group of nuclei then more complex multiplet is observed. For example in the system CH_3-CH-CH_2 the resonance for CH will be a triplet of quartets as every line of a triplet resulting from the interaction with CH_2 is split in to quartets due to the interaction with CH_3. There are thus the total twelve lines (3 times 4).

For other most frequent cases the observed idealised spectra shapes are presented below.


dublet
J (stała sprzężenia)
δ (przesunięcie chemiczne)
tryplet
J (stała sprzężenia)
J
J
δ (przesunięcie chemiczne)
dublet dubletów
J1 (pierwsza stała sprzężenia)
J2
J2 (druga stała sprzężenia)
δ (przesunięcie chemiczne)
dwa dublety
J1
J2 (dwie stałe sprzężenia, najczęściej różne)
δ1 δ2 (przesunięcia chemiczne)

1.4 Nuclei equivalence- nonequivalence

In order to perform the discussion about the number of resonanses (signals) and number of coupling constants we have to define the terms **equivalence and nonequivalence** of nuclei. The chemical equivalence is necessary to predict the number of nuclei with the same chemical shift. Chemically equivalent nuclei are characterised with the same chemical shift. However, for the prediction of the number of coupling constants the analysis of magnetic equivalence is necessary as the coupling is observed between all magnetically not equivalent nuclei, even if they are chemically equivalent.

Chemical equivalence

Mislow and Raban (K. Mislow, M. Raban: W:*Topics in Stereochemistry*. Red. A. L. Allinger, E.L. Eliel. Wiley, New York **1967**, t.1, s. 1-38) divided the relation between atoms or a group of atoms in the molecule into four categories, which allows to predict satisfactory the number of NMR signals (chemically equivalent or not equivalent nuclei)

-homotopic

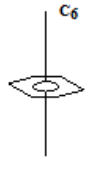
-enantiotopic

-diastereotopic

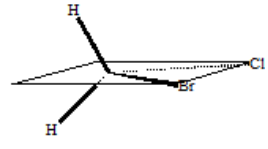
-heterotopic

The categorization relies on two features: symmetry and topology.

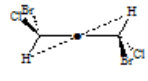
Symmetry criteria are based on basic isometric transformations and a combination of them.



symetry axis



symetry plane

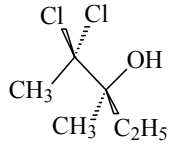


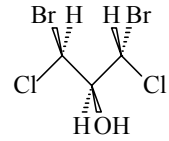
inversion point

Atoms or a group of atoms are homotopic if they are related by a symmetry axis (like all hydrogen atoms in benzene) or enantiotopic if they are related by a symmetry plane or inversion point (like hydrogen atoms in chlorobromometane, or 1,2-dichloro-1,2 dibromoetane).

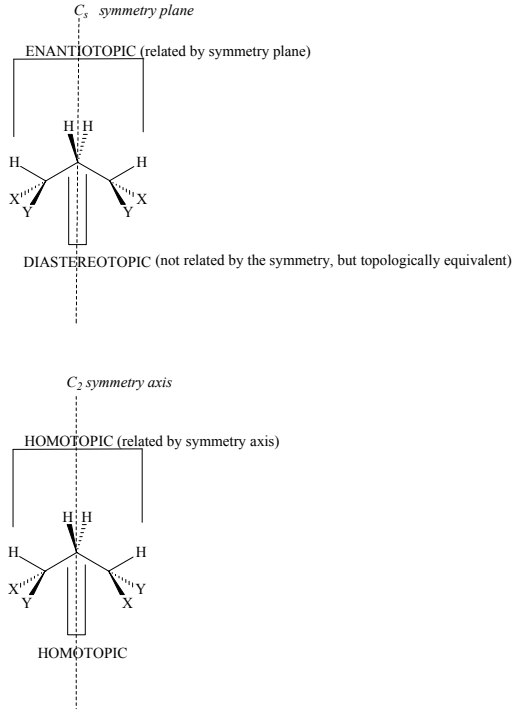
Atoms or a group of atoms, if they are not related by any symmetry operation are diastereotopic (if they are topologically equivalent) or heterotopic (if they are topologically nonequivalent).

Topological equivalence might be also explained as connectivity equivalence. In the pictures below two chlorine atoms are topologically equivalent, however, they are not symmetry related. Two methyl groups in one of the structure below are topologically different (different connectivity, different neighbors).





The examples of enantiotopic and diastereotopic relations are presented in the diagram below.



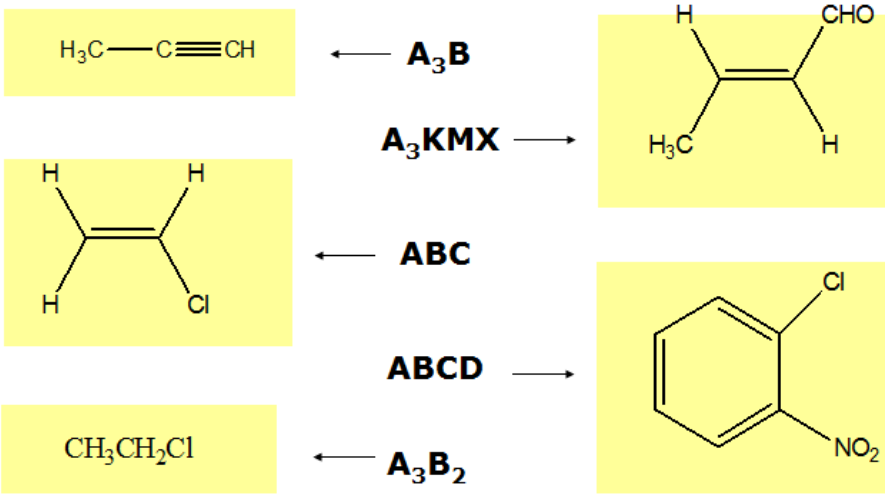
Magnetic equivalence

If two equivalent groups of nuclei, for example A and B have such a relation that every A nucleus has identical relations with every B nucleus then they are "magnetically equivalent". Otherwise they are not magnetically equivalent. See examples below.

1.5 Spin system labeling

1. Groups of chemically equivalent nuclei are given the same capital letter (A,B,C….)
2. The subscript describes the number of nuclei in the group (A_2, B_3, C_5……..)

Examples:

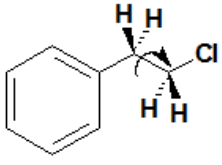


3. If the group contains magnetically not equivalent nuclei it is marked by the corresponding superscript (AA', B,B'B"……..). See example below.

Example:

It is very important to recognize the magnetically nonequivalent nuclei since the spectra become significantly different, For example, system A_2B_2 is characterized by one coupling constant J_{AB} whereas system AA'BB' by four coupling constants: J_{AB}, $J_{A'B}$, $J_{AB'}$, $J_{BB'}$.

In the example below one should expect eight lines for every group (ddd) due to AA'BB' spin system and not three (t) as for A_2B_2 system.



AA'BB'

A_2B_2

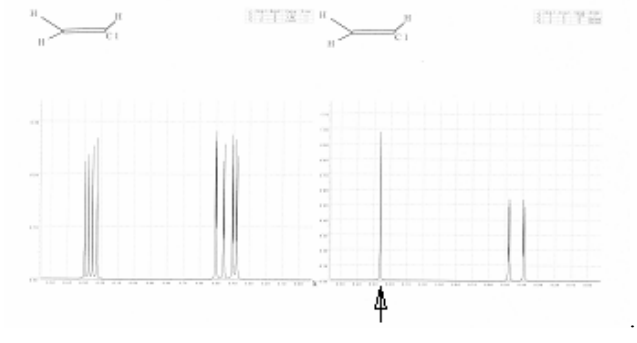
ddd ddd

t t

1.6 NMR one dimensional techniques

Homodecoupling

A simple description how the spin decoupling operates might be explained on the spin system ABC shown below (left part of the figure). Every proton signal is observed as ddd (a doublet of doublets of doublets). The application of the additional external field selectively to the chosen frequency (marked with the arrow in the figure below) results in a rapid frequency change of the selected spins between alpha and beta states. The observed effect on the spectrum – the "removal" of the coupling effect from the irradiated spins is shown in the figure below (the right part). The spin system is reduced to BC and two doublets of doublets are observed.



The procedure described above is a selective decoupling and is marked as $^1H\{^1H\}$

Broadband decoupling

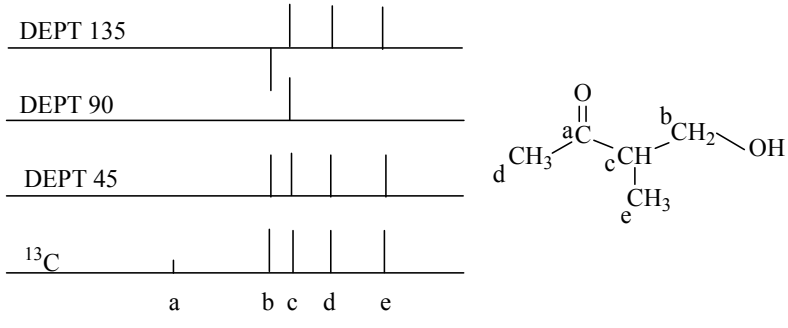
To “remove” the effect of coupling with other nuclei than observed a broad band decoupling is applied. For example, a typical procedure during the observation of the carbon nuclei is characterized with a broadband decoupling with frequencies covering all proton resonances - $^{13}C\{^{1}H\}$. The resulting spectrum contains singlets for all carbon resonances disregarding the number of protons attached.

DEPT (Distortionless Enhancement by Polarisation Transfer)

The method is very suitable to distinguish CH_3, CH_2 and CH carbon atoms. Without going into details, the special pulse sequence results in that some signals disappear, some are positive and other are negative, as shown in the table below.

	DEPT 45	DEPT 90	DEPT 135
CH	+	+	+
CH_2	+	0	-
CH_3,	+	0	+

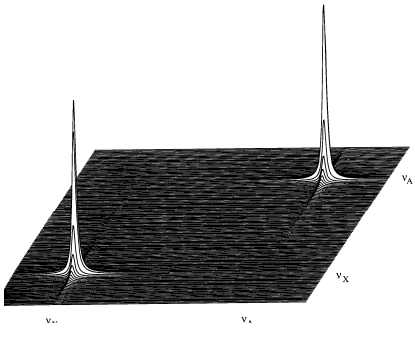
The Figure below shows the example of ^{13}C carbon spectrum and DEPT 45, DEPT 90, and DEPT 135 for 4-hydroxy-3-methyl-2-butanone

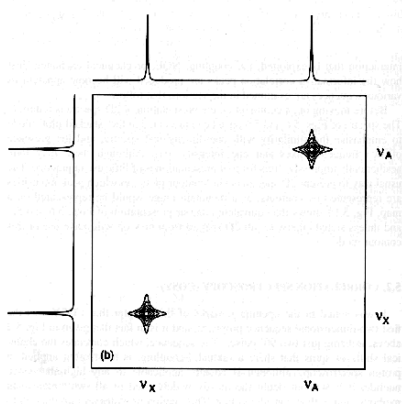


1.7 Two dimensional methods in NMR

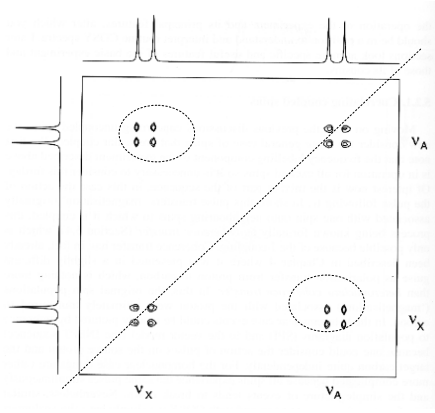
The basic two dimensional techniques refer to two-frequency dimensions. In fact the previously discussed one-dimensional methods were two-dimensional. The intensity of the signal, however, is usually never included when describing the dimensionality of the experiment.

In two-dimensional NMR techniques two frequency dimensions may represent any combination of a chemical shift or scalar couplings. The third dimension is intensity. For the two uncoupled nuclei like in proton – proton correlation spectroscopy, for example in spectra taken for CH_3COOCH_3, the two spectra representations are shown below.





If the two nuclei are coupled, signals out of the diagonal appear, indicating the interaction of the nuclei. The example below is given for a spin system AB like in $X_2C\mathbf{H}\text{-}C\mathbf{H}Y_2$

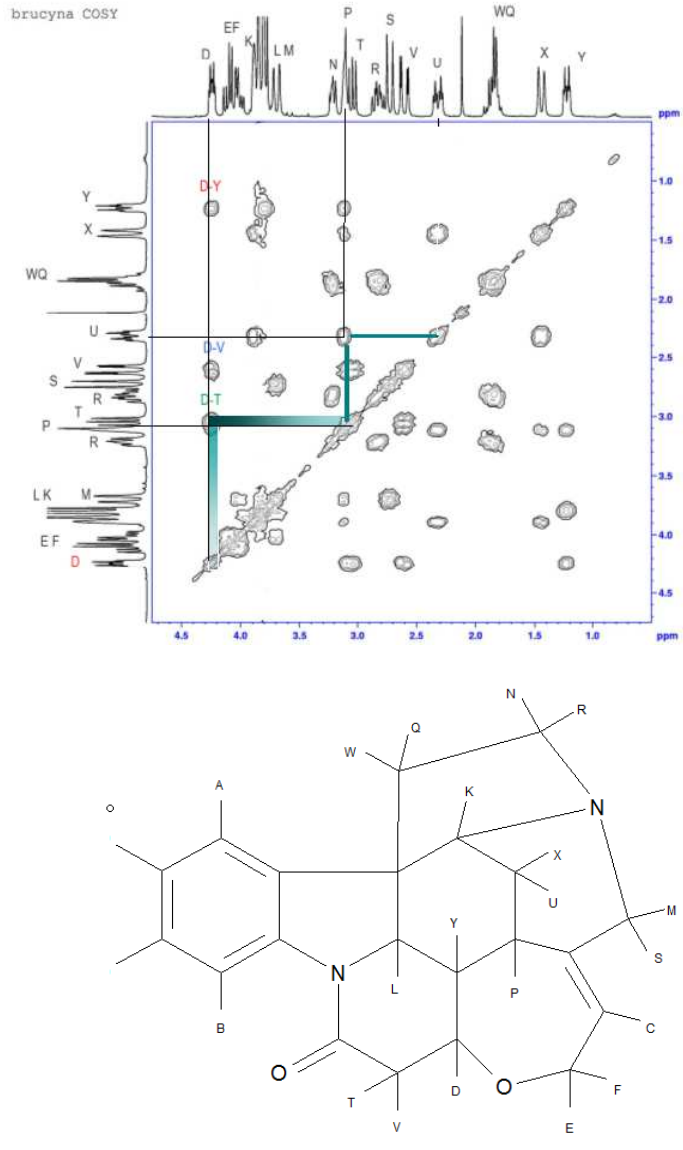


The projection of the spectrum on the two dimensions (frequency, intensity), the left and top part of the drawing, as well as on the diagonal plane (the dashed line) is typical for AB spectrum composed of two doublets. The out of the diagonal signals (the dashed circles) are named cross peaks and represent the interaction of the two interacting nuclei. Various interactions can be expressed as shown in the following sections.

More recently any combinations of data are also applied; for example a chemical shift and data from HPLC.

HHCOSY

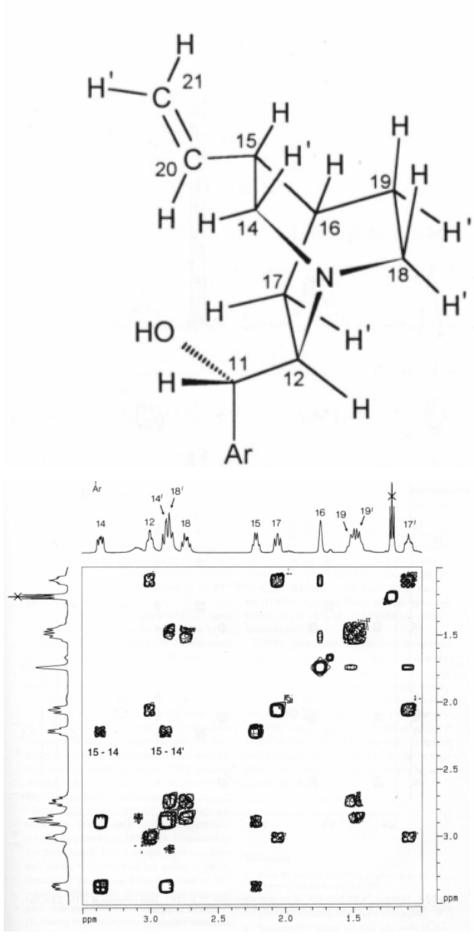
^{1}H ^{1}H correlation spectroscopy for brucine



Two kinds of peaks appear in the spectrum: diagonal and cross peaks (oustside the diagonal). Both coordinates (in the chemical shift) for the diagonal peaks are the same, however, for the cross peaks the coordinates are different ω_D and ω_P ,which means that D and P are coupled. Applying a similar analysis we find that D is also coupled to T and Y, whereas P is coupled to U.

TOCSY (Total correlation spectroscopy)

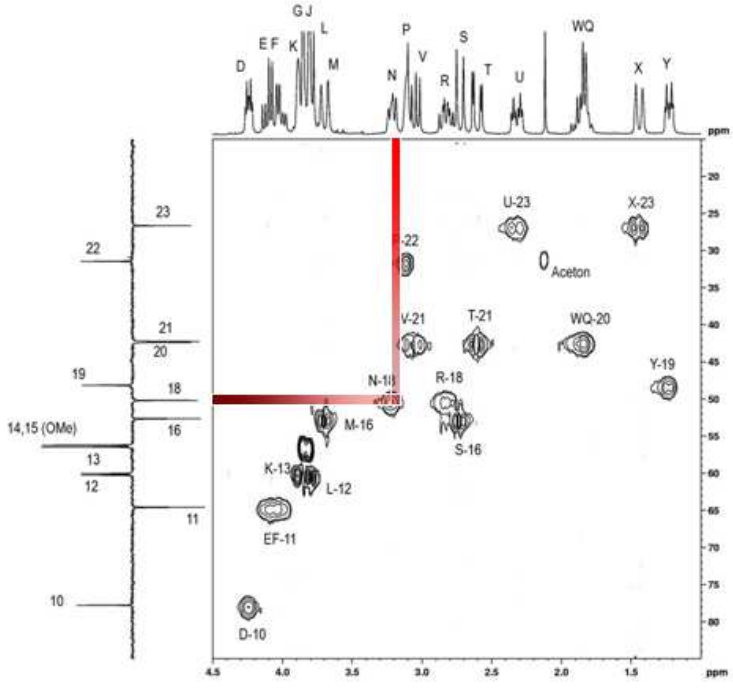
TOCSY is similar to COSY to such an extend that cross peaks indicate an interaction between spins. COSY, however, shows only cross peaks between directly coupled spins, whereas TOCSY shows all in an “unbroken chain of spins”. If A is coupled to B and B to C and C to D then the relations between all those spins are shown even if A and D are not directly coupled. In the figure below the comparison is shown between COSY and TOCSY. The COSY spectrum shows only the relation of 19 to 18 and 16. TOCSY indicates the spin system 19, 16, 18, and 17, 12, 15, 14,12. Such an analysis is very useful in structure elucidation as shown for polysaccharides since each saccharide unit represents an independent spin system.

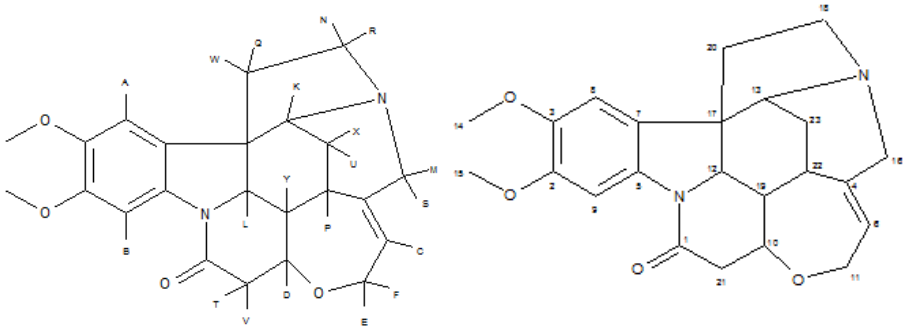


COSY TOCSY

HMQC

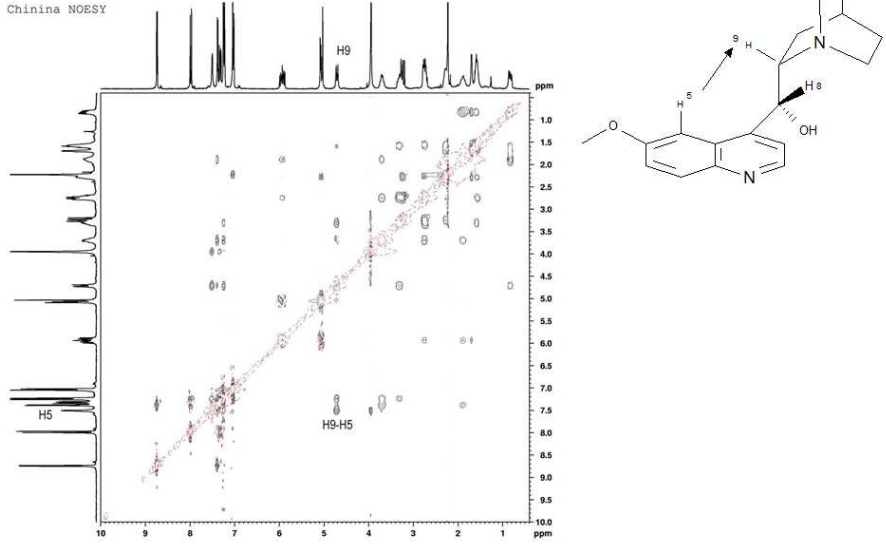
^{1}H ^{13}C correlation through one bond for brucine. The crosspeaks indicate the direct connection between the carbon and proton nuclei.



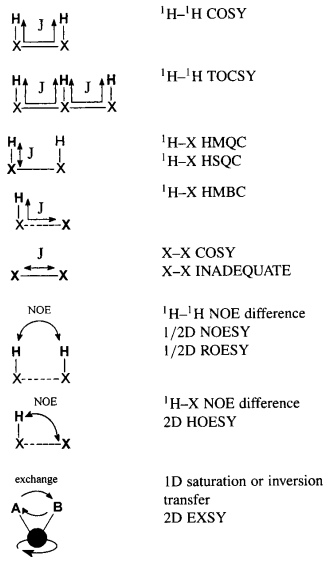


NOESY

The cross peaks of a **NOESY** spectrum indicate which protons are close in space. This is different from COSY, which relies on J-coupling to provide a spin-spin correlation. By measuring a cross peak intensity, distance information can be extracted. The example is given below.



Principal correlations established by NMR



1.8 Other methods

Other techniques are not discussed due to limited space they are suggested for individual studies

1D methods

spin echo

1D NOE

1D TOCSY

1D ROESY

INEPT

APT

INADEQUATE

2D methods

2D TOCSY

HETCOR

HMBC

HSQC

DQF COSY

1.9 Other problems for further individual studies

Energy levels and NMR

Quantum model

Vector model of NMR

Rotating frame model

Spin spin relaxation

Spin latice relaxation

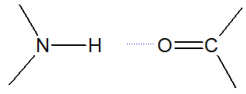
Pulse technique and fourier transfomation

1.10 NMR in hydrogen bonding and dynamic processes studies

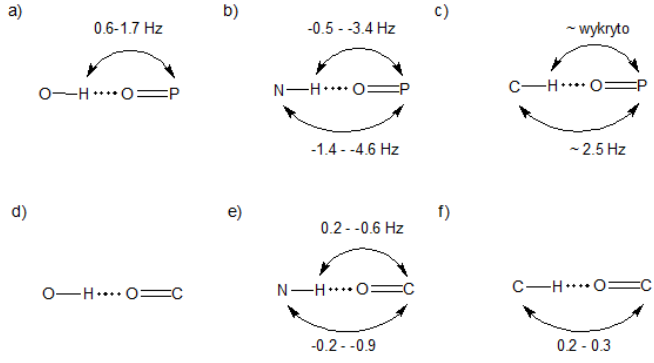
Coupling through hydrogen bond

The coupling can also be observed through the hydrogen bond.

$^{3h}J_{NC}$ $^{2h}J_{HC}$



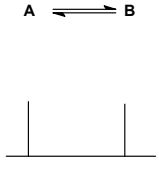
The magnitude of the effect is not very strong



but such effects are very important in the analysis of the substrate-enzyme interaction or in the analysis of the complementary nucleotides in a DNA structure.

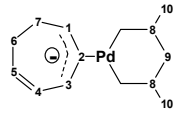
Spin saturation transfer

If we have a dynamic system in which two nuclei exchange the NMR in low an exchange condition indicates the presence of two separate signals.

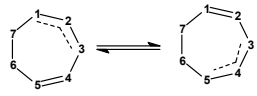


When one nucleus, for example A is saturated by applying the irradiation with frequency ν_A, it also results also in the saturation of the nucleus B unless the rate of relaxation is faster than the rate of the exchange.

Example:
For the organometalic compound;



the rapid exchange was proposed as shown below

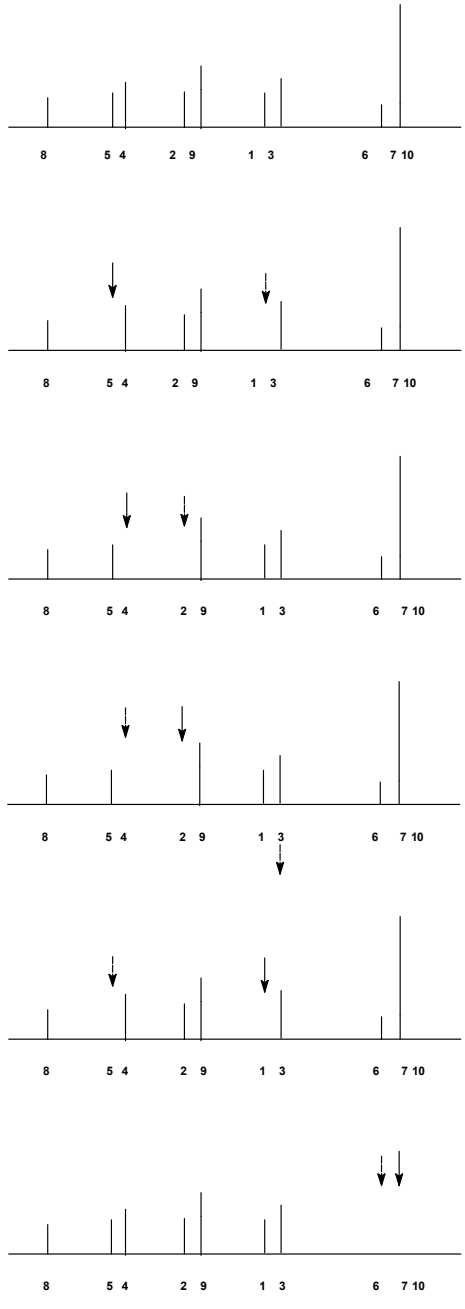


The process can be represented by the permutation groups.

(1,5)(2,4)(3)(6,7)

It means that there are four groups. In the first group atom 1 replaces atom 5 and 5 replaces 1, in the second group 2 replaces 4 and 4 replaces 2, in the third one 3 is unchanged, in the fourth group 6 replaces 7 and 7 replaces 6.

The spectra attached confirmed the proposed mechanism. The first spectrum is the original one without irradiation In each of the following spcrum the solid arrow indicates the signal which was saturated and the dashed arrow indicates the signal which also becomes also saturated due to a dynamic exchange and in consequence it disappeared.


8 5 4 2 9 1 3 6 7 10
8 5 4 2 9 1 3 6 7 10
8 5 4 2 9 1 3 6 7 10
8 5 4 2 9 1 3 6 7 10
8 5 4 2 9 1 3 6 7 10
8 5 4 2 9 1 3 6 7 10

2. MASS SPECTROMETRY

2.1 Introduction

Mass spectrometry is one of the best methods of structural investigations of organic compounds, especially macromolecules. It is advantageous to other methods of analysis in its ability to separate and analyze mixtures of compounds. MS is also an extremely sensitive technique with current levels of accuracy within the range of one mass unit. The method is still developing for the reason of high importance of applications in structural studies like: finding of molecular weight of organic compounds, investigations of structural isomers structural investigations of biopolymers and naturally occurring compounds (proteomics and metabolomics).

MS is not a spectroscopic method but rather spectrometry. The first step of measurement is to find a way to "charge" an atom or molecule (ionization). Next, the charged atom or molecule is placed in a magnetic field or subjected it to an electric field and its speed or radius of curvature relative to its mass-to-charge ratio (mass analyzer) is measured. The last part is a detection of the signal of separated ions (Figure 1).

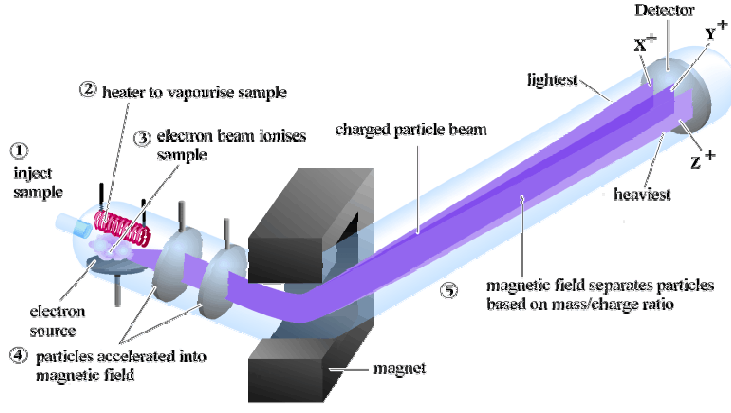


Figure 1: General scheme of typical mass spectrometer.

A typical mass spectrum (Figure 2) is characterized by sharp, narrow peaks. X-axis position indicates the m/z ratio of a given ion (for singly charged ions this corresponds to the mass of the ion). The height of a peak indicates the relative abundance of a given ion (which characterizes the stability of the obtained fragments). The big advantage of mass spectrometry is its sensitivity, which makes this method a powerful tool for analytical investigations. For

small organic molecules the MW can be determined to within 5 ppm or 0.0005% which is sufficiently accurate to confirm the molecular formula from mass alone. In the case of large biomolecules the MW can be determined within an accuracy of 0.01% (within 5 Da for a 50 kD protein).

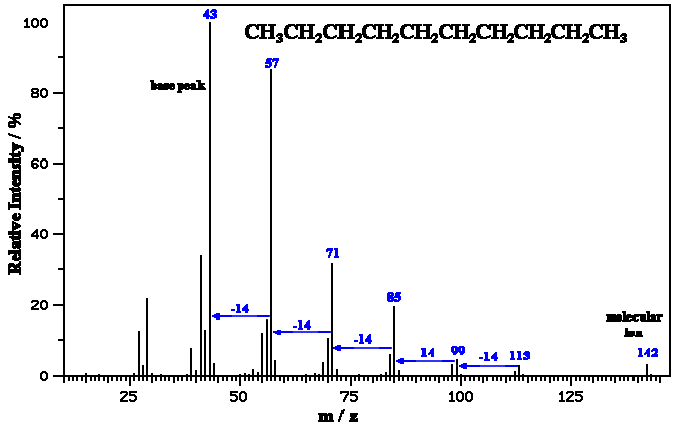


Figure 2: Typical mass spectra.

Another fact, very important in a case of analytical studies of a wide spectrum of molecules (starting from small organic compounds and ending on proteins) is a possibility of coupling a mass spectrometer with a chromatograph. Thanks to this fact, the process of purification and structure identification could be done in one step. For small organic compounds a mass spectrometer should be coupled with a gas chromatograph – GC-MS system (Figure 3), while for macromolecules liquid chromatography is advised, namely LC-MS method.

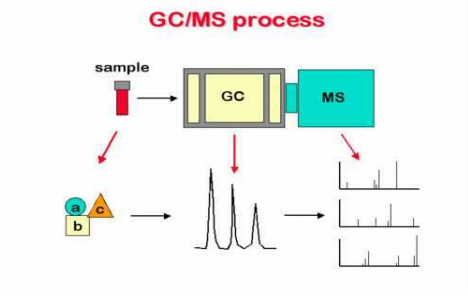


Figure 3: Scheme of GC-MS system.

2.2 Ionization methods

In general, the choice of the method which should be used for a mass spectra recording is very strongly connected with the type of an investigated sample. Two parameters of a mass spectrometer are determining the possibilities of its application – they are s type of ionization and an analyzer. The most popular methods of ionization could be classified in two groups:

1) Methods dedicated for small molecules investigations:

• **electron impact ionization (EI):** In this method a sample is introduced into an instrument by heating it until it evaporates. After that gas phase the sample is bombarded with electrons coming from rhenium or tungsten filament (energy = 70 eV). In effect the molecule is being split into fragments (70 eV >> 5 eV bonds), which are sent to a mass analyzer. This is a hard method of ionization devised to low MW organic compounds and its big disadvantage is the fact, that the sample should be introduced as a gas phase

• **chemical ionization (CI):** In this technique ions obtained from ammonium, methane, water, butane are used as "reagent" gas. In the case of methodology this type of ionization is very similar to electron impact ionization, however, it is not so hard and the obtained fragmentation of the sample is smaller. In this case, the sample also has to be introduced in a gas phase. CI is performed in two steps. In the first step a reagent gas at relatively high pressure (~ 1 Torr) is subjected to electron impact ionization. The ions obtained in this way are then available for the second stage of the CI process, in which they undergo an ion - molecule reaction with the molecule of interest. The rate constants for this type of ion - molecule reaction are very high, because the activation energy for any such favored reaction is minimal.

2) Methods used for macromolecules ionization:

• Fast Atoms Bombardment (FAB),

• ESI (Electrospray ionization),

• MALDI (Matrix Assisted Laser Desorption Ionization)

Being important in the analysis of biomolecules, these methods will be described more precisely in the next part of the paragraph.

2.3 Analyzers

The second element of a mass spectrometer, which is critical for possibilities and the quality of measurement is an analyzer. The most important parameters of an analyzer are:

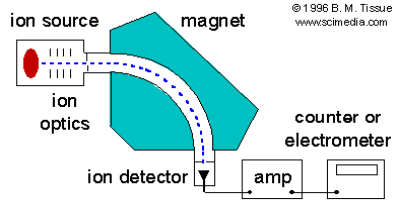
• resolution,

• the range of molecular weight for which it could be used,

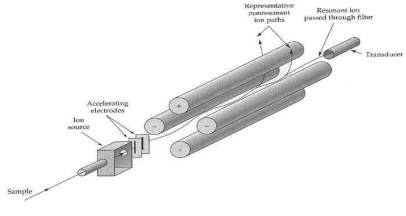
• sensitivity.

The fragments obtained as a result of ionization are accelerated and differentiated depending on their m/z (mass to charge) ratio in an analyzer. There are several types of mass analyzers:

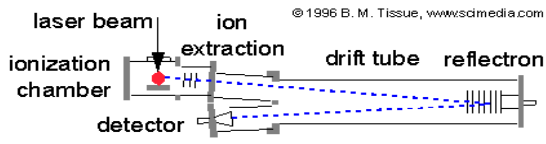
1) **Magnetic Sector Analyzer**: based on the use of an electromagnetic field, the most popular type of analyzer, which is characterized by high resolution and gives the exact mass value. Changing the voltage or the magnetic field of the magnetic-sector analyzer, the individual ion beams are separated spatially and each has a unique radius of curvature, according to its mass/charge ratio.



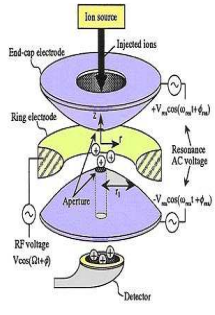
2) **Quadrupole Analyzer:** consists of four parallel metal rods with different charges. Two opposite rods have an applied potential of (U+Vcos(wt)) and the other two rods have a potential of -(U+Vcos(wt)). The applied voltages affect the trajectory of ions traveling down the flight path. For given voltages, only ions of a certain mass to charge ratio pass through the quadrupole filter and all other ions are thrown out of their original path. The disadvantage of this analyzer is low resolution, however, it gives a possibility of fast and cheap measurement.



3) **Time-of-Flight Analyzer (TOF):** making use of the fact that ions of different mass/charge ratio and all with the same initial translational energy require different times to traverse a given distance in a field-free region. In this case a differentiation of fragments is dependent on their time of travel and a drift tube length. This type of analyzer is not recommended to upper values of m/z, however, it is good for kinetic studies of fast reactions and for use with gas chromatography to analyze peaks from a chromatograph.



4) **Ion Trap Mass Analyzer:** ion traps are ion trapping devices that make use of a three-dimensional quadrupole field to trap mass-analyzed ions. This permits the ejection of ions with a *m/z* ratio lower than a prescribed value and the retention of these with higher mass. It depends on the application of radio frequency voltages between a ring electrode and two end-cap electrodes to confine the ions in a circular path. The choice of these voltages determines the *m/z* ratio below which ions are ejected. This analyzer is characterized by good resolution and often called an "all-in-one" mass analyzer.



5) **Ion Cyclotron Resonance (FT-ICR):** the use of a powerful magnet (5-10 Tesla) gives a possibility to create a miniature cyclotron, additionally the FT approach allows many ion masses to be determined simultaneously. The big advantage of this analyzer is the highest resolution and the exact mass measurement but this is very expensive equipment.

2.4 Principles of fragmentation process

The process of fragmentation very strongly depends on the type of ionization and usually is important for hard methods, which are used for investigations of small molecule structures and a determination of their molecular weights. As an effect of an ionization factor three kinds of structures could be obtained: inert molecules, cations and free radicals. The number of created fragments is bigger for hard types of ionization. As this is presented in Figure 4, in

the case of electron impact ionization, the obtained fragmentation could be very complicated, even for relatively simple molecules.

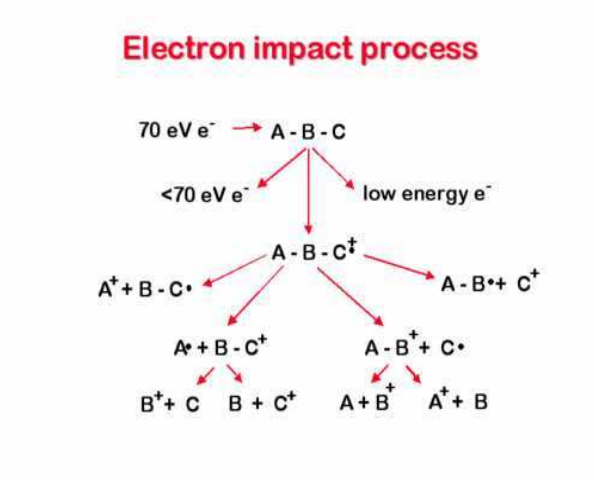


Figure 4: Scheme of process of fragmentation as electron impact (EI) ionization.

This process is strongly connected with the structure of an analyzed molecule. For the reason of limited size of this paragraph, the detailed patterns of fragmentation will not be discussed, however, there are known several general rules of this process:

- in fragmentation of an aliphatic chain the break of C-C bonds is favourised,
- in alkyl substituted aromatic rings a cleavage is very probable at the bond β to the ring giving the resonance stabilized benzyl ion or the tropylium ion,
- the C-C bond far by two bonds from a double bond are frequently cleaved,
- the C-C bonds next to the heteroatom is frequently cleaved.

In the case of biomolecules, like peptides and proteins, the use of soft methods of ionization is necessary – in this way the fragmentation process is for instance stopped at the level of a single amino acid residue.

2.5 Methods of ionization of macromolecules

In the case of molecules which are characterized by a higher molecular weight, the main problem was to avoid the necessity of introducing the substance in a gas phase. It was

mentioned before that there are three predominant methods of ionization which are advised for macromolecules studies: FAB, ESI and MALDI.

Fast Atoms Bombardment (FAB)

The technique of FAB involves the bombardment of a solid analyte and matrix mixture by a fast particle beam. The matrix (glycerol or 3-nitrobenzyl alcohol) is used to keep a "fresh" homogeneous surface for bombardment, thus extending the spectral lifetime and enhancing sensitivity. The particle beam is a neutral inert gas, typically Ar or Xe, at bombardment energies of 4-10 KeV. FAB is a comparatively soft ionization technique and is well suited to the analysis of low volatility species.

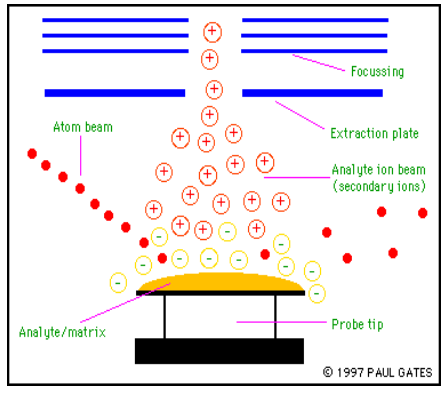


Figure 5: Fast Atom Bombardment scheme.

FAB was the first good method for investigations of bigger molecules, presently other two methods of ionization (ESI and MALDI) are predominant, The advantages of the FAB application are connected with the facts, that the method is easy and fast to operate and spectra are relatively simple to interpret. Moreover, the source itself is easily retro fitted on most MS instruments. Unfortunately, there are also several problems in the use of this method:

- moderate sensitivity due to the high concentration of the organic liquid matrix required (typically 80 to 95% glycerol),
- matrix cluster ions can, in some cases, dominate the mass spectrum,
- intense chemical background arising from the damage to the matrix caused by the particle bombardment,
- in some cases the matrix also directly reacts with the analyte, forming radical anions or causing the reduction of the analyte.

Electrospray ionization (ESI)

In this method a sample is dissolved in a polar, volatile buffer (no salts) and pumped through a stainless steel capillary (70 - 150 mm) at a rate of 10-100 mL/min. After that, the strong voltage (3-4 kV) applied at the tip along with the flow of nebulizing gas causes the sample to "nebulize" or aerosolize. The obtained aerosol is directed through the regions of higher vacuum until the droplets evaporate to a near atomic size (still carrying charges). The mechanism of ESI is based on fact that the charged droplets contain ions of both polarities, with an excess of the ions of the same polarity as the applied spray voltage. During the solvent evaporation and under the electric field the ions go to the surface of the droplet. When a critical value is reached, the ions are emitted in the gas-phase from the droplet.

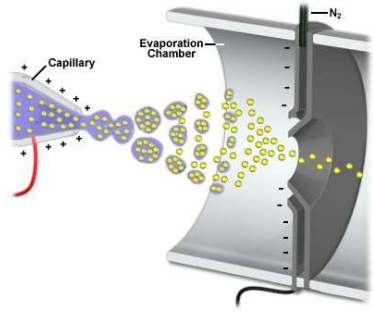


Figure 6: Electrospray ionization scheme.

ESI is specially advised for structural investigations of peptides, the advantages of the method are:

- solutions for transfer of ions to the gas phase are a wide range of solvents acceptable,
- sample introduction through LC and direct infusion
- compatible with magnet, quadrupole and time of flight analyzers,
- high sensitivity with femtomole to low picomole,
- multiple charging allows for an analysis of high mass ions with a relatively low m/z range instrument.

In the case of ESI also some disadvantages are known:

- the presence of salts and complex mixtures can reduce sensitivity,
- simultaneous mixture analysis can be poor,
- multiple charging can be confusing, especially in a mixture analysis,
- sample purity is important.

Matrix Assisted Laser Desorption Ionization (MALDI)

In this method of ionization, a sample is mixed with a UV absorbant matrix (for instance 4-hydroxycinnaminic acid for peptides) and bombarded by laser light. The light wavelength matches that of the absorbance maximum of a matrix so that the matrix transfers some of its energy to an analyte. As an effect of dissociation of the matrix, the phase changes to super-compressed gas and a charge transfer to analyte molecule is observed. This type of ionization bases on the fact that in the presence of an aromatic matrix, large molecules like peptides ionize instead of decomposing. Although the mechanism remains uncertain, it may involve the absorption of UV light by the matrix, followed by the transfer of this energy to the peptide, which then ionizes into the gas phase as a result of a relatively large amount of energy absorbed. To accelerate the resulting ions into a flight-tube in a mass spectrometer they are subjected to a high electrical field.

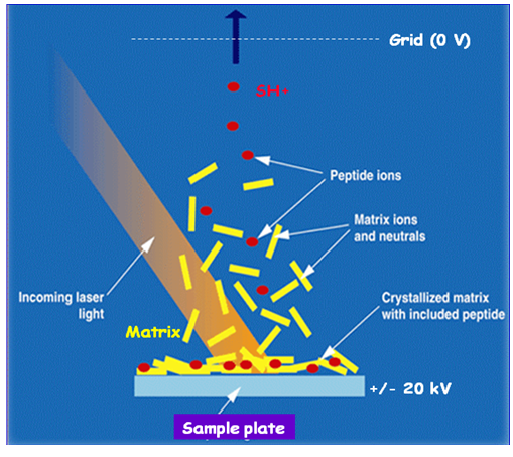


Figure 7: MALDI ionization scheme.

The choice of matrix is crucial for success in MALDI experiments and for the control of fragmentation, usually three types of matrix are in application:

- solid (most common) where co-crystallization of an analyte and matrix is necessary,
- liquid - IR or UV absorbing liquid containing an analyte,
- combination of two phases: liquids with an absorbing solid.

MALDI is an ionization method advised for proteins sequence studies, which posseses some big advantages like:

- Practical mass range of up to 300,000 Da. Species of much greater mass have been observed using a high current detector.
- Typical sensitivity to the order of low femtomole to low picomole.
- Soft ionization with little fragmentation is observed.
- Tolerance of salts in millimolar concentrations.
- In contrast to ESI, suitable for the analysis of complex mixtures.

On the other hand, several disadvantages of MALDI are known:

- Matrix background, which can be a problem for compounds below a mass of 700 Da. This background interference is highly dependent on the matrix material .
- Possibility of photodegradation of an analyzed substance by laser desorption/ionization.
- Acidic matrix used in MALDI may cause degradation of some compounds.

2.6 Methodology employed in investigations of peptides and proteins

Recording and interpretation of mass spectra of biopolymers, like peptides and proteins, for the reason of high value of a molecular weight are more complicated. Besides the types of ionization and an analyzer, also a special methodology of measurements is necessary. The most typical in this situation is the application of tandem mass spectrometry (MS/MS) – thanks to this there are possibilities to make a strict analysis of peptides or proteins sequence. In fact, MS/MS technique is analogical to the use of a mass spectrometer coupled with chromatography. In this case, two analyzers are separated by a collision cell into which an inert gas (argon, xenon) is admitted to collide with the selected sample ions and bring about their fragmentation. The analyzers can be of the same or of different types, the most common combinations being:

- quadrupole - quadrupole
- magnetic sector - quadrupole
- magnetic sector - magnetic sector
- quadrupole - time-of-flight.

The first analyzer is used to select user-specified sample ions arising from a particular component; usually molecular-related ions. These chosen ions pass into the collision cell, where they are bombarded by gas molecules which cause fragmentation. These fragment ions are analyzed by the second analyzer. All the fragment ions arise directly from the precursor

ions specified in the experiment, and thus produce a fingerprint pattern specific to the compound under investigation. The differences in the methodology and philosophy of measurements of tandem mass spectrometry as compared to a standard experiment are shown in Figure 8. Of course, more developed systems, like triple or even quadrupole mass spectrometers are known.

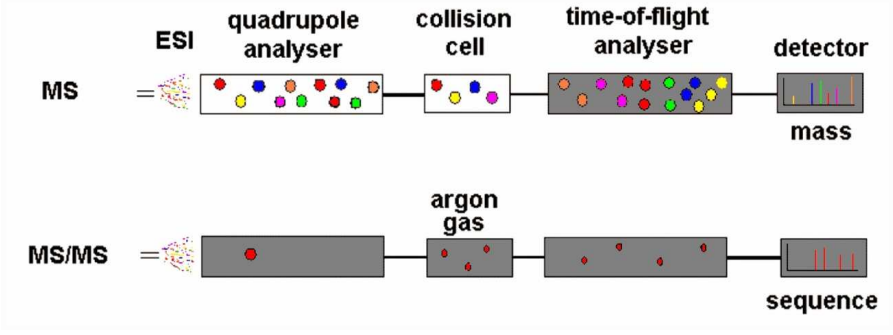
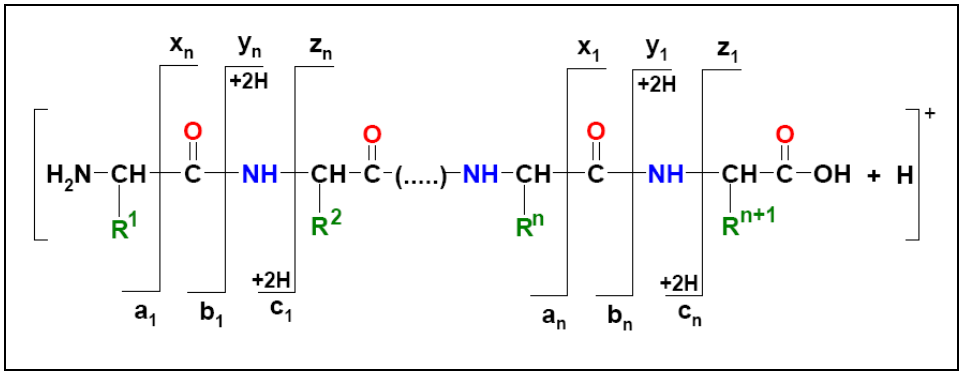


Figure 8: The differences in standard and tandem mass spectrometers.

There are three different types of bonds which can fragment along the amino acid backbone: **NH-CH**, **CH-CO**, and **CO-NH** bonds. Each bond breakage gives rise to two species, one neutral and the other one charged. Only the charged species is monitored by the mass spectrometer. The charge can stay on either of the two fragments, depending on the chemistry and relative proton affinity of the two species. Hence, there are six possible fragment ions for each amino acid residue and these are labelled as in the diagram, with the a, b and c" ions having the charge retained on the N-terminal fragment, and the x, y" and z ions having the charge retained on the C-terminal fragment. The most common cleavage sites are at the CO-NH bonds, which give rise to the b and/or the y" ions. The mass difference between two adjacent b ions, or y ions is indicative of a particular amino acid residue.



Another method which is used, especially in the case of proteins is Protein Mass Fingerprinting (PMF). Using sequence specific proteases (for example trypsin) we can cut the protein chain selectively at certain amino acids in the case of the mentioned above trypsin only at lysine and arginine residues. The protein fragments have distinct weights. They can be accurately measured by mass spectrometry. A list of these weights is like a fingerprint which is unique to the protein and can be used to identify it. In the next part of measurement, also in this case, tandem or even triple mass spectrometry could be used.

The general rules and methodology in the interpretation of peptide/proteins spectra are:

1. Choose on intense peak
2. Look for ions that are different in mass from the fragment by a specific amino acid mass. If there are, look for other fragment that differs by a specific amino acid mass to build a partial sequence.
3. Decide the type of fragments that corresponds to the above sequence by either a mass difference to the precursor or complimentary fragments.
4. Look for other complimentary fragments and low mass ions to confirm the partial sequence.
5. Choose another unassigned peak and repeat the procedure until the sequence is determined.

Supplementary sources:

1. R.M. Silverstein, F.X. Webster, D.J. Kiemle „Spektroskopowe metody identyfikacji związków organicznych" PWN Warszawa 2007

2. R.A.W. Johnstone, M.E. Rose „Spektrometria mas" PWN 2001.

http://elchem.kaist.ac.kr/vt/chem-ed/ms/ms-intro.htm

http://masspec.scripps.edu/information/intro/

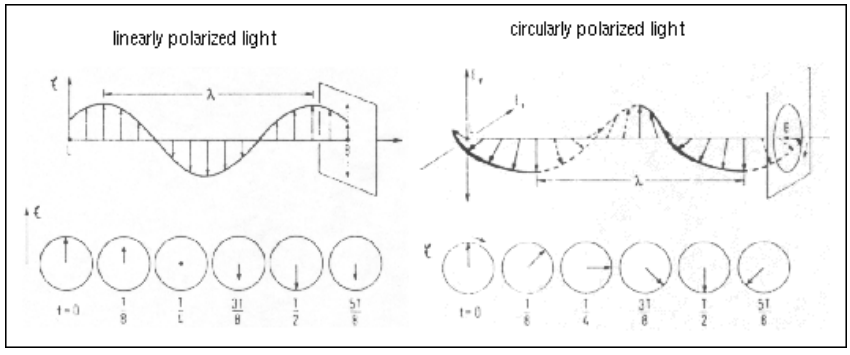
http://ms.mc.vanderbilt.edu/tutorials/ms/ms.htm

http://www.abrf.org/ABRFNews/1996/September1996/sep96iontrap.html

http://www.chem.arizona.edu/massspec/intro_html/intro.html

3. CIRCULAR DICHROISM SPECTROSCOPY

The electromagnetic radiation is a complex wave form which can be considered to behave as two simple wave motions at right angles to each other. One of these waves is magnetic (M) and the other is electric (E). Electromagnetic waves are generated by oscillating electric or magnetic dipoles and are propagated at the speed of light (c). Since the E- and the M-components are always perpendicular to each other, it is sufficient, in many cases, to consider only the E-component in describing the wave. The most typical - unpolarized light - contains oscillations of the E-components in all directions perpendicular to the direction of propagation. Linearly polarized light results when the direction of the E-component is restricted to a plane perpendicular to the direction of propagation while its magnitude oscillates. For another form of polarization - circularly polarized light - the magnitude of the oscillation is constant and the direction oscillates.



Circular dichroism (CD) spectroscopy is a form of light absorption spectroscopy that measures the difference in absorbance of right- and left-circularly polarized light (rather than the commonly used absorbance of isotropic light) by the substance. The CD measurement results from differences in their extinction coefficients. This is measured experimentally as an absorption difference:

$$\mathbf{\Delta A = A_L - A_R}$$

The CD signal is usually expressed as:

Molar circular dichroism: $\Delta\varepsilon(\lambda) = \varepsilon_L(\lambda) - \varepsilon_R(\lambda)$

Molar ellipticity: $\theta = 3298\Delta\varepsilon(\lambda)$

The primary criterion for CD spectra recording is the existence of least one center of chirality in a molecule (in the case of a racemic mixture the effect is not observed). It has been known that CD spectra between 260 and approximately 180 nm can be analyzed for the different secondary structural types: α-helix, β-sheet, turn and others. Actually, two main currents are observed in applications of this method in structural and conformational studies:

1) Conformational preferences of peptides and proteins.

2) Investigations of metal ions complexes' geometry.

For molecules that absorb in the ultraviolet and a visible range, CD can be quite useful for determining the secondary structure of proteins of an unknown structure as well as for probing conformational changes occurring in various biological processes or manipulations. The simplest method of extracting secondary structure content from CD data is to assume that a spectrum is a linear combination of CD spectra of each contributing secondary structure type – for instance "pure" α-helix, "pure" β-turn etc. weighted by its abundance in the

polypeptide conformation. For that reason, this spectroscopic method is very often is used for fast, preliminary conformational studies. The majority of CD measurements on peptides and proteins are performed from ~180-~220nm ("far-UV" region), where the signal is dominated by transitions of the amide chromophores. The main advantage of CD spectroscopy is that different secondary structures produce quite different signals, for instance:

- α-helix: negative band at 222nm and also negative and positive signals at 208nm and 190nm
- β-turn: single negative band at ~215nm and single positive band at ~198nm

Additionally, the general view of CD spectra gives preliminary information about conformational preferences of a molecule. In the case when CD signals are not very intensive, it means that this is the effect of equilibrium between at least several conformers so in general a molecule adopts disordered conformation in a solution.

The application of CD spectroscopy for conformational studies in these macromolecules can be largely classified as:

1) Monitoring of conformational changes (monomer-oligomer, substrate binding, denaturation).

2) Estimation of secondary structural content (this peptide is 35% helical under some conditions).

One of the simplest possible applications of CD spectroscopy is monitoring a protein secondary structure damage by thermal denaturation. In room conditions the observed spectra (see Figure 1) suggest that the predominant is α-helical conformation. With the increase of temperature, the shape of CD spectra is changing and for temperature 81°C a signal typical for a random coil is observed.

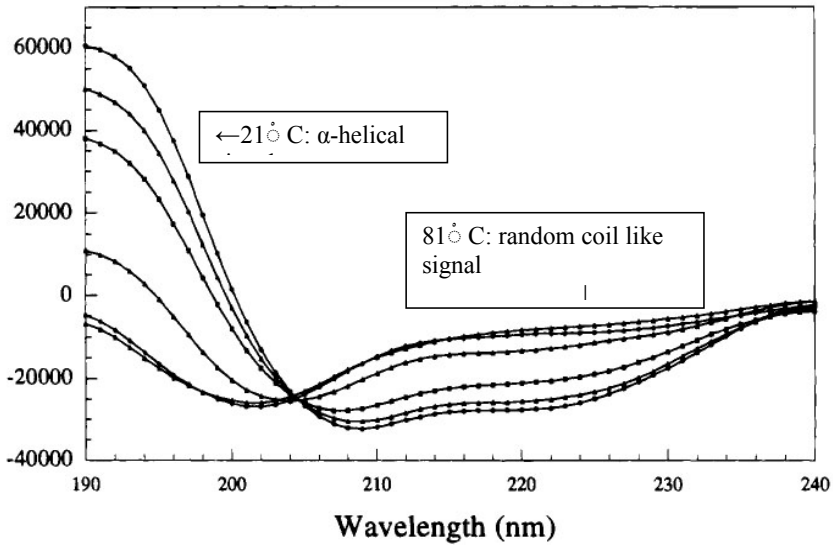


Figure 1: Changes of CD spectra of protein during thermal denaturation.

The another situation when CD spectroscopy is also a very sensitive tool for investigations of the changes in conformational preferences of peptides is presented in Figure 2. CD spectra of tripeptide Boc-G-Δ^ZF-F-OMe measured in four different solvents show that a molecule adopts similar, not ordered conformations in two solvents – acetonitrile and DMSO. On the other hand, high intensity of the absorption peaks for the spectra measured in methanol suggests more ordered conformation of the peptide in this solvent.

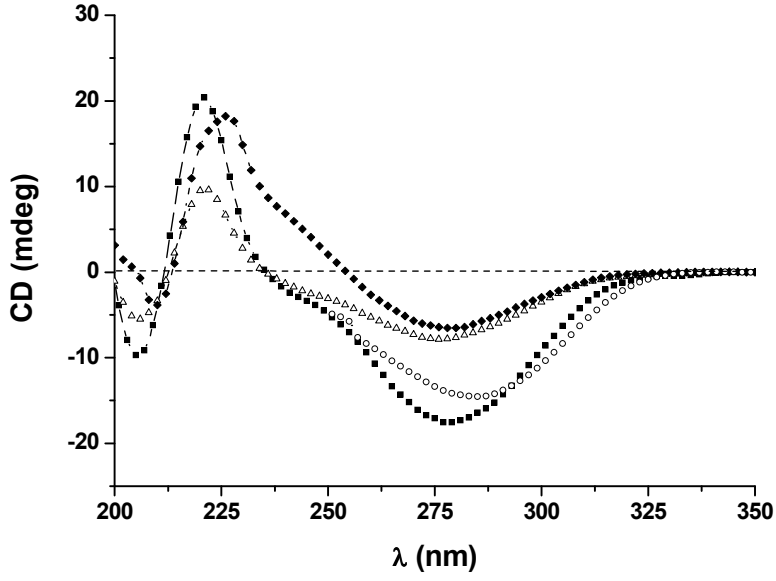


Figure 2: CD spectra of tripeptide Boc-G-Δ^ZF-F-OMe measured in: methanol (quadrats), dichloromethane (circles), acetonitrile (triangles) and DMSO (rhombs). The spectra made in dichloromethane were recorded in a narrow range for the reason of technical problems.

As it was mentioned before, circular dichroisn spectroscopy is also widely used in investigations of metal ions complexes, especially when peptides are ligands. Thanks to high sensitivity and resolution, this method gives much more information about complex geometry than UV-Vis spectroscopy. In Figure 3 the process of coordination of hexapeptide SRHWFL-NH_2 with Cu^{2+} ions is presented. Water solution, where the concentration of the peptide was 6×10^{-4} M, has been titrated in six steps by the portions of 1×10^{-4} M solution of metal ions. The bands of absorption observed for the wavelength 560nm ($\Delta\varepsilon$ = -0.52) and 482nm ($\Delta\varepsilon$ = +0.42) are the effect of d-d transitions in Cu^{2+} ions which are coordinated by four nitrogen atoms of the peptide. The increase of intensity of the signal for the value 314 nm ($\Delta\varepsilon$ = +0.63) is characteristic for one of the charge transfer transitions, which is also connected with creating of the complex.

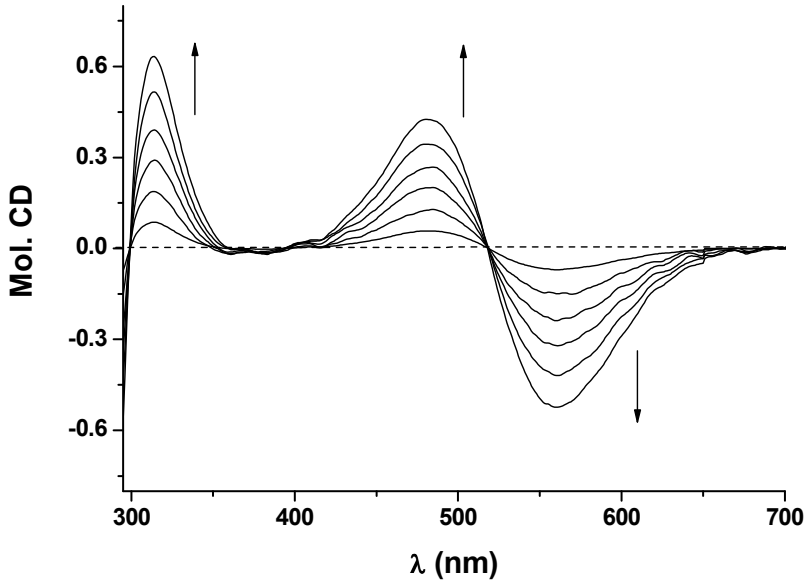


Figure 3: CD spectra of complexes of hexapeptide SRHWFL-NH_2 with Cu^{2+} ions. The arrows on the graph show the increase of coordination process which is the effect of the increase of metal ion's concentration.

Supplementary sources:

1. R.W. Woody "Circular dichroism of peptides" in The peptides: analysis, synthesis, biology, vol.. 7, academic Press 1985

2. I.Z. Siemion "Biostereochemia" PWN 1985

4. Computational techniques in spectroscopic methods useful in medicinal chemistry.

The present stage of development of computational techniques, both hardware and software meanings, does not permit to waste such a potential and not to use hese methods in the spectroscopic field. The application of the computational techniques in spectroscopic methods, in the aspect of bioactive compounds, could be divided at least into two fields:

a) simulation/prediction of theoretical spectra
b) application of spectroscopic data to calculation of the structure and conformation of the investigated compounds

4.1 Theoretical simulation of the spectra of real compounds.

Before any simulation of olecular spectra could be possible, information about the optimal structure and the conformation of the investigated molecule is required. The available software enables us to use several methods, based on a different level of precision, which could be applied to calculation of the molecular structure. In *Table 1*there are presented examples of the available software. The calculation method selected for The structural optimization should be as good as the complexity of the investigated molecule and the available hardware capacity allows. In the case of calculation methods, they could be selected from the most imprecise molecular mechanics, meant for large systems like proteins to semi-empirical methods, meant for medium systems like small proteins or peptides, up to very accurate *ab initio* methods which are meant for small molecules.

A Classical molecular mechanics (*MM*) method relies on description of atoms building an investigated molecule, by a specific atom type. The collection of these atom types is called Force Field (FF). In literature there could be found numerous type of force fields, with the mostly used Amber[i], GROMOS96 [ii,iii], CHARMM[iv], XPLOR[v] and others. From the point of view of this methods the most important are nuclei, whereas all electrons are neglected during the calculation process, therefore the molecular mechanic is the method which is completely useless to simulation of properties where electrons play a crucial role. In view of further application in a spectroscopic simulation, force field methods are the least significant methods.

Table 1. Example of available software used in computational chemistry

Software name	**Implemented theory level**	**Commercial/GNU license**
Gaussian[vi]	Molecular mechanic (Amber, Dreiding, UFF), Semi-empirical (AM1 PM3, PM3MM, PM6), quantum chemistry (HF, DFT, TD-DFT, MP, CI, CC)	commercial
Gamess[vii]	Semi-empirical (AM1 PM3, PM3MM), quantum chemistry (HF, DFT, TD-DFT, MP, CI, CC)	Freeware/open source
Turbomole[viii]	Molecular mechanic (UFF), quantum chemistry (HF, DFT, MP, CC)	Commercial
MOPAC	Semi-empirical (MINDO, MNDO, AM1, PM3, PM5)	Open source
Q-Chem [ix]	quantum chemistry (HF, DFT, CC, CI)	Commercial

Semi-empirical methods represent a compromise between an inaccurate and mostly qualitative structure obtained from force filed methods and quantitative, but time-consuming, results available from *ab initio* methods. Only valence electrons are taken heed during semi-empirical calculation and other electrons, from the inner orbital, called core electrons, are neglected. In fact, core electrons are not ignored during semi-empirical calculation, they are built in parameters, which are determined from the experimental data. Because semi-empirical methods take into account valence electrons, these methods could be applicable not only for quite an accurate way to predict he molecular structure, but also for calculation of its properties, like UV spectra.

Ab initio is a Latin term for "from the principles". These methods are the most computionally time-consuming, but on the other side they give the most accurate structures of the

investigated compounds as well as quantities information about the their molecular properties. The quality of the results depends on at least three factors:

- selected methods (Hartree-Fock, Density Functional Theory *DFT*, Moller-Plesset perturbation theory *MP*, Configuration Interaction *CI,* Coupled cluster *CC*)
- selected basis set[x]
- starting geometry

The more accurate results required a bigger basis in general sometimes they also demanded a specialist basis (for example for a magnetic resonance shielding tensor or a spin-spin coupling constant[xi]) and more complicated methods to be used. This pursuit for better and better results evokes the increasing of the time cost of calculation (Table 2).

Table 2. Depending of calculation methods and size of basis set on calculation time consuming.

Methods	**Scaling of calculation time**
Semi-empirical	N^2 (small- to medium-size molecules)
HF	N^2 - N^3
DFT	N^3
MP	N^5 - N^{10}
CC	N^5 - N^8
CI	N^6 - N^{10}

4.1.1 Theoretical calculation of IR spectra.

Infrared spectroscopy is one of the most commonly used analytical techniques for molecular analysis. IR spectra could be divided into two regions. The first, called a *fingerprint region*, is the region from about 1500 to 500 cm^{-1}, which usually contains a very complicated series of vibration. The second region, from about 4000 to 1500 cm^{-1}, contains vibration characteristic for a functional group. As the molecular weight is increasing, the complexity of the observed spectra is also growing. Because of that, a rational and accurate interpretation of IR spectra of new, unknown compounds, especially in the finger print region, is sametimes really difficult.

In that situation computational chemistry could be helpful, which allows to calculate theoretical spectra, moreover, it also allows to simulate these spectra reflected solvents influences.

The performance of the vibrational analysis with Gaussian, essential for assigning IR spectra, is made in a mass weight Hessian matrix (*Eq. 1*). After generating coordinates in a rotating and translating frame and transforming the Hessian into internal coordinates, it is possible to calculate the frequencies and force constants, which are essential to determine IR

$$f_{MWC_{ij}} = \frac{\left(\partial^2 V / \partial \xi_i \partial \xi_j\right)_0}{\sqrt{m_i m_j}} \qquad \textbf{Eq. 1}$$

The present quantum chemical software allows to predict IR spectra with using *ab initio* and semi-empirical methods.

Example of calculation of infrared spectrum.

All the calculation presented here were made in Gaussian'09 when using Polarizable Continuum Model (PCM). The first step in calculations should be a structure optimization. It was realized at the DFT level with using B3Lyp hybrid functional and basis 6-311g with a polarization function on the protons and heavy atoms (*6-311g(d,p)*). The structure of Gaussian input file is shown bellow (*Calc. 1*). Against On the yellow background they are defined:

- after % sign checkpoint file, where important results are written
- after # sign chosen methods and basis separated by / sign (B3Lyp/6-311g(d,p)), the type of calculation OPT (optimization), there is also information that a calculation should be performed with using the solvent model PCM and there is information about a selected solvent (SCRF=(PCM,Solvent=DMSO))

Against the gray background there are comments/notes/titles. On the are defined total charge and multiplicity of calculated molecule (0 1). After this section (against the blue background) the molecular structure in the internal coordination is defined and after that, against the green background, the starting values of variables are defined.

Calc. 1 Structure of input file used to optimization of the alanine structure.

```
%chk=ala.chk
#b3lyp/6-311g(d,p) opt SCRF=(PCM,Solvent=DMSO)

Optimization

0 1
n
c   1 cn2
c   2 cc3       1 ccn3
o   3 oc4       2 occ4       1 dih4
o   3 oc5       2 occ5       4 dih5
h   1 hn6       2 hnc6       3 dih6
c   2 cc7       3 ccc7       1 dih7
h   2 hc8       3 hcc8       1 dih8
h   7 hc9       2 hcc9       3 dih9
h   7 hc10      2 hcc10      3 dih10
h   7 hc11      2 hcc11      3 dih11
h   1 hn12      2 hnc12      6 dih12
h   1 hn13      2 hnc13      12 dih13

cn2         1.460000
cc3         1.510000
ccn3        111.600
oc4         1.220000
occ4        121.000
dih4        113.000
oc5         1.220000
occ5        122.500
dih5        180.000
hn6         1.020000
hnc6        121.000
dih6        61.000
cc7         1.536000
```

ccc7	**110.100**
dih7	**122.910**
hc8	**1.100000**
hcc8	**108.234**
dih8	**-118.193**
hc9	**1.100000**
hcc9	**110.300**
dih9	**179.999**
hc10	**1.100000**
hcc10	**110.300**
dih10	**60.000**
hc11	**1.100000**
hcc11	**110.300**
dih11	**-60.326**
hn12	**1.008000**
hnc12	**109.000**
dih12	**120.000**
hn13	**1.008000**
hnc13	**109.000**
dih13	**120.000**

After the normal termination of the calculation, this information is given at the end of the output file, in this same part of the output there is also information about the time of calculation (*Figure 1*).

```
Job cpu time:  0 days  0 hours 35 minutes 17.3 seconds.
File lengths (MBytes):  RWF=  54 Int=     0 D2E=   0 Chk=  4 Scr=     1
Normal termination of Gaussian 09 at Thu Jan 13 11:01:49 2011.
```

Figure 1. Terminal part of Gaussian output file

When the optimal structure of alanine was calculated, next step could be performed and the calculation of the IR could be done (*Calc. 2*). If a checkpoint file from an optimization step is available, the optimal structure required for a frequency calculation could be read from it by

using commands *Geom=Checkpoint Guess=Read*, and for IR calculation keyword *FREQ* is required (*Calc. 2)*.

Calc. 2 Structure of input file used fir calculation of IR spectrum.

%chk=ala.chk

#b3lyp/6-311g(d,p) freq *Geom=Checkpoint Guess=Read* SCRF=(PCM,Solvent=DMSO)

Optimization

0 1

After the normal termination of the calculation, the pertinent information at the end of the output file is written (similar to *Figure 1*). The information, required to assign the IR spectrum, is written in the part of output starting from the information about the calculated integral values of the second derivatives,All the calculated frequencies are listed in the line beginning with "Frequencies" (*Figure 2*). In the following lines it could be found the information about reduced masses (*Red. Masses*), force constants (*Frc consts*) and the intensity of a signal on the IR spectrum (*IR Inten*). It is worth to knowing that the method and basis chosen for the calculations had an influence on the obtained results and the obtained theoretical IR spectrum was comparable with the experimental one. It was necessary to use the precomputed scaling factor. These factors could be found, for example, on the website of the National Institute of Standards and Technology (*http://cccbdb.nist.gov/vibscalejust.asp*)

	1	**2**	**3**
	A	**A**	**A**
Frequencies	**72.7478**	**249.8945**	**253.2637**
Red. Masses	**3.9457**	**1.3102**	**2.1107**
Frc consts	**0.0123**	**0.0482**	**0.0798**
IR Inten	**4.0488**	**4.7095**	**19.5402**

Figure 2. Part of Gaussian output file with information required to assign IR spectra.

To generate the IR spectrum based on the Gaussian output file several softwares could be used, like commercial GaussView[xii], or freeware Molden[xiii] and GaussSum[xiv]. In *Figure 3* it is presented the calculated IR spectrum visualized in the GaussSum with scaling factor 0.967.

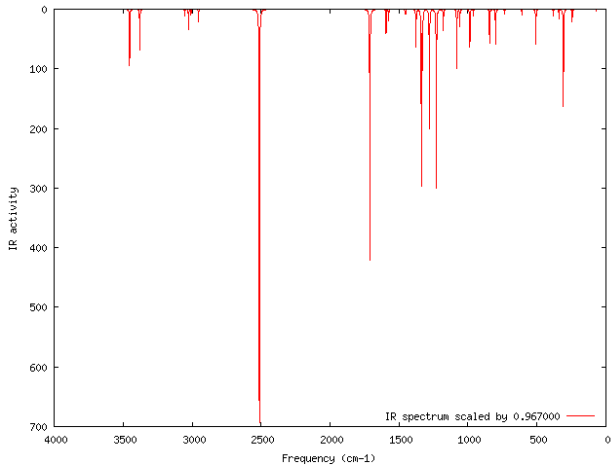
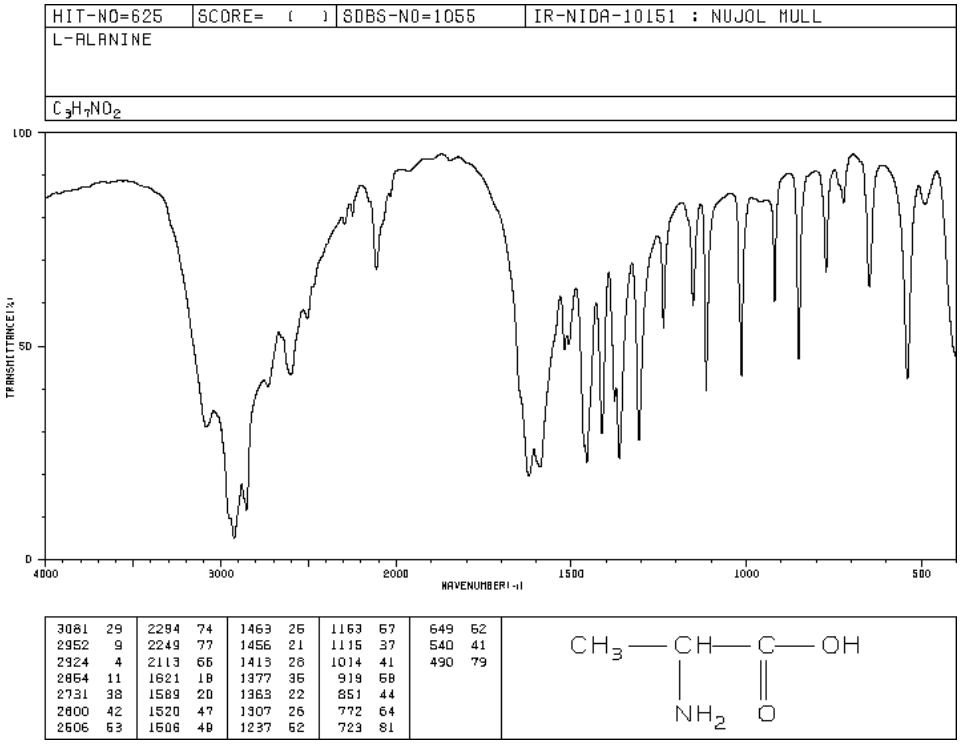


Figure 3. Visualization of the calculated IR spectrum.



3081	29	2294	74	1463	25	1163	57	649	62
2952	9	2249	77	1456	21	1115	37	540	41
2924	4	2113	66	1413	28	1014	41	490	79
2864	11	1621	18	1377	35	919	68		
2731	38	1589	20	1363	22	851	44		
2600	42	1520	47	1307	26	772	64		
2506	53	1506	49	1237	52	723	81		

Figure 4. IR spectrum of $_{L}$-Ala in nujol (http://riodb01.ibase.aist.go.jp/sdbs/) with the strong signal from nujol (2950-2800cm^{-1}, 1465-1450 cm^{-1}, 1380-1370 cm^{-1}).

4.1.2 Theoretical calculation of the NMR spectra.

Nuclear Magnetic Resonance (NMR) spectroscopy is a very useful technique for obtaining structural information. This is because a small change of the molecular structure strongly influences the NMR spectra. On the other hand, because of this sensitivity the theoretical calculation of these spectra is very difficult.

The prediction of the NMR spectra is possible when using three different methods: *ab initio*, semi-empirical and empirical.

In the case of *ab initio* methods chemical shifts are determined by subtracting the isotropic shielding tensor for the molecule of interest from the values obtained for the compound treated as a reference. Selecting as a reference an object molecule similar, to the investigated one, is very useful and enable us to obtain values closer to the experimental ones. It is especially important in the case of 13C NMR. In Table 3 there are presented reference molecules commonly used in theoretical calculations. The calculation of the shielding tensor is possible using several techniques, but the most popular and the most frequently used one is GIAO (*Gauge Invariant Atomic Orbitals*). This technique is supported in the mostly used quantum chemical packages, like Gaussian or Turbomole.

The semi-empirical methods used for the calculation of NMR spectra are implemented in a program called HyperChem[xv]. The obtained spectra are better for compounds similar to the molecules used to parameterize the methods. This kind of calculations are useful for large molecules, but predicted spectra are not so accurate as those from *ab initio*. On the other hand, the application of *ab initio* methods to the calculation of shielding tensor for such large molecules is, at this moment, very rare because of hardware limitations.

Table 3. Reference molecules commonly used in calculation of chemical shift.

Type of spectra	**Reference compound**
1H	TMS
13C	TMS (aliphatic carbons)
13C	Benzene (aromatic carbon)
15N	nitromethane

The estimation of the NMR spectra is also possible at the empirical level. The simplest technique uses the additive character of the chemical shifts. The biggest advantage of this method is its rapidility. It is possible to perform these calculations in a real time scale on a personal computer for small organic compounds with common functional groups. This kind of methods are the least accurate from from all the available ones. Empirical methods are implicated, for example, in ChemOffice.

Example of application of the selected methods to calculate the ^{1}H and ^{13}C NMR spectra of alanine.

Ab initio

For the proper calculation of the chemical shift isotropic shielding tensor for a reference molecule should be calculated. As a reference object for 1H and 13C NMR tetramethylsilane (TMS) could be chosen. The performed calculation used Polarizable Continuum Model (PCM) available in Gaussian09, with this same solvent model used for all calculations for TMS and alanine.

The first step of calculations for both TMS and alanine was structure optimization. It was realized at the DFT level with the use of B3Lyp hybrid functional and basis 6-311g with a polarization function on the protons and heavy atoms (*6-311g(d,p)*). The optimization of the alanine structure is similar to this shown in chapter *1.1.1* and the optimization of TMS was made with the following input file Calc. 3.

Calc. 3. Gaussian input file for optimization of TMS structure.

```
%chk=tms-for-nmr.chk
#b3lyp/6-311g(d,p) opt freq gfinput SCRF=(PCM,solvent=water)

Optimization  of the TMS geometry

0 1
si
c   1 cs2
c   1 cs3      2 csc3
c   1 cs4      3 csc4      2 dih4
c   1 cs5      2 csc5      3 dih5
```

h	3 hc6	1 hcs6	2 dih6
h	3 hc7	1 hcs7	6 dih7
h	3 hc8	1 hcs8	6 dih8
h	4 hc9	1 hcs9	3 dih9
h	4 hc10	1 hcs10	9 dih10
h	4 hc11	1 hcs11	9 dih11
h	5 hc12	1 hcs12	2 dih12
h	5 hc13	1 hcs13	12 dih13
h	5 hc14	1 hcs14	12 dih14
h	2 hc15	1 hcs15	3 dih15
h	2 hc16	1 hcs16	15 dih16
h	2 hc17	1 hcs17	15 dih17

cs2	1.710000
cs3	1.710000
csc3	109.471
cs4	1.710000
csc4	109.471
dih4	120.000
cs5	1.710000
csc5	109.471
dih5	120.000
hc6	1.089000
hcs6	109.471
dih6	180.000
hc7	1.089000
hcs7	109.471
dih7	120.000
hc8	1.089000
hcs8	109.471
dih8	240.000
hc9	1.089000
hcs9	109.471
dih9	180.000

hc10	**1.089000**
hcs10	**109.471**
dih10	**120.000**
hc11	**1.089000**
hcs11	**109.471**
dih11	**240.000**
hc12	**1.089000**
hcs12	**109.471**
dih12	**180.000**
hc13	**1.089000**
hcs13	**109.471**
dih13	**120.000**
hc14	**1.089000**
hcs14	**109.471**
dih14	**240.000**
hc15	**1.089000**
hcs15	**109.471**
dih15	**180.000**
hc16	**1.089000**
hcs16	**109.471**
dih16	**120.000**
hc17	**1.089000**
hcs17	**109.471**
dih17	**240.000**

If the checkpoint file from the optimization stage for both molecules is available, the information bout the optimal structure could be read directly into the computer RAM (*Geom=Checkpoint Guess=Read*) when the NMR calculations are started. The Gaussian keyword *NMR* is used to perform the calculation of the isotropic shielding tensor (*Calc 4 , Calc. 5),* the combination of the *NMR* keyword with the *SpinSpin* keyword and results in the information about the spin-spin coupling constant. In general, if accurate, the MNR data are required. They require a specific basis set, specially written for the NMR calculation (for example Jensen's basis for chemical shield psJ-2, and for spin-spin coupling constants psS-2).

The combination of B3Lyp DFT theory with 6-31g(d) basis set is a reasonable minimum for a calculation method and basis set.

Calc. 4. Gaussian input file for calculation of isotropic shielding tensor of TMS..

%chk=tms-for-nmr.chk

#b3lyp/6-311g(d,p) NMR *Geom=Checkpoint Guess=Read* SCRF=(PCM,solvent=water)

NMR calculation for

0 1

Calc. 5. Gaussian input file for calculation of isotropic shielding tensor of alanine.

%chk=ala-for-nmr.chk

#b3lyp/6-311g(d,p) NMR *Geom=**Checkpoint Guess=Read*** **SCRF=(PCM,Solvent=water)**

Optimization

0 1

After finishing the calculations, the outputs give the predicted shielding values for each atom in a bout molecule (*Figure 5* and *Figure 6*). *Figure 5* shows the TMS output file. In a red color the reference values for 1H and 13C spectroscopy are marked.

Calculating GIAO nuclear magnetic shielding tensors.
SCF GIAO Magnetic shielding tensor (ppm):
1 Si Isotropic = 339.8796 Anisotropy = 0.0000
XX= 339.8796 YX= 0.0000 ZX= 0.0000
XY= 0.0000 YY= 339.8796 ZY= 0.0000
XZ= 0.0000 YZ= 0.0000 ZZ= 339.8796
Eigenvalues: 339.8796 339.8796 339.8796
2 C Isotropic = **185.2419** Anisotropy = 9.4361
XX= 185.2419 YX= 3.1454 ZX= -3.1454
XY= 3.1454 YY= 185.2419 ZY= -3.1454
XZ= -3.1454 YZ= -3.1454 ZZ= 185.2419
Eigenvalues: 182.0965 182.0965 191.5326

6 H Isotropic = **31.9939** Anisotropy = 9.2815
XX= 32.7927 YX= 4.6351 ZX= -1.8320
XY= 4.6351 YY= 32.7927 ZY= -1.8320
XZ= -1.5939 YZ= -1.5939 ZZ= 30.3965
Eigenvalues: 28.1576 29.6427 38.1816

Figure 5. Selected part of output file for shielding calculation for TMS.

Calculating GIAO nuclear magnetic shielding tensors.
SCF GIAO Magnetic shielding tensor (ppm):
1 N Isotropic = 196.6474 Anisotropy = 45.4089
XX= 194.7910 YX= -8.4598 ZX= -10.0530
XY= -2.8238 YY= 221.7456 ZY= 8.9426
XZ= -10.3472 YZ= 16.4816 ZZ= 173.4057
Eigenvalues: 167.4784 195.5438 226.9200
2 C Isotropic = ***126.5275*** Anisotropy = 43.6977
XX= 138.3503 YX= -18.2855 ZX= -11.1902
XY= -13.2190 YY= 137.6366 ZY= 4.4708
XZ= -7.3726 YZ= 4.9884 ZZ= 103.5955
Eigenvalues: 101.2525 122.6707 155.6592
3 C Isotropic = ***8.6197*** Anisotropy = 88.7876
XX= -62.0118 YX= -18.3465 ZX= -33.8621
XY= -25.3744 YY= 33.1865 ZY= -16.4205
XZ= -34.6700 YZ= -17.5991 ZZ= 54.6845
Eigenvalues: -77.2858 35.3335 67.8115
4 O Isotropic = 9.5278 Anisotropy = 490.6427
XX= -80.6359 YX= 33.1949 ZX= -73.2970
XY= -56.9084 YY= -176.8168 ZY= -142.6333
XZ= -106.8950 YZ= -115.2671 ZZ= 286.0360
Eigenvalues: -219.5782 -88.4614 336.6229
5 O Isotropic = 29.2605 Anisotropy = 401.1931
XX= -79.8481 YX= 52.6623 ZX= -64.9764
XY= 175.5297 YY= -91.6126 ZY= -70.2843
XZ= -27.4542 YZ= -113.8685 ZZ= 259.2422

Eigenvalues: -202.8734 -6.0677 296.7226

6 H Isotropic = ***18.8343*** Anisotropy = 23.0429

XX= 31.9871 YX= 2.6872 ZX= 4.4775

XY= 6.1090 YY= 14.4471 ZY= 1.8388

XZ= 5.8161 YZ= 0.7696 ZZ= 10.0687

Eigenvalues: 8.8991 13.4075 34.1962

7 C Isotropic = ***165.1006*** Anisotropy = 29.3933

XX= 160.6180 YX= 8.0038 ZX= -10.8475

XY= 1.3160 YY= 175.4674 ZY= -8.7308

XZ= -13.1336 YZ= -12.1189 ZZ= 159.2165

Eigenvalues: 147.1010 163.5048 184.6962

8 H Isotropic = ***28.5043*** Anisotropy = 5.7805

XX= 25.2551 YX= -0.9521 ZX= -1.1758

XY= -1.4463 YY= 28.1044 ZY= 0.2616

XZ= 0.0615 YZ= 1.1649 ZZ= 32.1533

Eigenvalues: 24.8067 28.3482 32.3579

9 H Isotropic = ***30.4286*** Anisotropy = 8.8291

XX= 34.7869 YX= 3.0317 ZX= 1.7208

XY= 3.2918 YY= 29.4117 ZY= -0.1797

XZ= 0.3674 YZ= -0.6437 ZZ= 27.0871

Eigenvalues: 26.5531 28.4179 36.3146

10 H Isotropic = ***30.0941*** Anisotropy = 8.6913

XX= 29.1223 YX= -1.0513 ZX= -1.1187

XY= -2.3023 YY= 35.4028 ZY= 0.9833

XZ= -1.0469 YZ= 0.1491 ZZ= 25.7571

Eigenvalues: 25.4382 28.9557 35.8883

11 H Isotropic = ***30.9191*** Anisotropy = 7.5964

XX= 28.1783 YX= 0.9638 ZX= -1.5590

XY= 1.1241 YY= 29.4245 ZY= -1.8408

XZ= -1.5878 YZ= -1.3630 ZZ= 35.1545

Eigenvalues: 27.5275 29.2464 35.9833

12 H Isotropic = ***28.5679*** Anisotropy = 12.6431

XX= 30.1291 YX= -1.8970 ZX= -6.8283

XY= -1.3814 YY= 26.4760 ZY= 1.5069

XZ= -7.1485 YZ= 0.8023 ZZ= 29.0986

Eigenvalues: 22.5863 26.1208 36.9966

13 H Isotropic = ***27.6983*** Anisotropy = 14.8870

XX= 25.0161 YX= -4.3152 ZX= 4.5920

XY= -4.6654 YY= 28.7096 ZY= -5.3453

XZ= 4.9806 YZ= -4.9677 ZZ= 29.3693

Eigenvalues: 21.5995 23.8725 37.6230

Figure 6. Selected part of output file for shielding calculation for alanine**.**

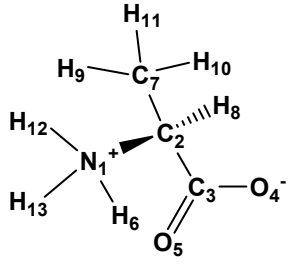


Figure 7. Structure of $_{L}$-alanine with numbered atoms.

To obtain the predicted shift for alanine proton and carbon atoms, its absolute value should be subtracted from the reference atoms of TMS. The values of the predicted chemical shift of alanine are shown in *Table 4*

As it is shown in *Table 4*, the theoretical values of chemical shifts of carbons are shifted in to a higher field in comparison to the experimental data. To minimize the observed dissonances the usage of a better basis set, a mean basis set meant for NMR calculations, Strona: 62 as well as more accurate functional methods

or other *ab initio* methods, like coupled clusters, is recommended. Also, the calculated proton chemical shifts are shifted in comparison to the experimental ones. Another difference between the calculated and experimental proton spectra is a dissonance in chemical shifts of beta protons. In the case of the experimental results all beta protons have the same value, whereas the calculation allows to distinguish between them. To explain this difference we have to remember, that calculations are performed for "static" structures, whereas

experimental values are recorded for the “dynamic” structures, what leads to averaged values of chemical shifts of alanyl methyl protons

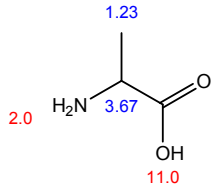
Table 4. Predicted and experimental chemical shifts for alanine.

Atom	**Shielding**	**Shielding of reference (*TMS)***	**Predicted chemical shift B3Lyp/6-311g(d,p)**	**Experimental chemical shift***
C2	126,5275	185,2419	58,714	51,32
C3	8,6197	185,2419	176,622	176,59
H6	27,8892	31,9941	3,7749	not observed
C7	165,1006	185,2419	20,141	17,02
H8	28,5043	31,9939	3,490	3,786
H9	30,4286	31,9941	1,565	1,482
H10	30,0941	31,9941	1,900	1,482
H11	30,9191	31,9942	1,075	1,482
H12	28,5679	31,9941	3,426	not observed
H13	27,6983	31,9941	4,296	not observed

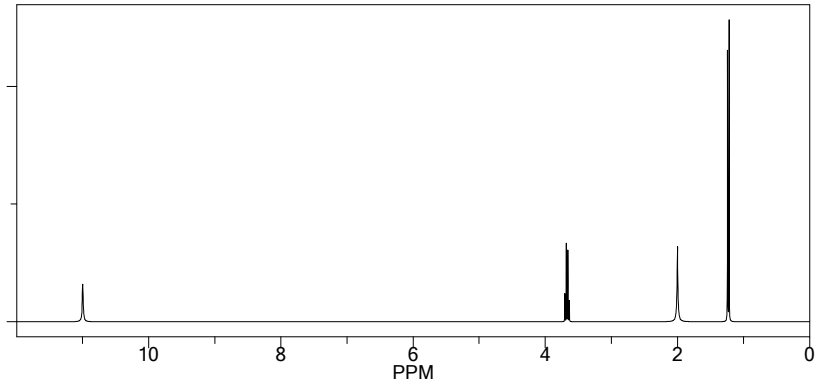
*- spectra recorded in D_2O, SDBSWeb : http://riodb01.ibase.aist.go.jp/sdbs/

Empirical methods

The empirical prediction of the chemical shifts of alanine was performed with using ChemDraw Ultra. This program required only a 2D structures as an input . The chemical shift prediction was based on the additive character of it. Below the results obtained from **ChemDraw** software are presented (*Figure 8* and 9). The outputs contain the information about the manner of calculation.



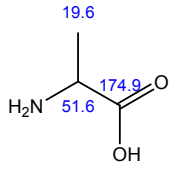
Estimation Quality: blue = good, magenta = medium, red = rough



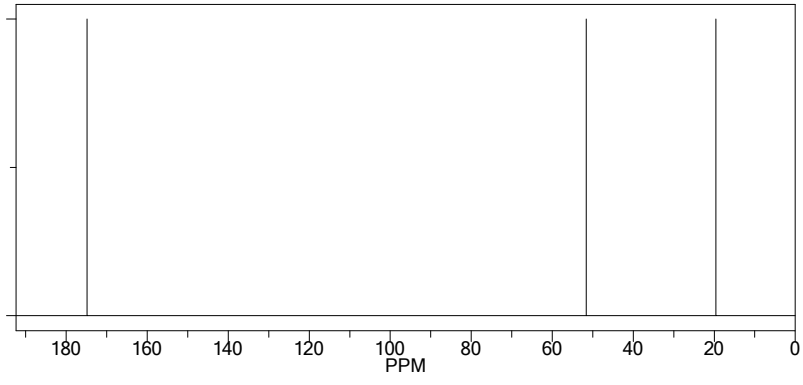
Protocol of the H-1 NMR Prediction:

Node	Shift	Base + Inc.	Comment (ppm rel. to TMS)
NH2	2.0	2.00	amine
CH	3.67	1.50	methine
		0.17	1 alpha -C
		1.13	1 alpha -N
		0.87	1 alpha -C(=O)O
CH3	1.23	0.86	methyl
		0.14	1 beta -N
		0.23	1 beta -C(=O)O
OH	11.0	11.00	carboxylic acid

Figure 8. Predicted 1H spectra based on the empirical methods.



Estimation Quality: blue = good, magenta = medium, red = rough

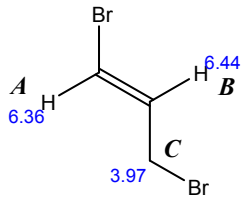



```
Protocol of the C-13 NMR Prediction:

Node      Shift    Base + Inc.   Comment (ppm rel. to TMS)

  CH       51.6         -2.3      aliphatic
                        21.8      1 alpha -C(=O)-O
                         9.1      1 alpha -C
                        28.3      1 alpha -N
                        -5.3      general corrections
 CH3       19.6         -2.3      aliphatic
                         9.1      1 alpha -C
                         2.0      1 beta -C(=O)-O
                        11.3      1 beta -N
                        -0.5      general corrections
   C      174.9        166.0      1-carboxyl
                        11.0      1 -C-C
                        -2.1      general corrections
```

Figure 9. Predicted 13C spectra based on the empirical methods.

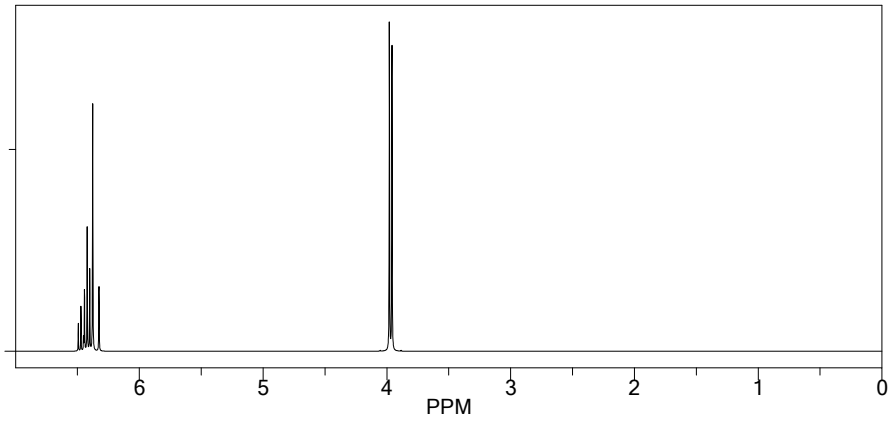
The comparison of the predicted shifts with the experimental ones (*Table 4*) shown a pretty good agreement, especially when taking into consideration that the estimation process, including structure drawing, takes not more than 1-2 minutes. Of course the rational application of the empirical methods is limited to molecules which are similar to these which

served to parameterize a algoritm. Additionally, the empirical methods cannot very accurately distinguish the influences of different types of isomers onto the observed chemical shifts, what is shown in the example of *(E)*-1,3-dibromoprop-1-ene and *(Z)*-1,3-dibromoprop-1-ene (*Figure 10* and *11*)



(*E*)-1,3-dibromoprop-1-ene

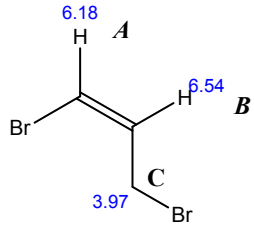
Estimation Quality: blue = good, magenta = medium, red = rough



Protocol of the H-1 NMR Prediction:

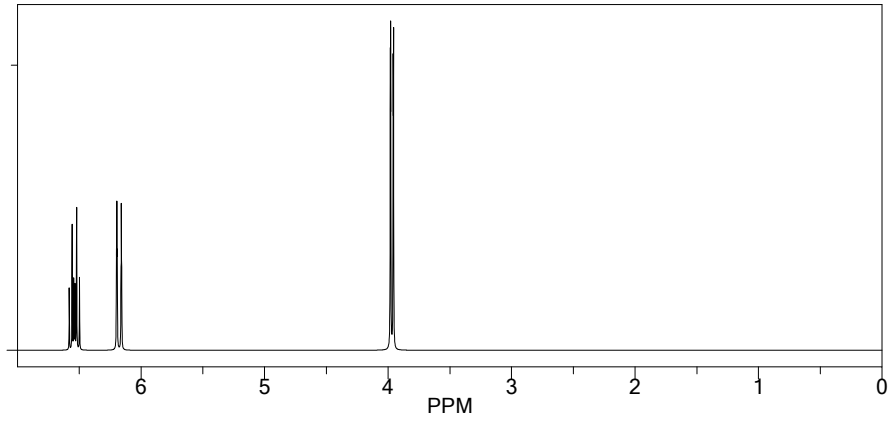
Node	Shift	Base + Inc.	Comment (ppm rel. to TMS)
H	6.36	5.25	1-ethylene
		1.07	1 -Br gem
		0.04	1 -C-Br cis
H	6.44	5.25	1-ethylene
		0.45	1 -Br cis
		0.74	1 -C-Br gem
CH2	3.97	1.37	methylene
		0.63	1 alpha -C=C
		1.97	1 alpha -Br

Figure 10. Esitmation of proton chemical Shift for *(E)*-1,3-dibromoprop-1-ene with empirical method.



(*Z*)-1,3-dibromoprop-1-ene

Estimation Quality: blue = good, magenta = medium, red = rough



Protocol of the H-1 NMR Prediction:

Node	Shift	Base + Inc.	Comment (ppm rel. to TMS)
H	6.54	5.25	1-ethylene
		0.55	1 -Br trans
		0.74	1 -C-Br gem
CH2	3.97	1.37	methylene
		0.63	1 alpha -C=C
		1.97	1 alpha -Br
H	6.18	5.25	1-ethylene
		1.07	1 -Br gem
		-0.14	1 -C-Br trans

Figure 11. Esitmation of proton chemical Shift for *(Z)*-1,3-dibromoprop-1-ene with empirical method.

Table 5. Comparison of experimental and predicted proton chemical shifts for *E* and *Z* isomers of 1,3-dibromoprop-1-ene

Proton	***Isomer (E)***		***Isomer (Z)***	
	Empirical	*Experimental**	*Empirical*	*Experimental**
A	*6.36*	*6.43*	*6.18*	*6.34*
B	*6.44*	*6.37*	*6.54*	*6.39*
C	*3.97*	*3.88*	*3.97*	*4.03*

*- spectra recorded in D_2O, SDBSWeb : http://riodb01.ibase.aist.go.jp/sdbs/

4.1.3 Theoretical calculation of the VCD spectra.

Vibrational Circular Dichroism (VCD) is exhibited by chiral molecules. VCD spectroscopy could be applied to structural and conformational investigations of numerous groups of compounds. This technique is often used in peptides and proteins investigations, because it allows to reveal a periodic secondary structure, such as α-helix and β-structure and, moreover, it also allows to evaluate their relative contributions[xvi]. This method is very useful for valuation of the predicted conformation[xvii]

Theoretical VCD intensities are determined by vibrational strengths, which are determined by vibrational electric and magnetic dipole transition moments.[xviii] A theoretical VCD spectrum is a graph of rotatory strengths and a wave numbers relation.

The calculation of the VCD spectra is possible at the *ab initio* level with using Hartree-Fock and DFT methods. This methodology is applied inter alia in the Gaussian software. Strona: 69

Figure 12 shows a comparison

of the experimental and theoretical VCD spectra for cyclo-(-Ile-Asp-Ser-*trans*-Achc-Leu-Asn-), where *trans*-Achc is (1R,2R)-2-aminocyclohexane carboxylic acid[16].

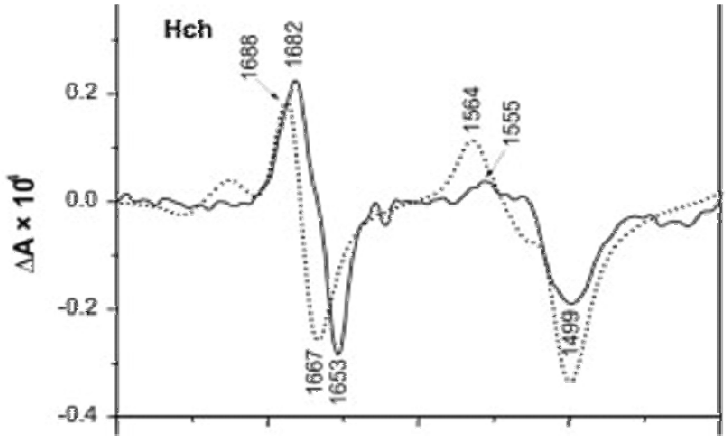


Figure 12. Comparison of theoretical (dotted) and experimental (solid) VCD spectra for cyclo-(-Ile-Asp-Ser-*trans*-Achc-Leu-Asn-)

4.2 Application of spectroscopic data to structural and conformational calculation.

Today developed molecular modeling techniques together with bigger computer capabilities, give a possibility to resort numerous experimental data to calculate a structure and conformation of the molecules characterized by higher MW. Those molecular modeling techniques are based on molecular mechanics and dynamics methods. Empirical information which is used during a molecular modeling stem from such experimental techniques like crystallography, nuclear magnetic resonance spectroscopy or infrared spectroscopy.

From the point of view of this script the most important is using NMR and IR techniques.

The most important structural information came from NMR spectroscopy when nuclear Overhauser effect was used, like 2D NOESY and ROESY experiments. The measurement of the diagonal and cross peaks intensity *a* allows to calculate inter atomic distances. Their calculation is possible on the basis of relation *Eq. 2* More detailed information about the calculation of interatomic distances could be found in Ammalahti et all paper[xix]

$$a_{ij} = 1/r_{ij}^{6}$$

Eq. 2

Another important experimental pieces of information refer to dihedral angles calculated from vicinal coupling constants[xx], values of chemical shifts (CSA), orientation of molecular parts in

proportion to the external magnetic field estimated from the residual dipolar coupling (RDC), the temperature factor of amid protons.

An example of a trategy using the combined molecular modeling methods and experimental data is shown in *Figure 13.*

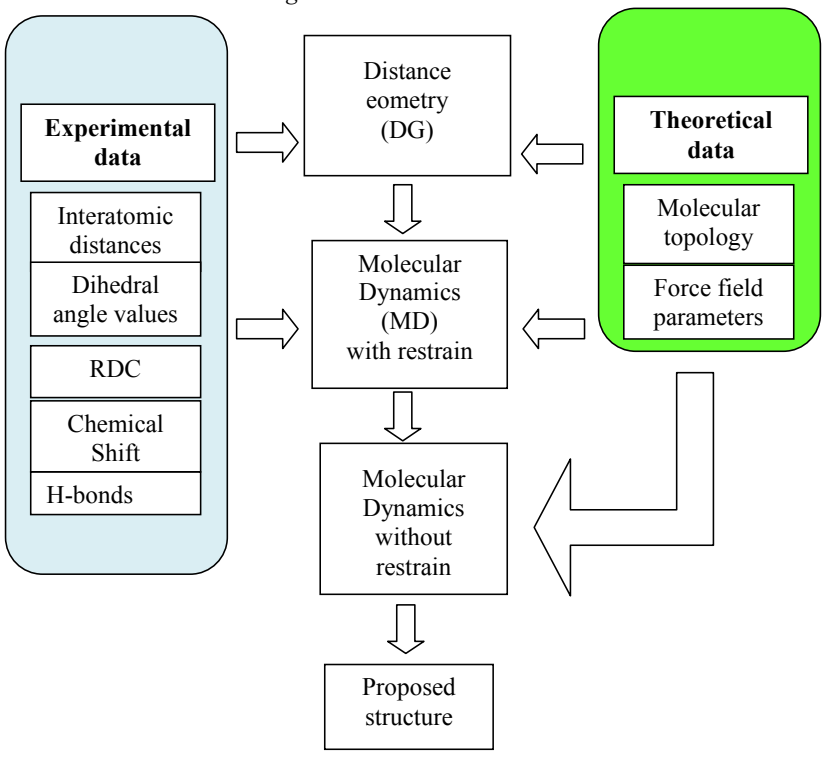


Figure 13. Example of calculation strategy using experimental information.

The structure of the third SH3 domain of CD2AP could serve as an example of using a similar calculation strategy (*Figure 14*) [xxi]

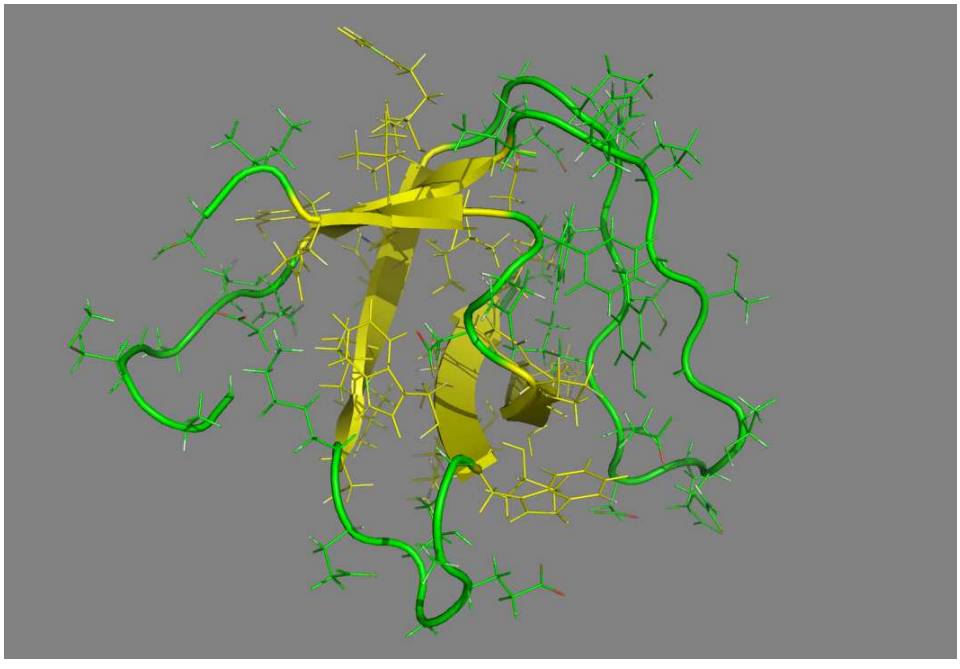


Figure 14. RDC refined high resolution structure of the third SH3 domain of CD2AP

Proposed literature:

1. David C. Young, "Computational chemisty", ISBN: 0-471-33368-9
2. Kenny B. Lipkowitz, Donald B. Boyd, „Reviews in computational chemistry volume18", ISBN: 0-471-21576-7
3. Leach A. R., "Molecular Modeling – principles and applications", ISBN 0-582-38210-6
4. Wuthrich K, "NMR of proteins and nucleic acids" ISBN:0-471-82893-9

References

[i] Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM Jr, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules". *J. Am. Chem. Soc.* 1995, vol. **117**, 5179–519

[ii] van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. Biomolecular Simulation: The GROMOS96 Manual and User Guide; vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.v.: Zürich, Groningen, 1996

[iii] The GROMOS Biomolecular Simulation Program Package, W.R.P. Scott, P.H. Huenenberger, I.G. Tironi, A.E. Mark, S.R. Billeter, J. Fennen, A.E. Torda, T. Huber, P. Krueger and W.F. van Gunsteren. *J. Phys. Chem. A*, vol. **103**,3596-3607

[iv] MacKerell, Jr. AD, *et al.* "All-atom empirical potential for molecular modeling and dynamics studies of proteins". *J Phys Chem B,* 1998, vol. **102**: 3586–3616

[v] Schwieters C.D, Kuszewski J.J., Tjandra N. and Clore G.M., "The Xplor-NIH NMR Molecular Structure Determination Package," *J. Magn. Res.*, 2003, vol. **160**, 66-74

[vi] www.gaussian.com

[vii] http://www.msg.chem.iastate.edu/index.html

[viii] http://www.turbomole-gmbh.com/

[ix] http://www.q-chem.com/

[x] http://www.gaussian.com/g_tech/g_ur/m_basis_sets.htm

[xi] Alkorta I., Elguero J., "Review on DFT and *ab initio* Calculations of Scalar Coupling Constants", *Int. J. Mol. Sci.* 2003,**vol 4**, 64-92

[xii] http://gaussian.com/g_prod/gv5.htm

[xiii] http://www.cmbi.ru.nl/molden/

[xiv] http://gausssum.sourceforge.net/

[xv] http://www.hyper.com

[xvi] Keiderling T.A, „Protein and peptide secondary structure and conformational determination with vibrational circular dichroism", *Curr. Opin. Chem. Biol.,* 2002, **vol. 6,** 682-688

[xvii] Vass E,Stijowski U, Wollschlager K, Mandity I.M, Szilvagyi G, Jewginski M, Gaus K, Royo S, Majer Z, Sewald N, Hollosi M, *J. Pept. Sci.*, 2010, **vol. 16**, 613-620

[xviii] Cheeseman J.R, Frish M.J, Devlin F.J, Stephens P.J, „*Ab initio* calculation of atomic axialy tensors and vibrational rotational strengths using density functional theory", *Chemical Physics Letters*, 1996, **vol. 252**, 211-220

[xix] Ammalahti E, Bardet M, Molko D, Cadet J, „Evaluation of distances from ROESY experiments with the intensity-ratio method", *J. Magn. Res. A*, 1996, **vol. 122**, 230-232

[xx] Contreras R.H, Peralta J,E, "Angular dependence of spin-spin coupling constants.", *Progress in Nuclear Magnetic Resonance Spectroscopy*, 2000, **vol 37**, 321-425

[xxi] http://www.pdb.org/pdb/explore/literature.do?structureId=2KRO