What can NMR do?

- Small molecules
- Macromolecules
  - 3D Structure
    - In solution
    - In cell
    - In membranes
  - Molecular motion
    - Fast internal motion
    - Sparsely populated states
- Interactions
    - Both transient & tight interactions
    - Residue specific interaction mapping
- Mixtures
    - Metabolomics
    - Compound screening

Biomolecular interactions by NMR

K_D	mM	μM	nM	pM
pK_D	3	6	9	12

Typical biochemical interactions

Marketed drugs

Fragments

Practical detection limit

Detection range ligand-observe NMR

Detection range protein-observe NMR

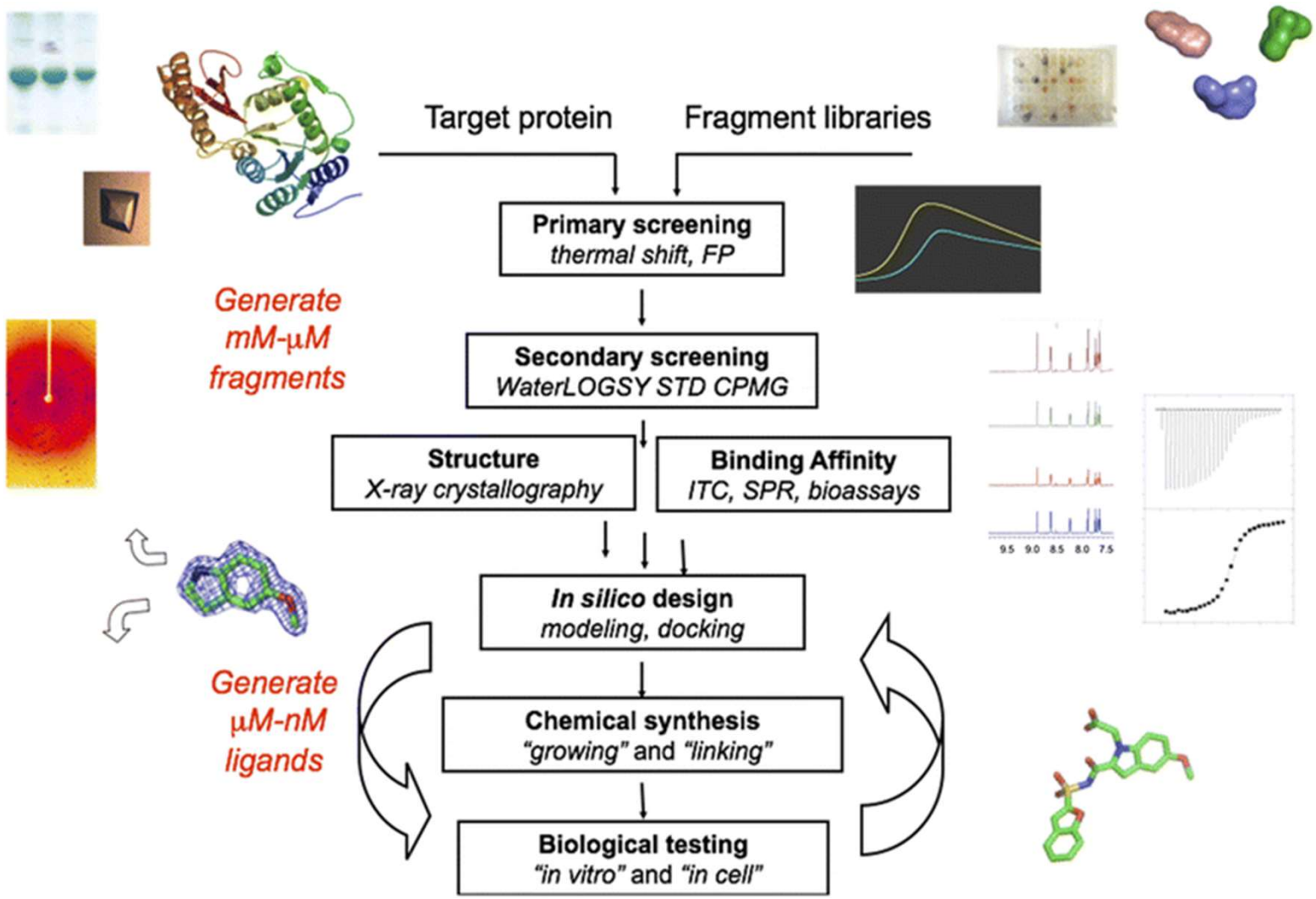
Focusing on small molecular ligands

Purpose

- High-throughput screening
- Lead generation – discovering binders
- Lead optimisation – increasing the strength of the binding
- Epitope mapping – which part(s) of the ligand are important for complex formation?

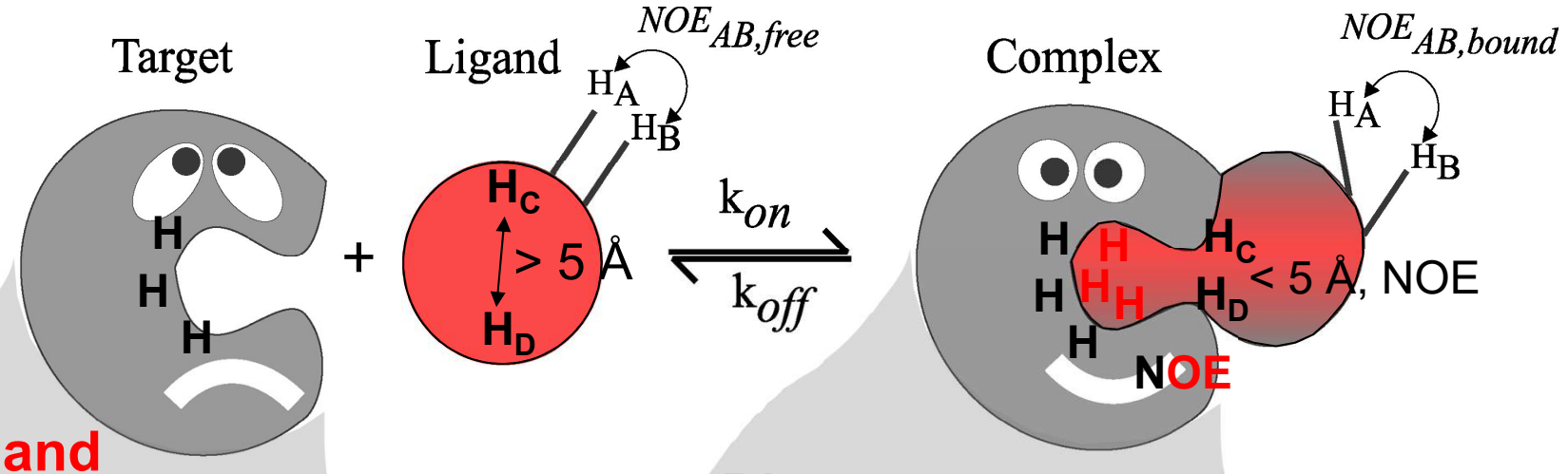
1) 1D NMR experiments typically used – usually fast
2) Experiments are more efficient with larger proteins
3) Most discussed methods are more efficient for weaker binders (K_D mM to μM)

Typical fragment based drug discovery process



Abell & Dagostin, in *Fragment-Based Drug Discovery*, 2015, pp. 1-18

Differences between large and small molecules



Free ligand

- 100-1000 D
- short rotational correlation times (τ_c)
- long relaxation times (T_1, T_2)
- fast translational diffusion
- weak positive NOE

Biomacromolecule

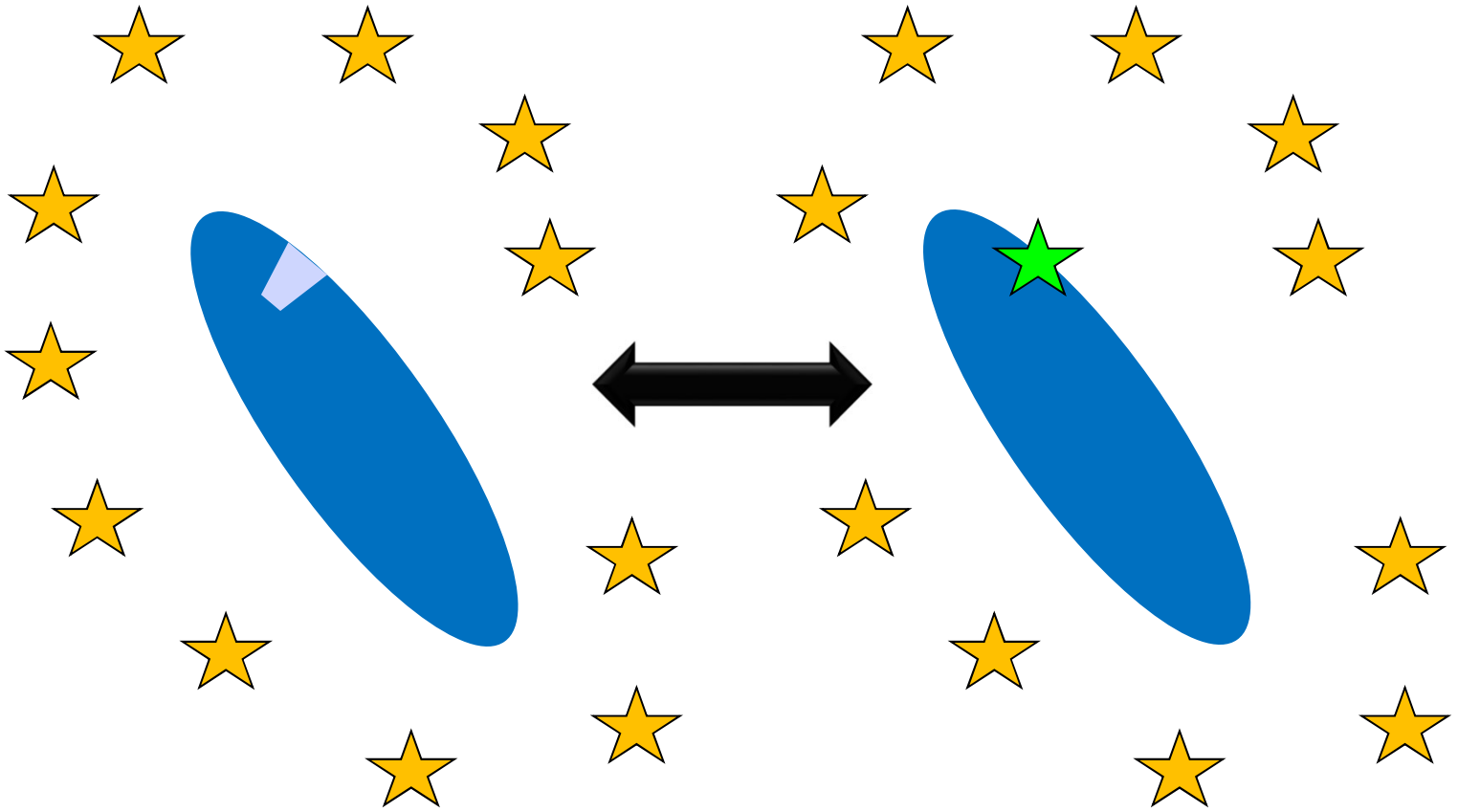
- > 5 kD (no upper limit)
- long rotational correlation times i.e. slow tumbling
- short relaxation times (T_2)
- slow translational diffusion
- strong negative NOE

How is binding detected?

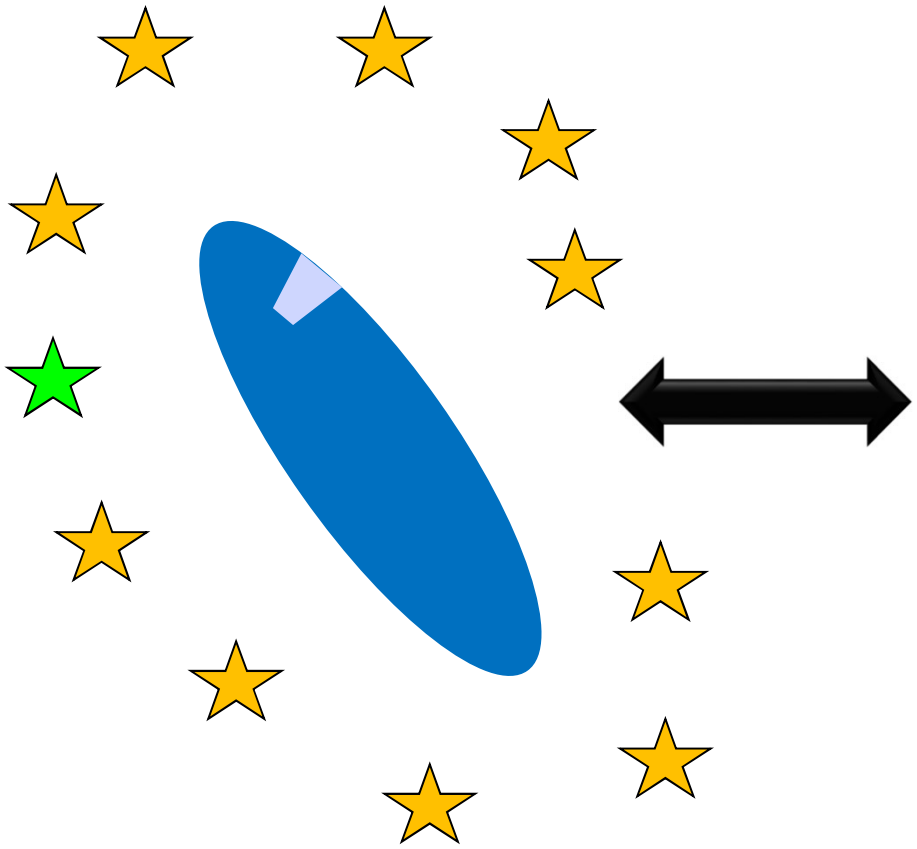
A. A small molecule behaving like a large molecule

B. Transfer of magnetisation from a macromolecule to a ligand

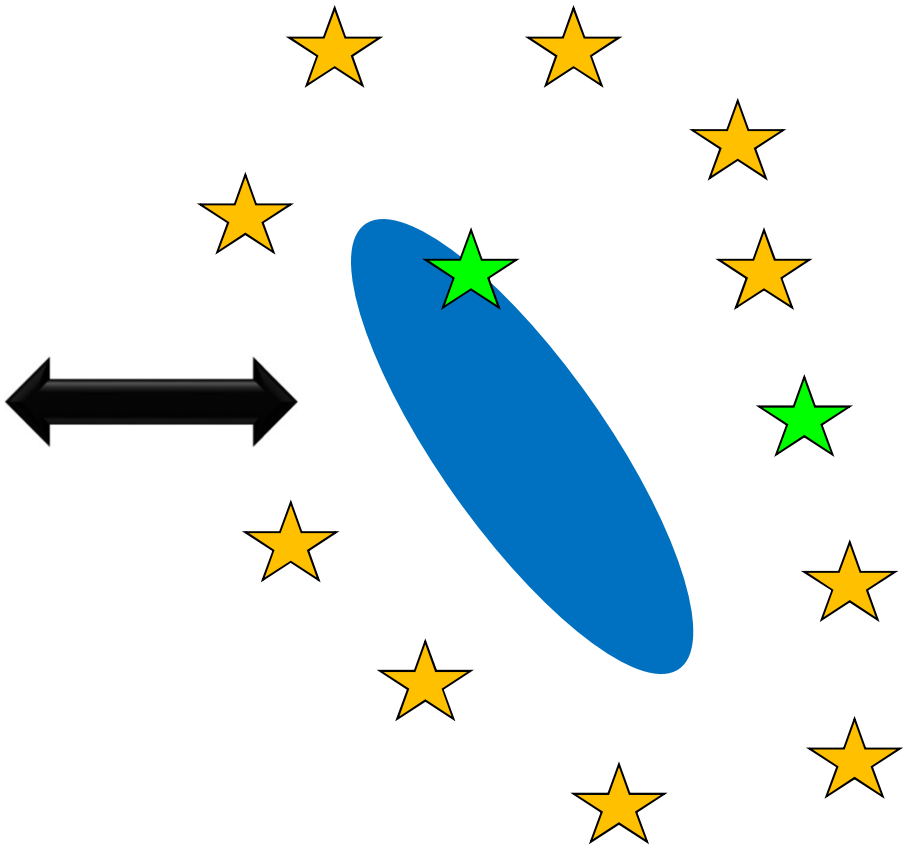
Ligand observe



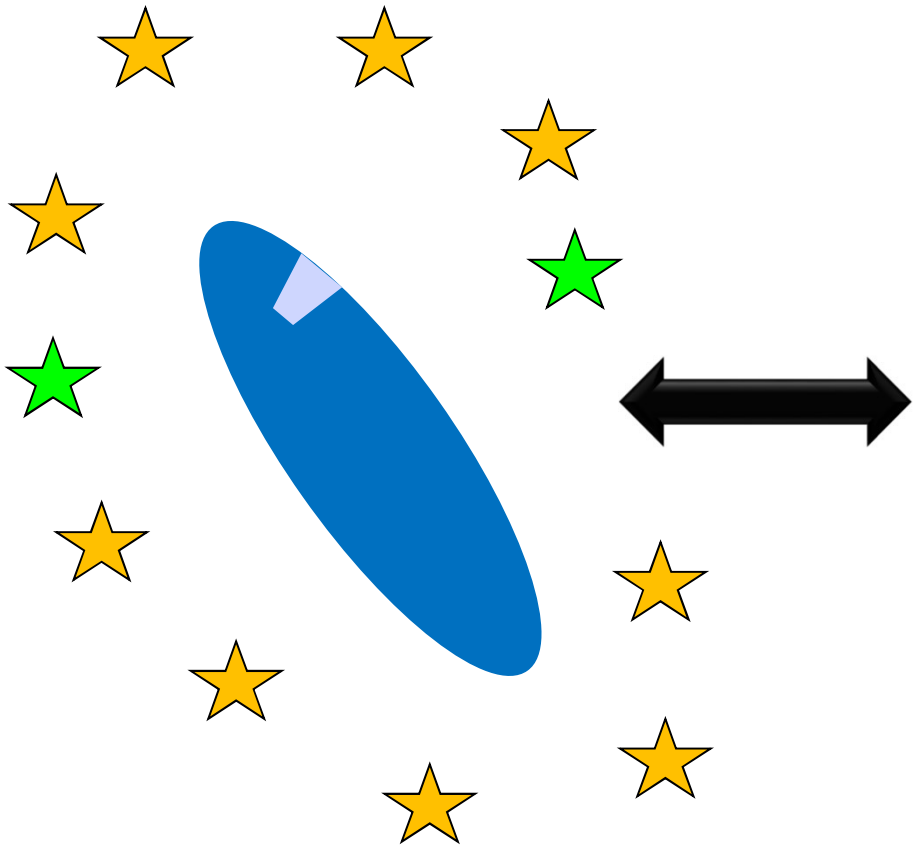
Ligand observe



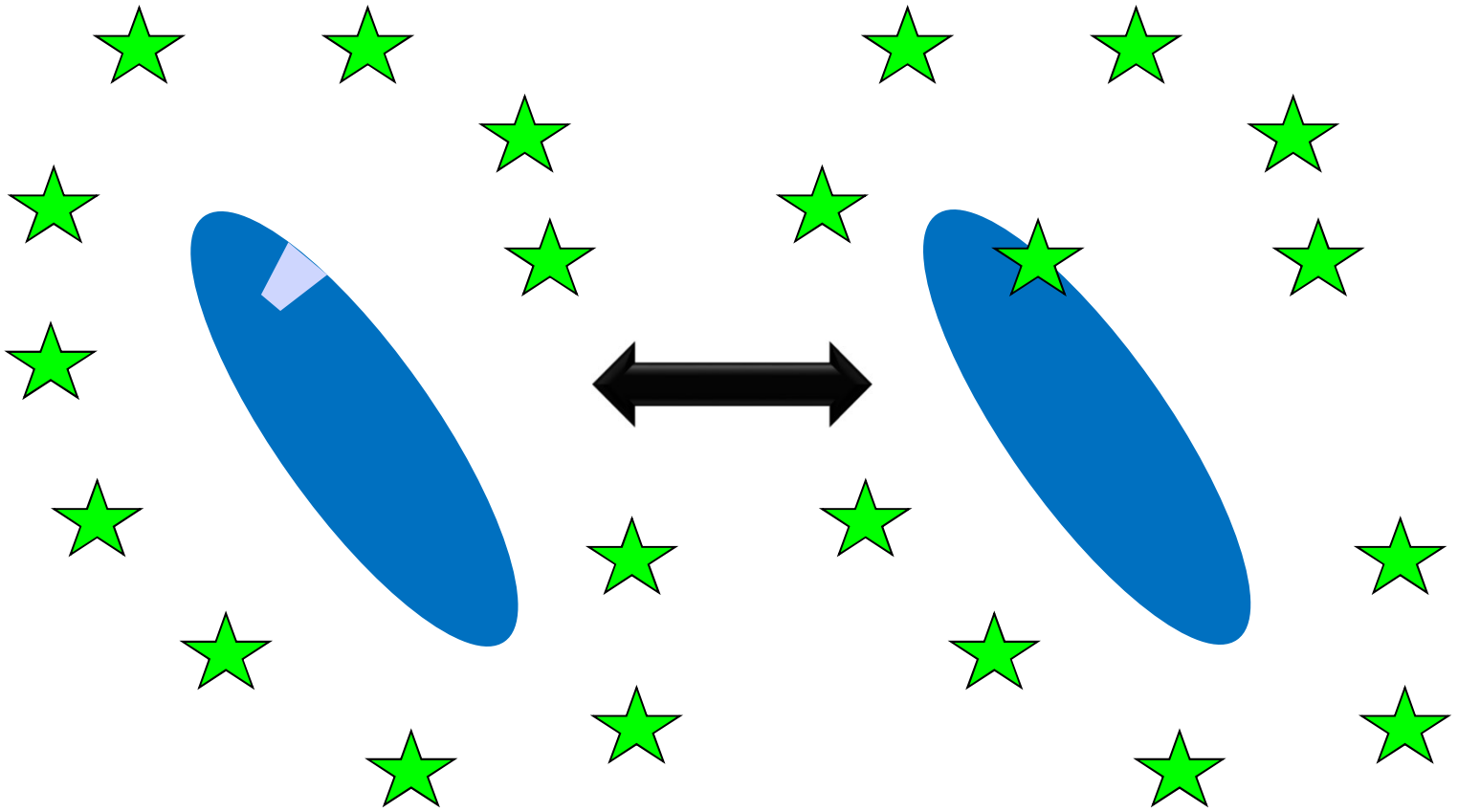
Ligand observe



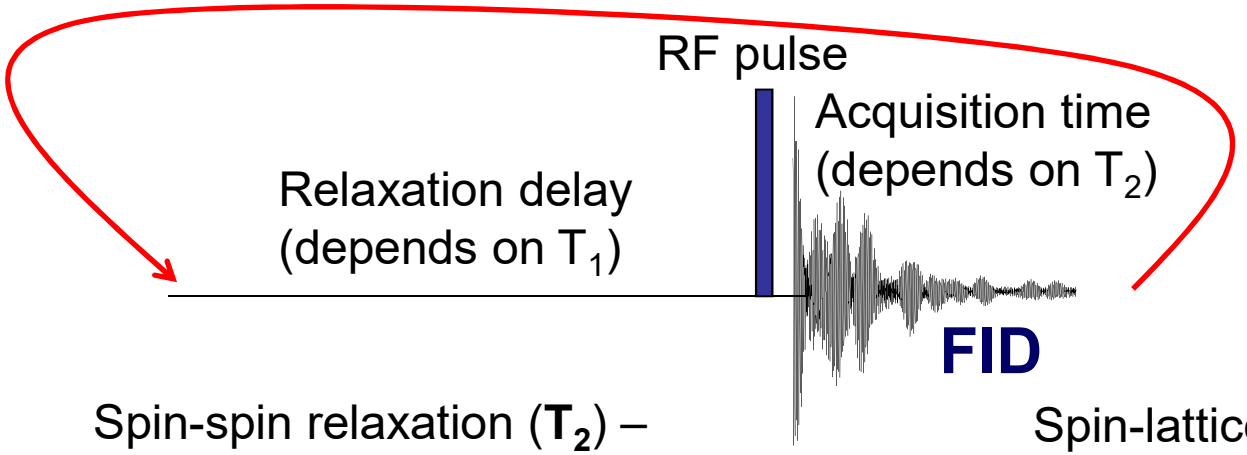
Ligand observe



Ligand observe



Relaxation of spins and NMR experiments

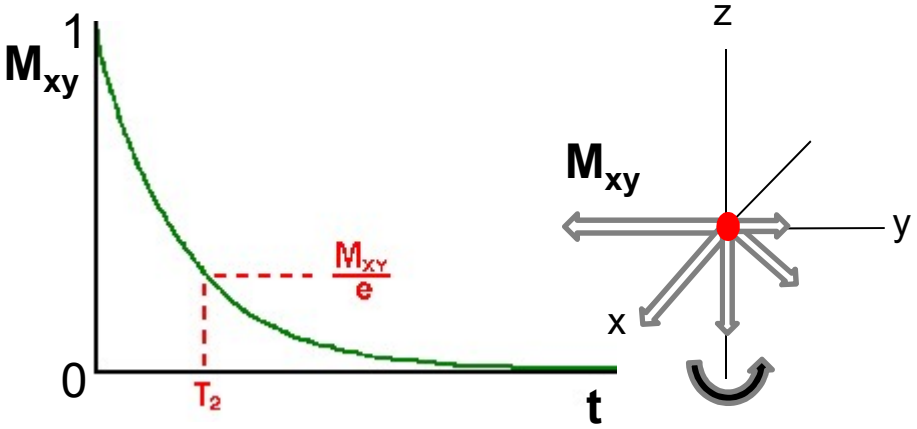


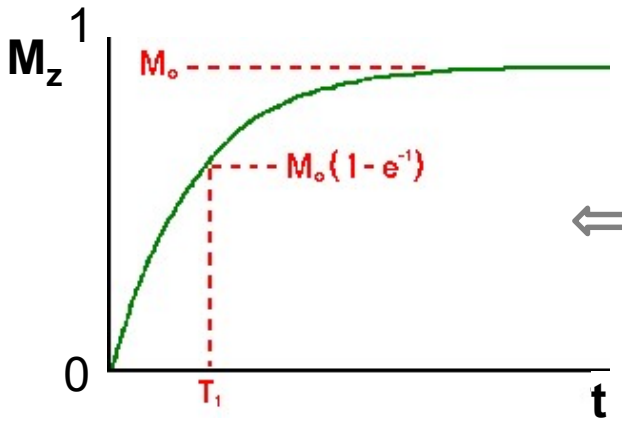
After an RF pulse, signal is acquired in the xy plane.

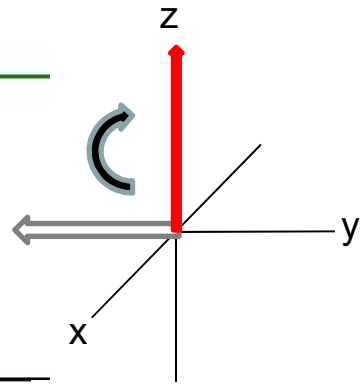
We need to wait before next scan for the magnetisation to return back to equilibrium

Spin-spin relaxation (**T_2**) – disappearance of the signal from the *xy* plane.

Spin-lattice relaxation (**T_1**) – return of the spins to the *z* axis.

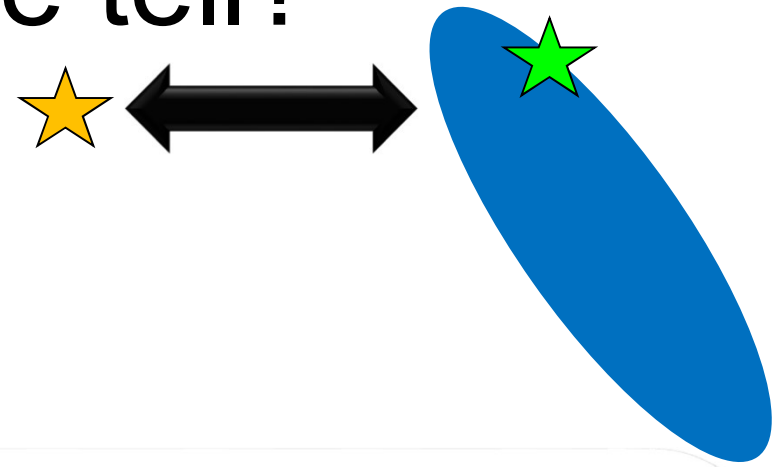




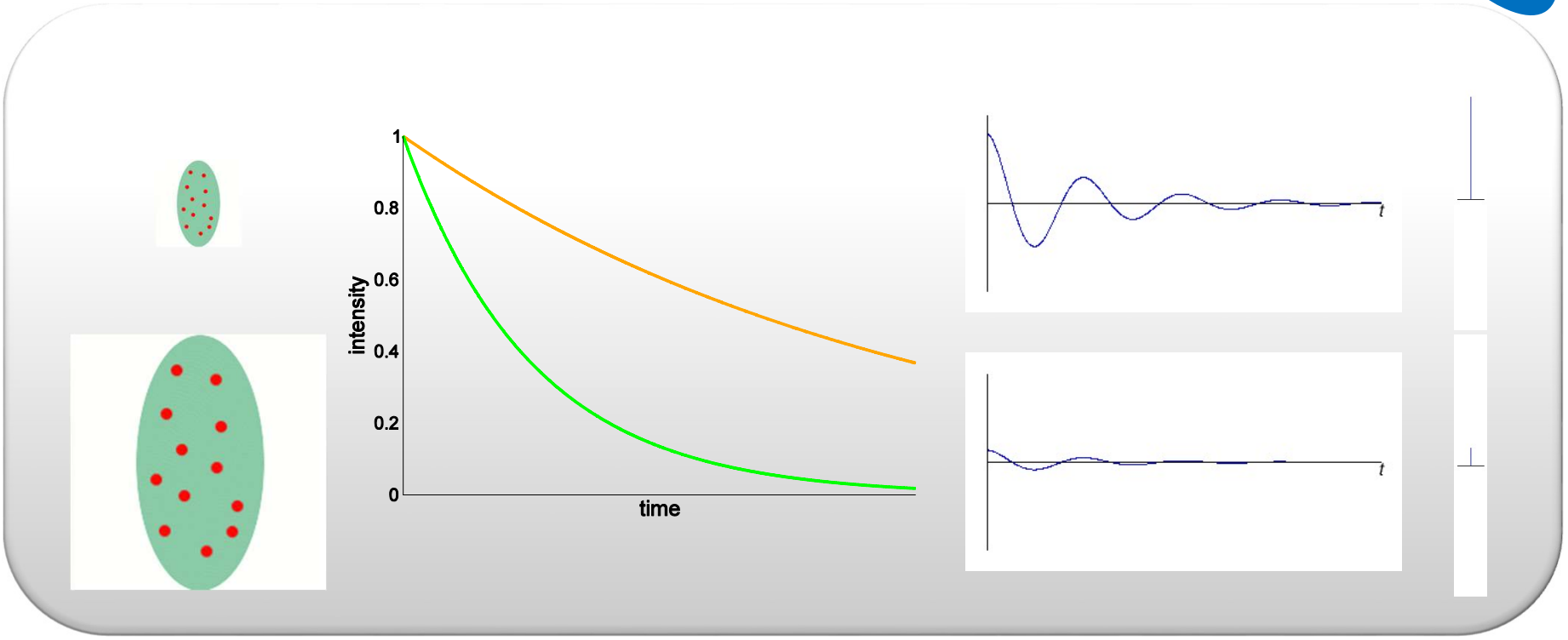


1. Two key relaxation times, T_1 and T_2. T_1 is longer than or equal to T_2.
2. For small molecules T_1 and T_2 are comparable and of order of seconds!
3. Large molecules have short T_2 relaxation (ms) and long T_1 (sec).

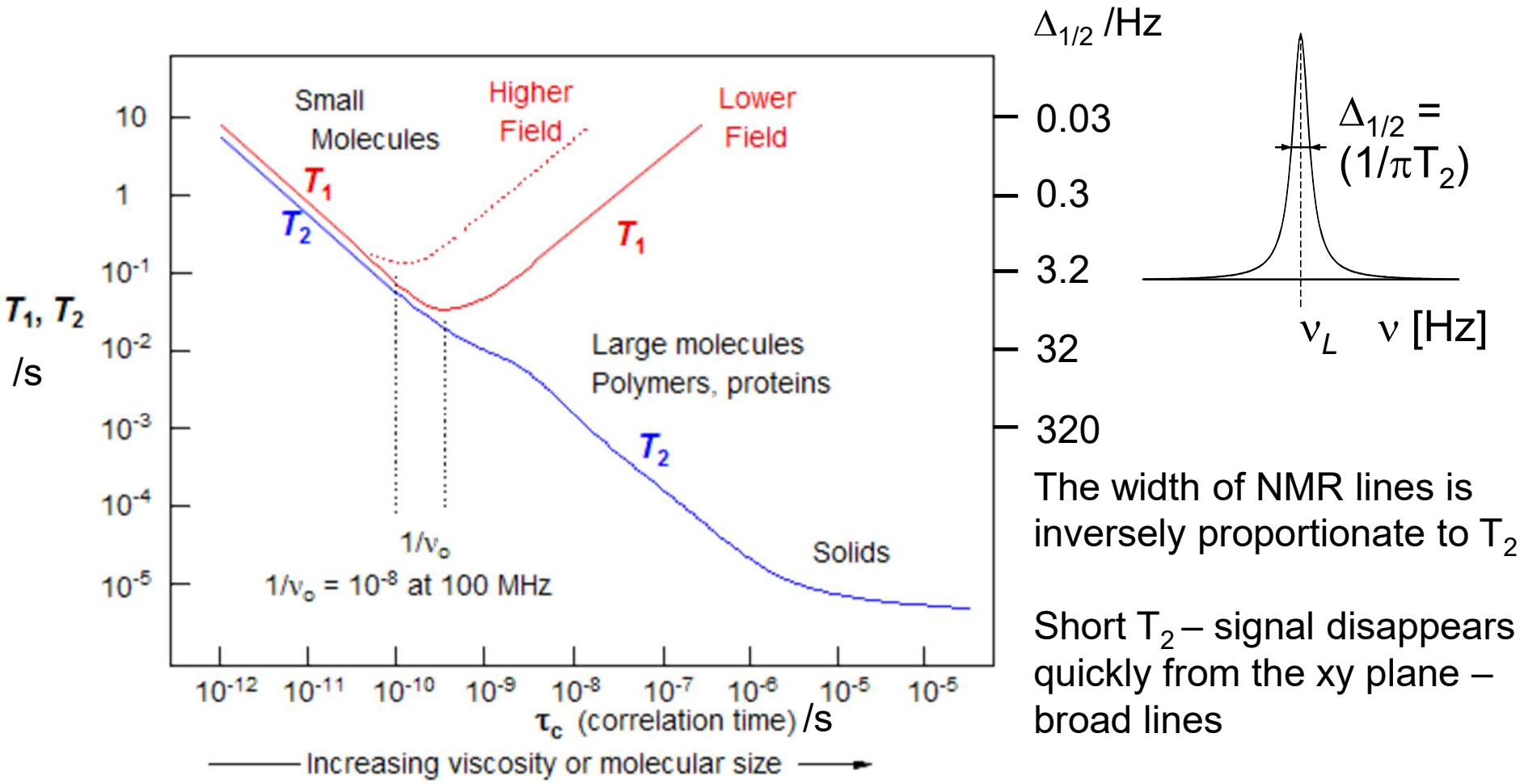
How can we tell?



- Relaxation
  - the T2 (CPMG) filter



Relaxation of spins in small and large molecules

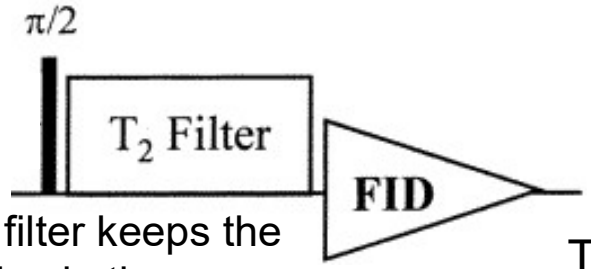


The width of NMR lines is inversely proportionate to T_2

Short T_2 – signal disappears quickly from the xy plane – broad lines

Differences in the T_2 relaxation between the large and small molecules can be exploited to detect binding

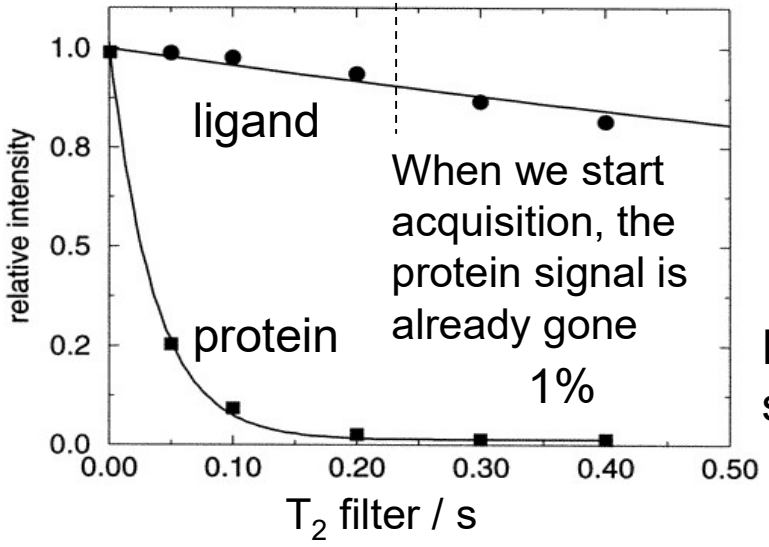
T_2 (spin-spin relaxation) edited spectra



T_2 filter keeps the spins in the transverse plane, where they relax

T_2 filter, small molecules only - all signals preserved

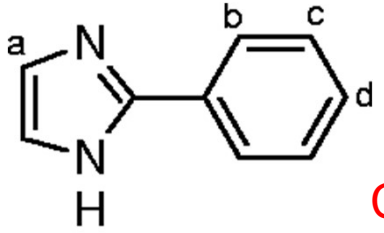
T_2 filter, protein + small molecules – some signals lost



A-B, binder visible

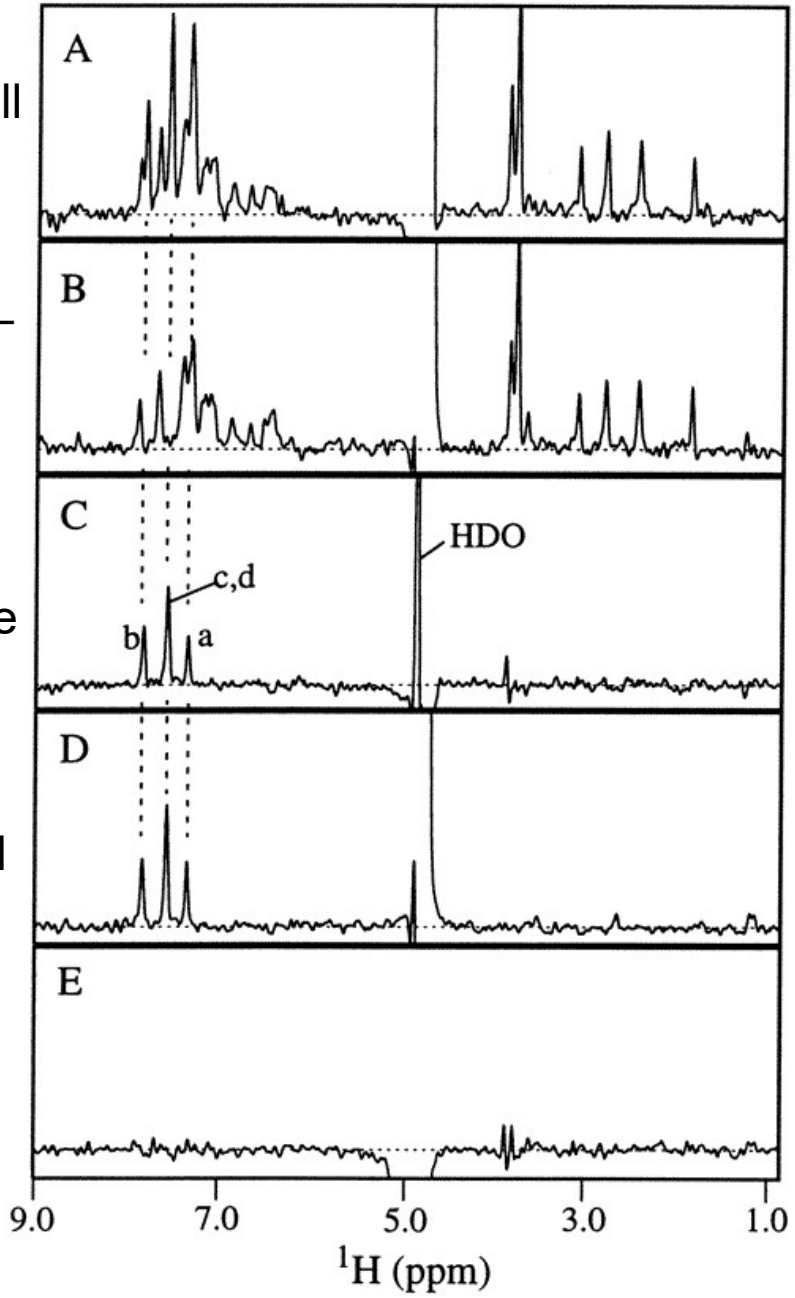
Reference spectrum of **1**

As C but **1** omitted from the mixture



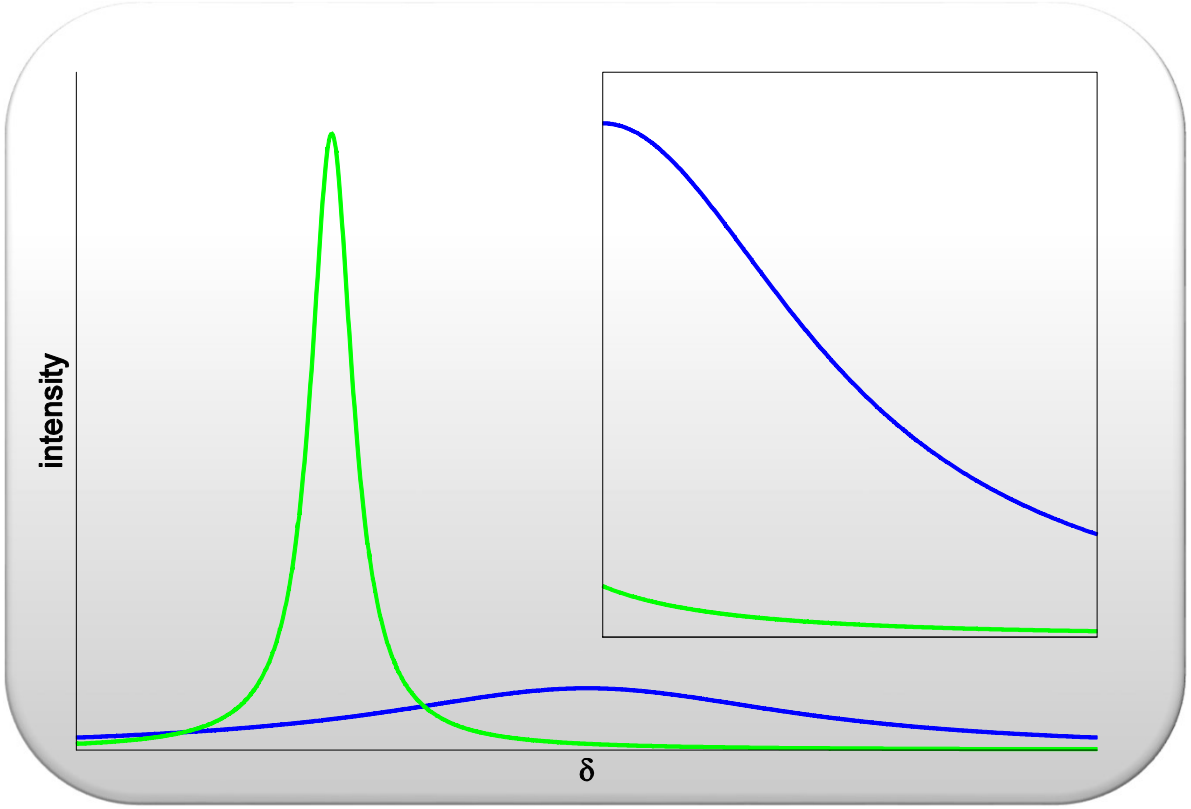
1

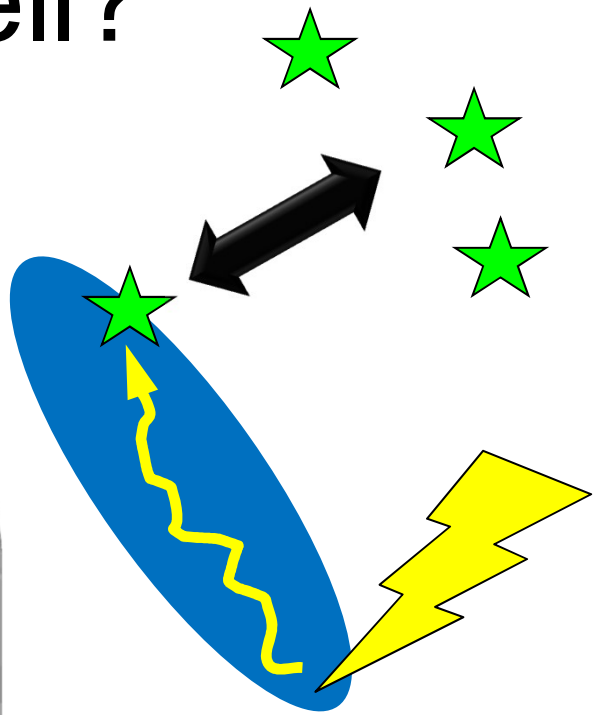
Compound **1** is a binder, in the presence of protein has short T_2



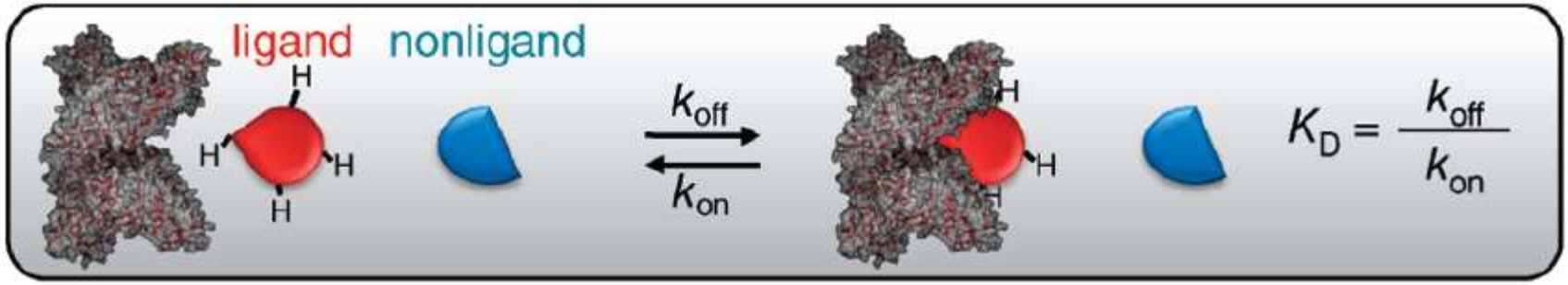
How can we tell?

- Cross-relaxation from protein
  - Saturation Transfer Difference

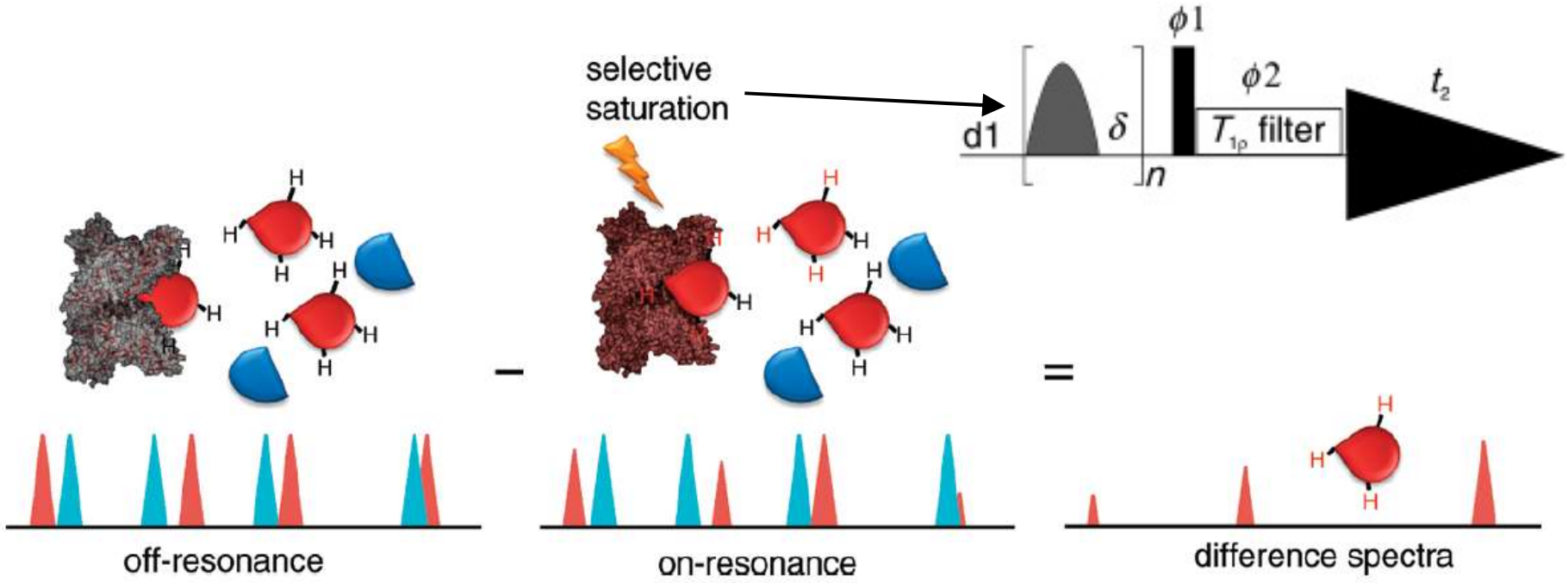




Saturation transfer difference (STD)



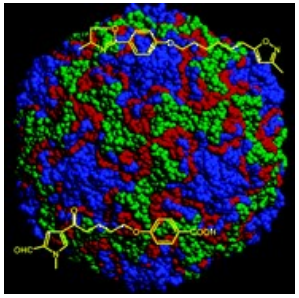
*J.Chem.Educ.*2011, **88**, 990−994



- When bound, saturation of protein protons is partially transferred to the ligand's protons
- The intensity of the ligand's protons, **detected in the free state**, is decreased
- The difference spectrum contains only signals of the binding ligand

STD – epitope mapping

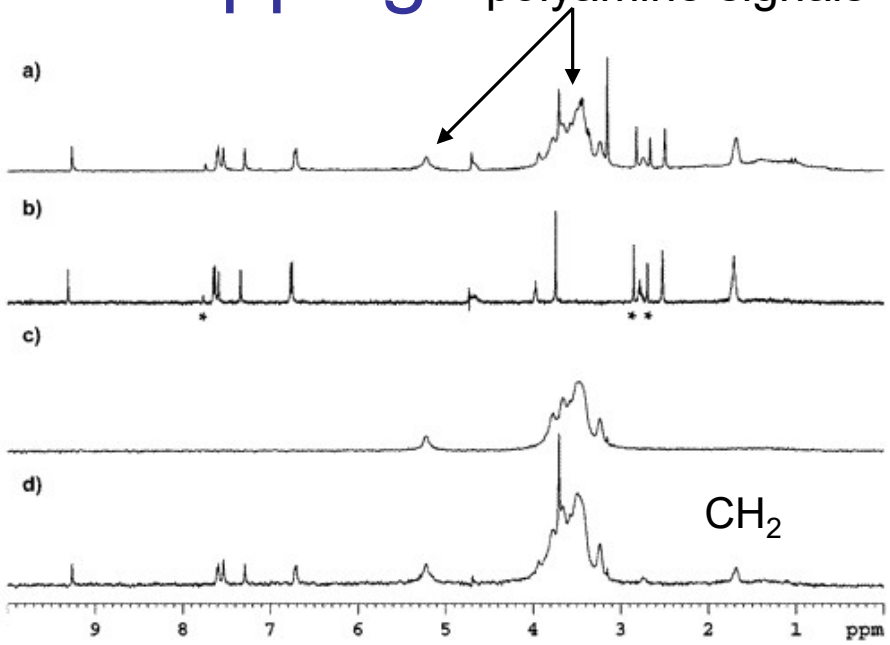
(a) Reference spectrum, compound **1**, HRV2 and polyamine counter ions)
(b) A 1D NMR spectrum of compound **1** alone
(c) STD spectrum for HRV2 in the absence of compound **1** (signals of polyamines that serve as counter ions for the RNA are visible)
(d) STD spectrum, for compound **1** (120 μM) with HRV2 (20 nM) at a molar ratio of 6000:1 (100:1 over the binding sites)

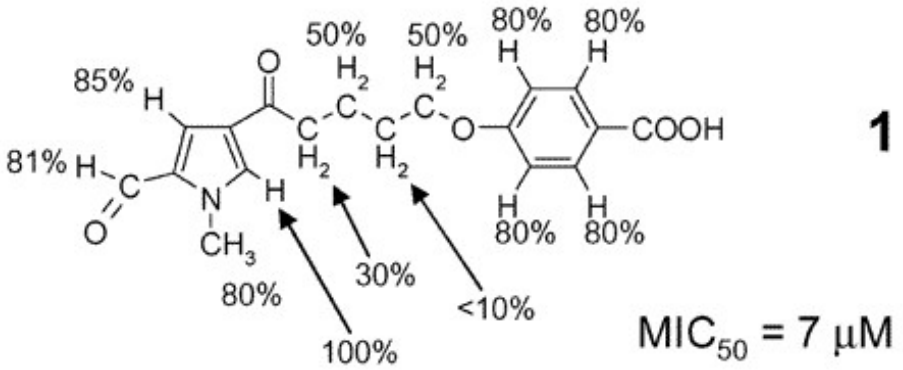


HRV2

Relative enhancements of individual signals in **1** indicates which protons are in a closest contact with the HRV2

J. Am. Chem. Soc., **125** (1), 14 -15, 2003

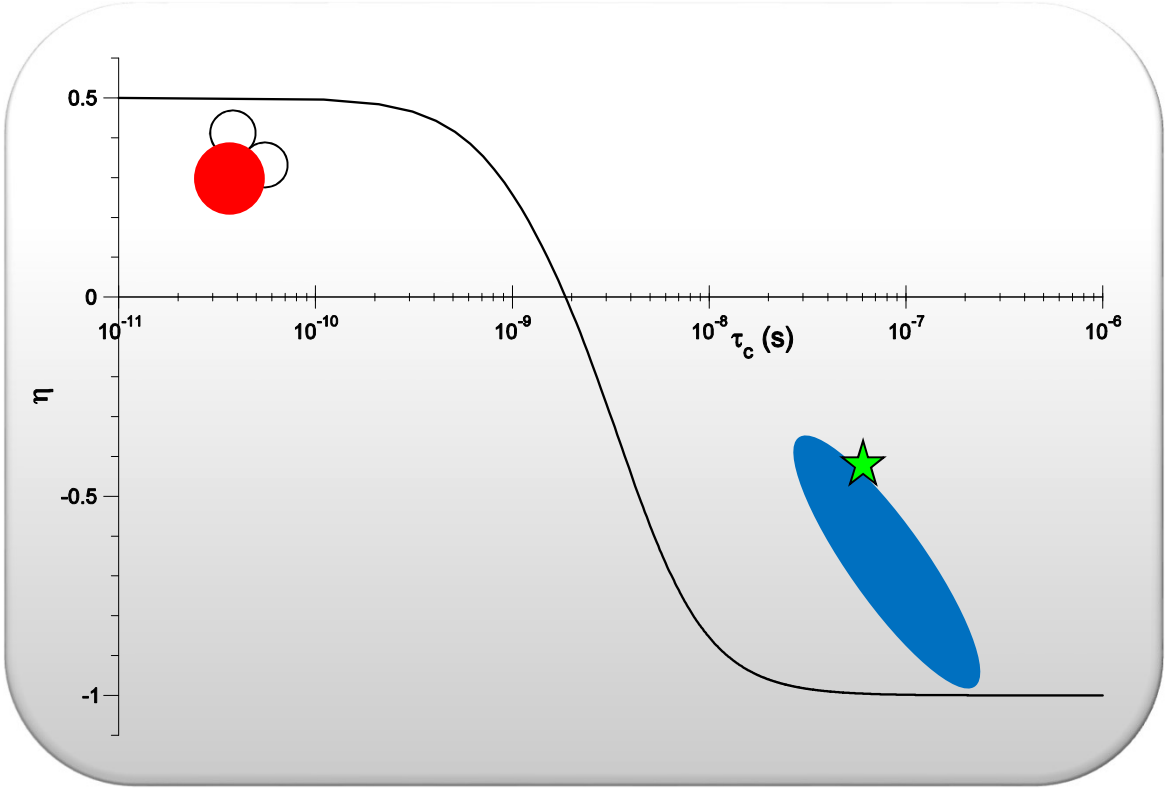


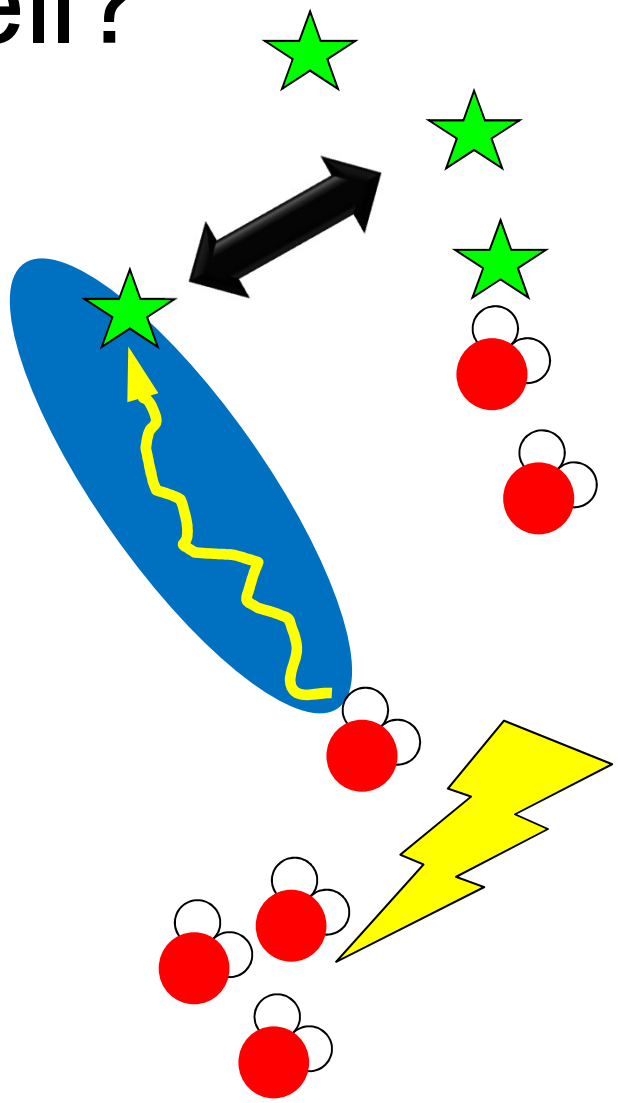


Aromatic parts of the ligand are more important for the binding

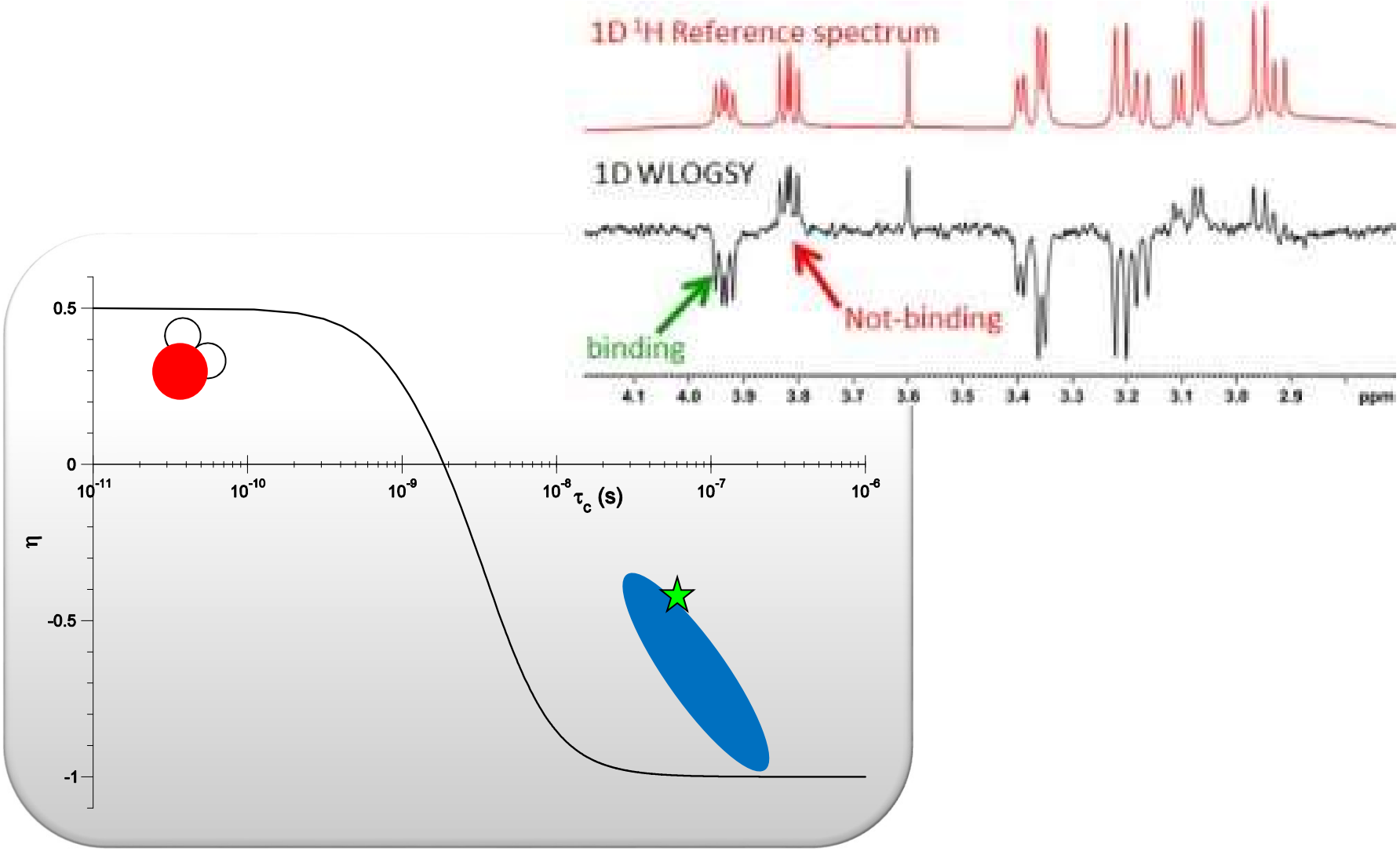
How can we tell?

- Cross-relaxation from water
  - waterLOGSY

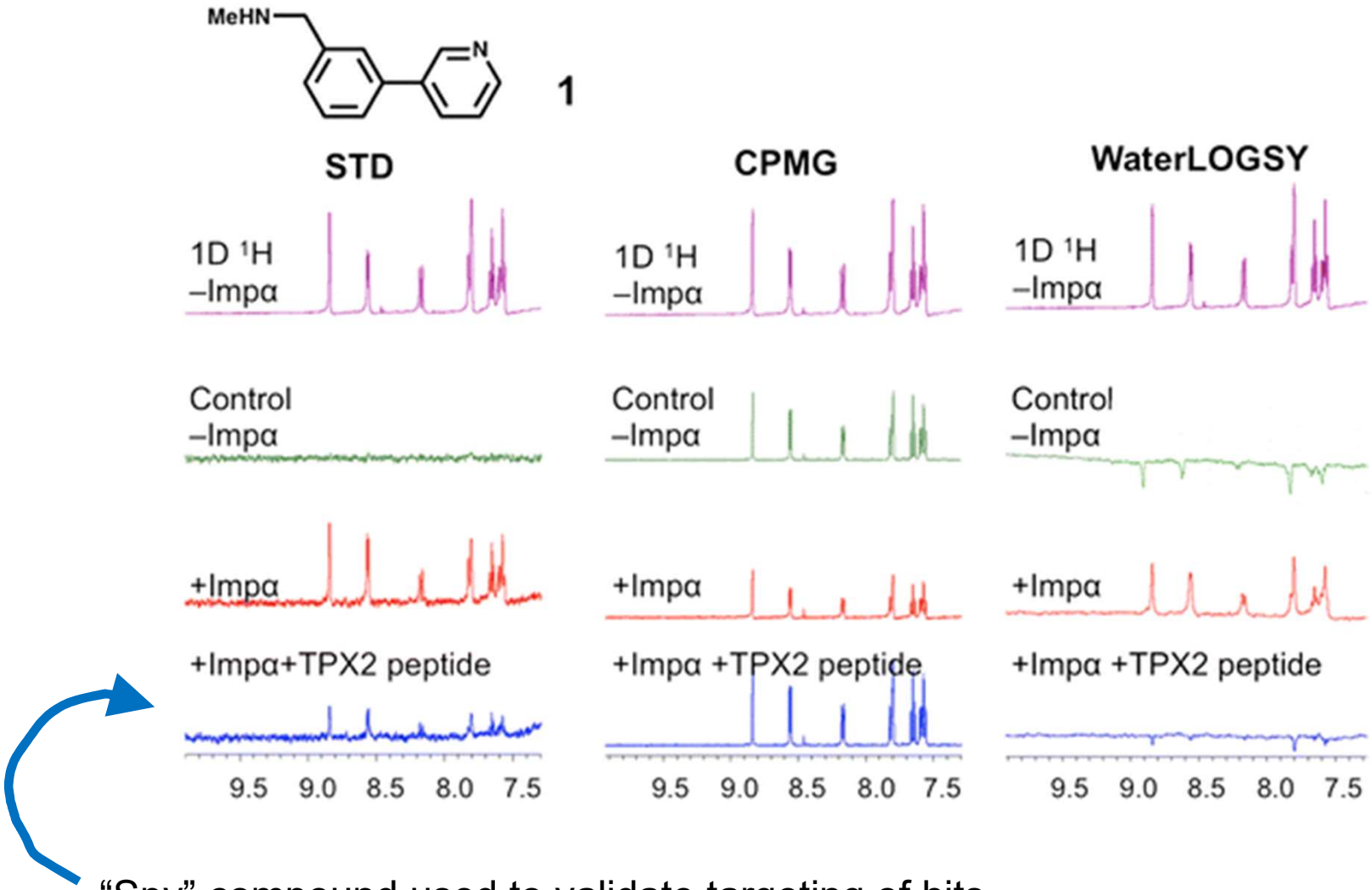




waterLOGSY


1D ^{1}H Reference spectrum
1D WLOGSY
binding
Not-binding
4.1
4.0
3.9
3.8
3.7
3.6
3.5
3.4
3.3
3.2
3.1
3.0
2.9
ppm
0.5
0
-0.5
-1
η
10^{-11}
10^{-10}
10^{-9}
10^{-8}
10^{-7}
10^{-6}
τ_c (s)

Techniques are typically combined



"Spy" compound used to validate targeting of hits

Abell & Dagostin, in *Fragment-Based Drug Discovery*, 2015, pp. 1-18

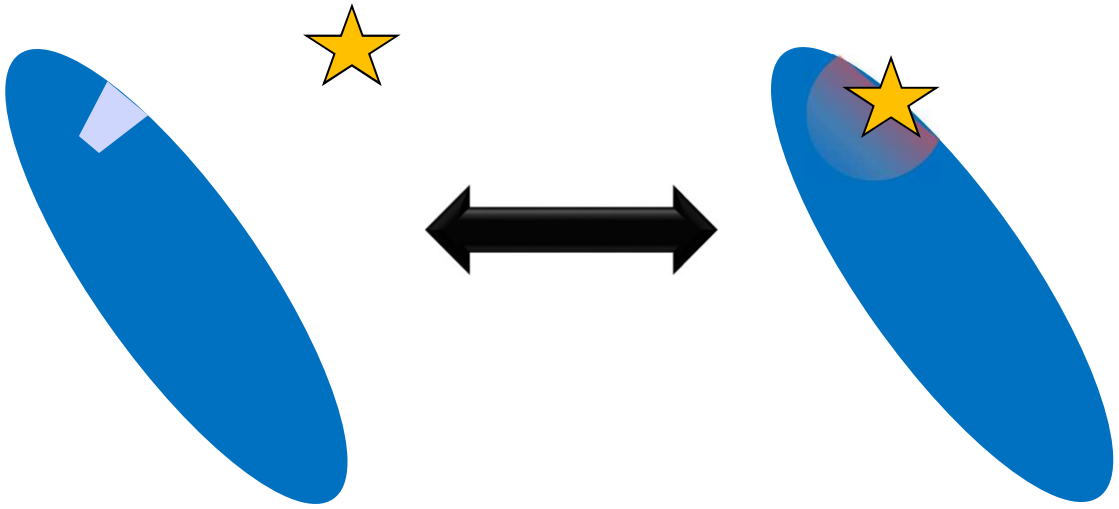
Focusing on macromolecules

Purpose

- High-throughput screening
- Lead optimisation – increasing the strength of the binding
- Mapping of the binding site
- Determination of dissociation constants, K_D

1) Typically used with smaller (5-30 kDa) ^{15}N labeled proteins
2) 2D NMR experiments, typically 2D ^{1}H,^{15}N HSQC
3) Efficient for stronger ($K_D \leq \mu M$) binders (not mM binders)

Protein observe



Chemical exchange

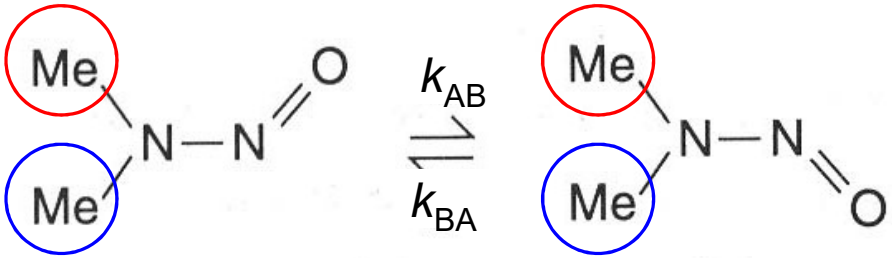
Dynamic equilibrium between two conformations with equal energy or a **symmetrical two-site exchange.**

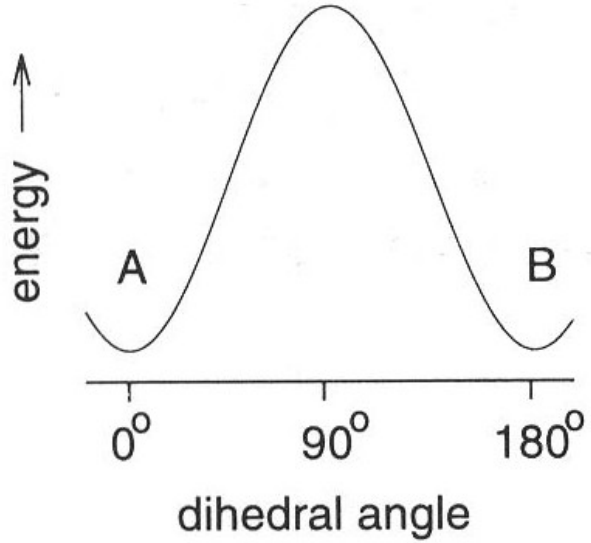
$$A \underset{k_{BA}}{\overset{k_{AB}}{\leftrightarrow}} B$$

$k = k_{AB} + k_{BA}$ [s^{-1}]

Average lifetime $\tau = 1/k$ [s]

Dimethylnitrosamine





- The rotation around the N-N bond is slowed down because of the conjugation of the free electron pairs on N and N=O electrons (partial double bond character).
- -N=O group undergoes 180° rotations.
- Through space effect of N=O group causes the Me groups to have different chemical shifts.
- Two conformations have identical energy (equal population of A and B).

The effect of chemical exchange on NMR spectra

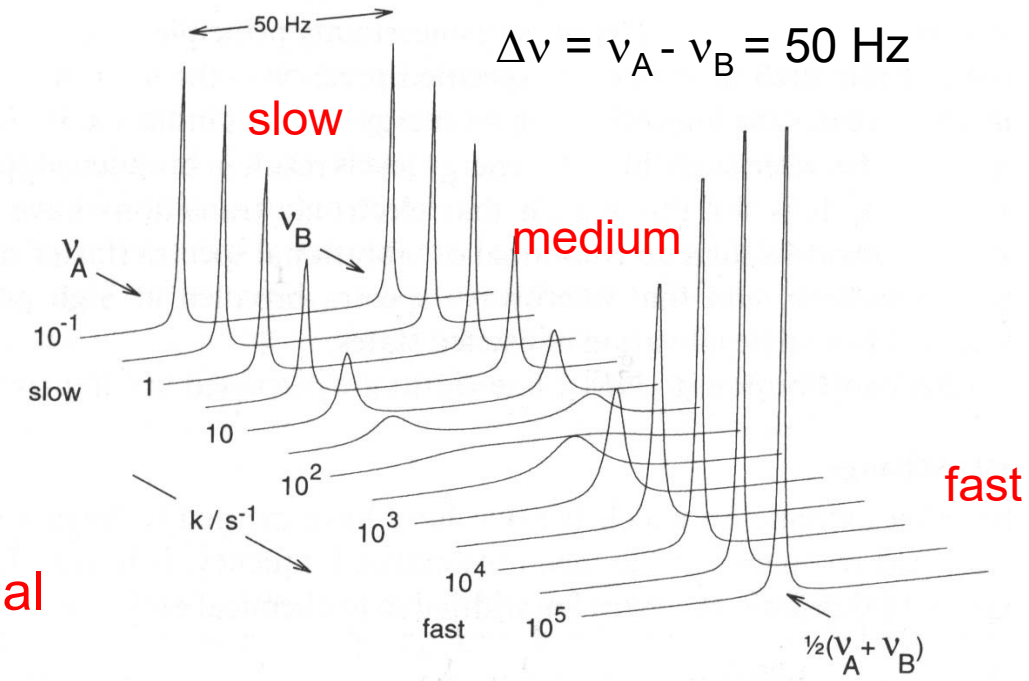
NMR spectra for a pair of nuclei exchanging between two sites with equal population.

$$A \underset{k_{BA}}{\overset{k_{AB}}{\leftrightarrow}} B$$

$|\nu_A - \nu_B| >> k$, two signals

$k \sim |\nu_A - \nu_B|$, signals merge

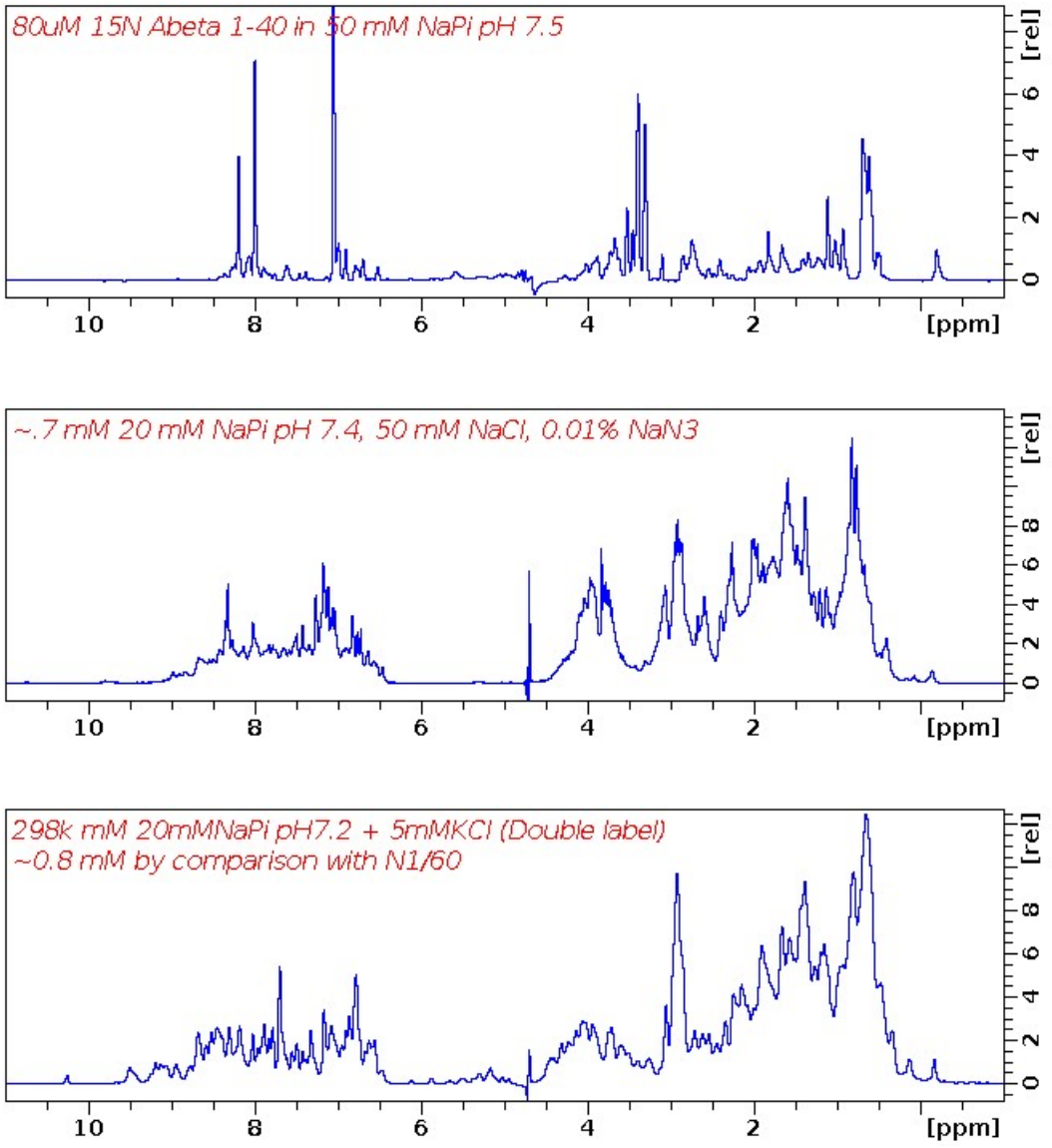
$k >> |\nu_A - \nu_B|$, one sharp signal



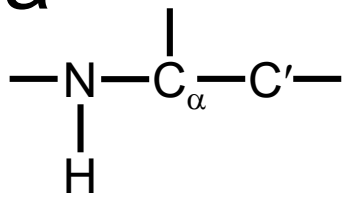
The process is slow (two signals), intermediate (broad signals) or fast (one signal) on the chemical shift time scale depending on the relation between $\Delta\nu$ and k.

- The same considerations apply when the free and bound ligand have different chemical shifts.
- In a protein/ligand equilibrium, P + L $\leftrightarrow$LP, signal of the ligand can therefore be broadened by this intermolecular chemical exchange.

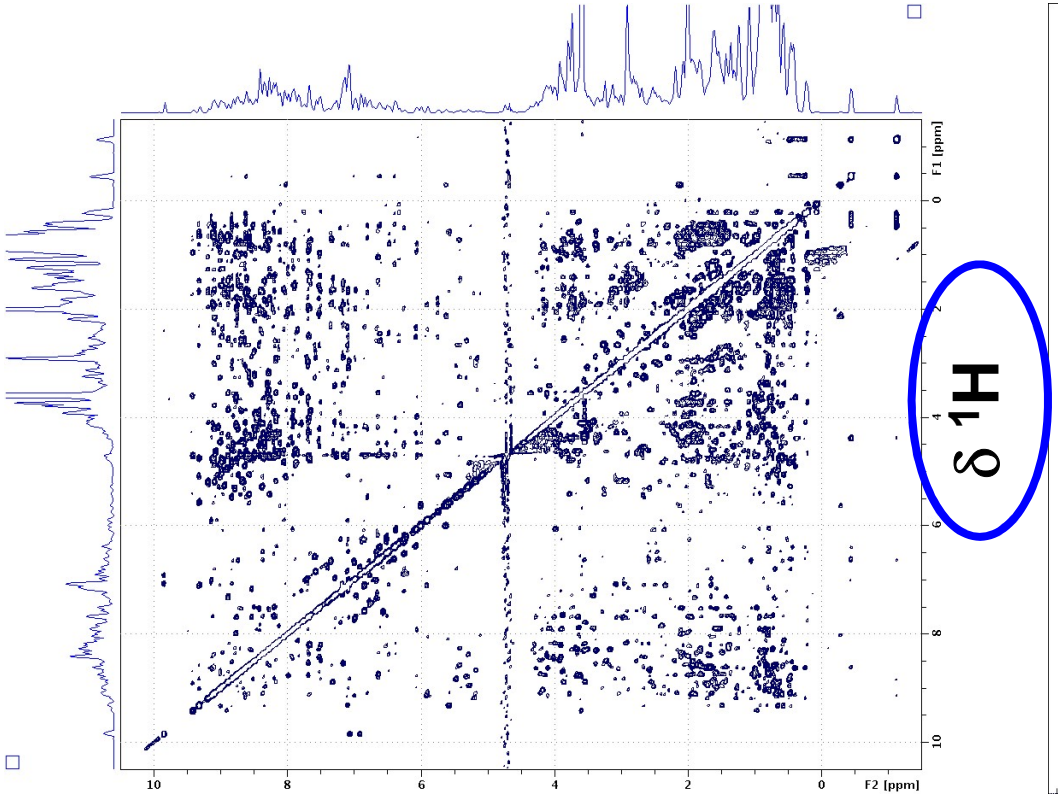
Protein NMR spectral quality



Protein multidimensional spectra



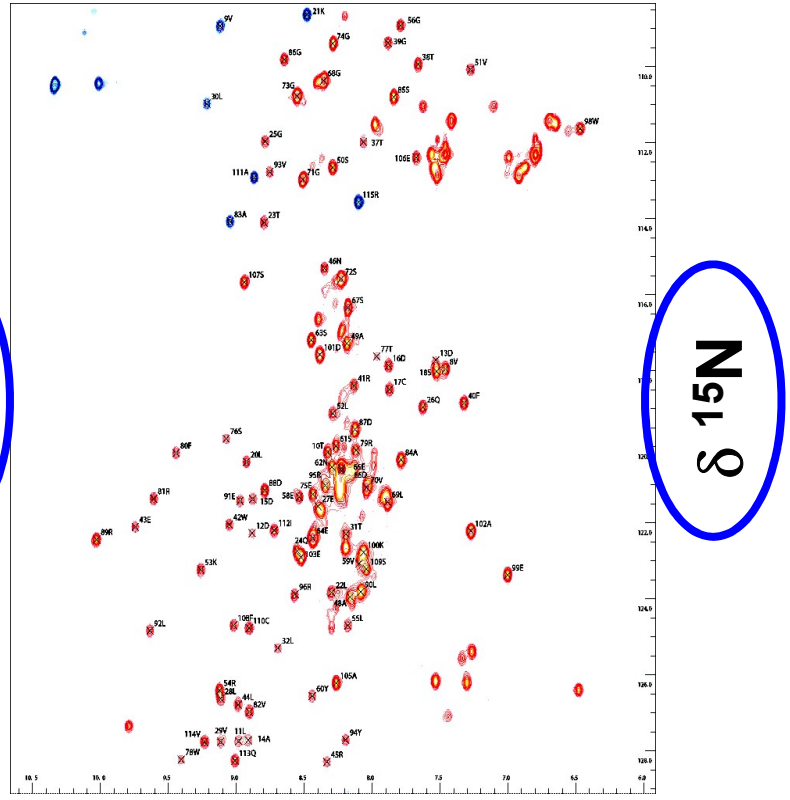
Homonuclear
e.g. NOESY, TOCSY, COSY



δ ¹H

δ ¹H

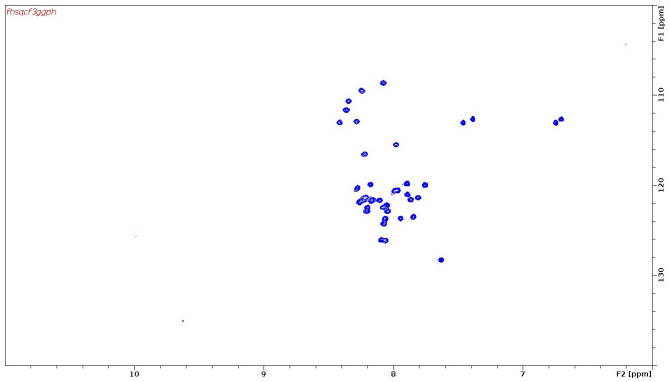
Heteronuclear
e.g. {^{15}N,^{1}H} HSQC

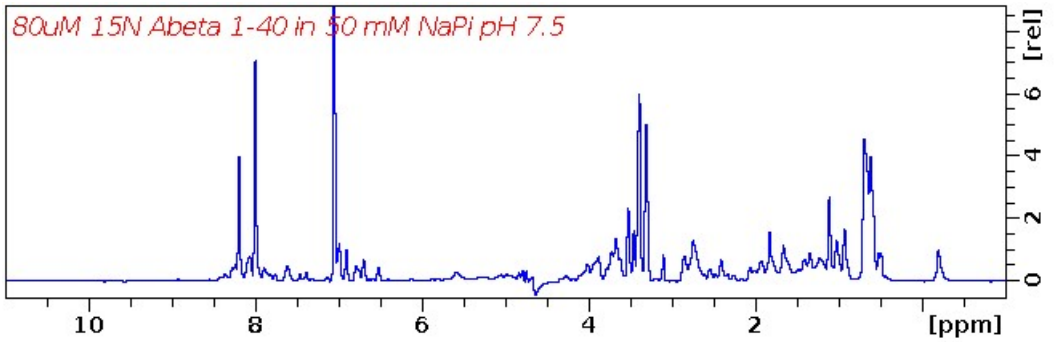


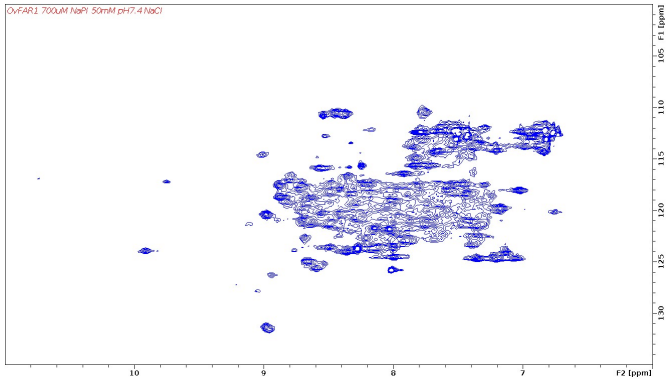
δ ^{15}N

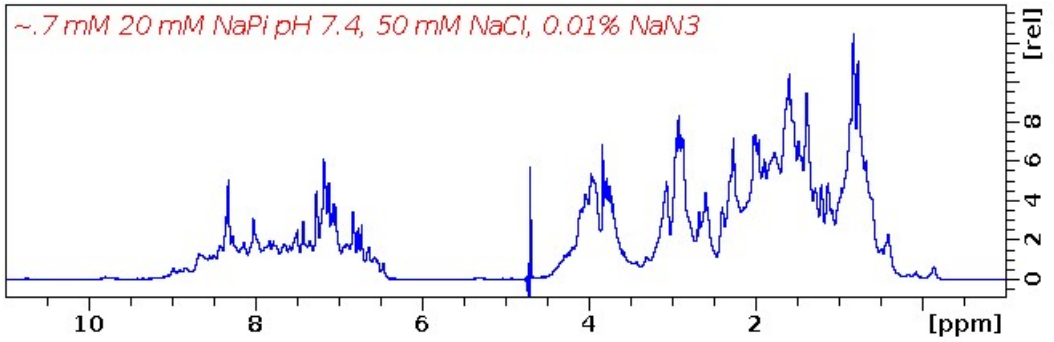
δ ¹H

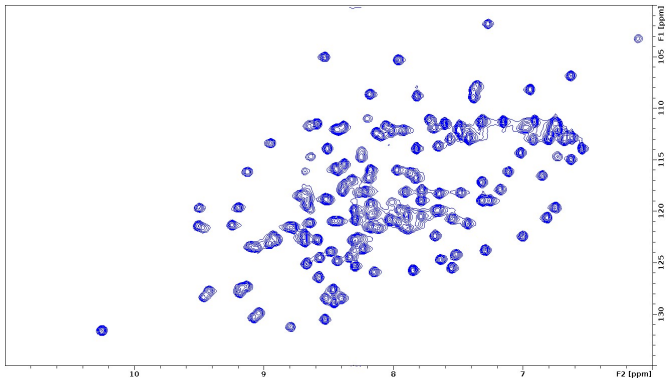
Protein NMR spectral quality

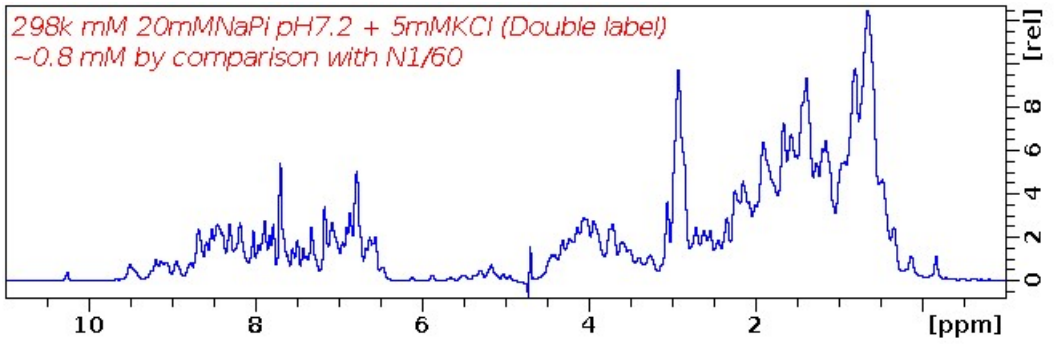












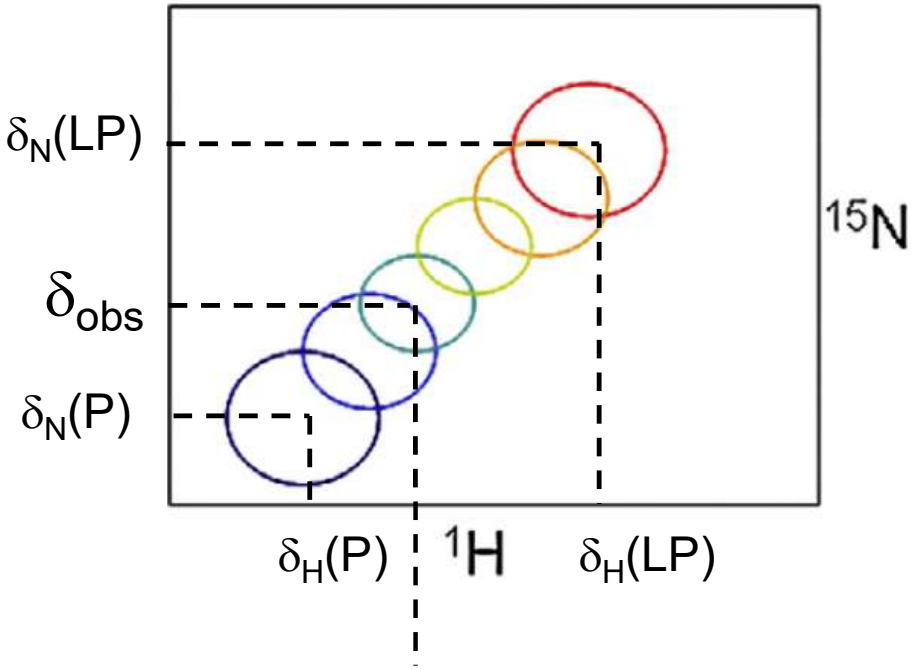
Titration followed by 2D $^{1}H,^{15}N$ HSQC

One NH cross peak,
increasing L concentration

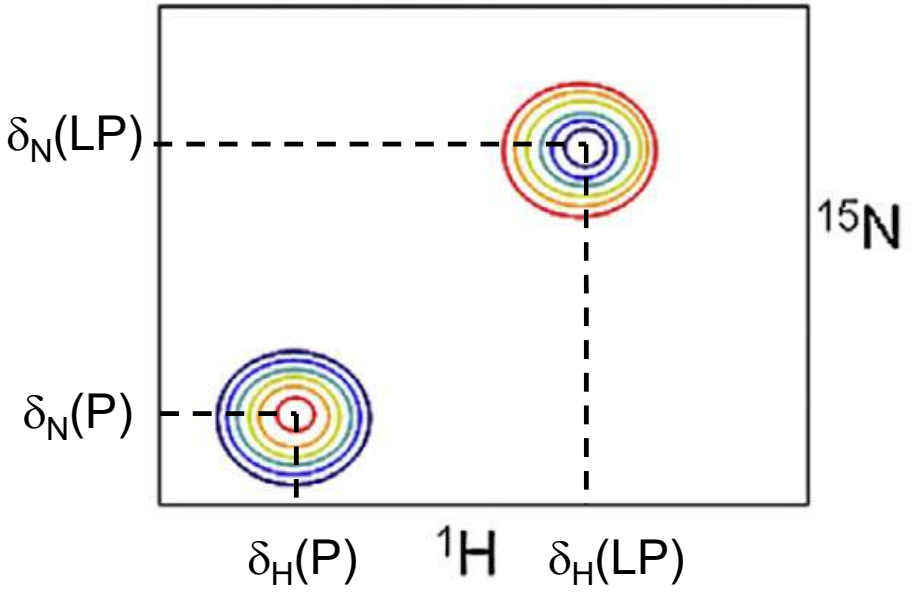
$$P + L \leftrightarrow LP$$

P – protein
L – ligand

Weak binding
Fast exchange



Strong binding
Slow exchange

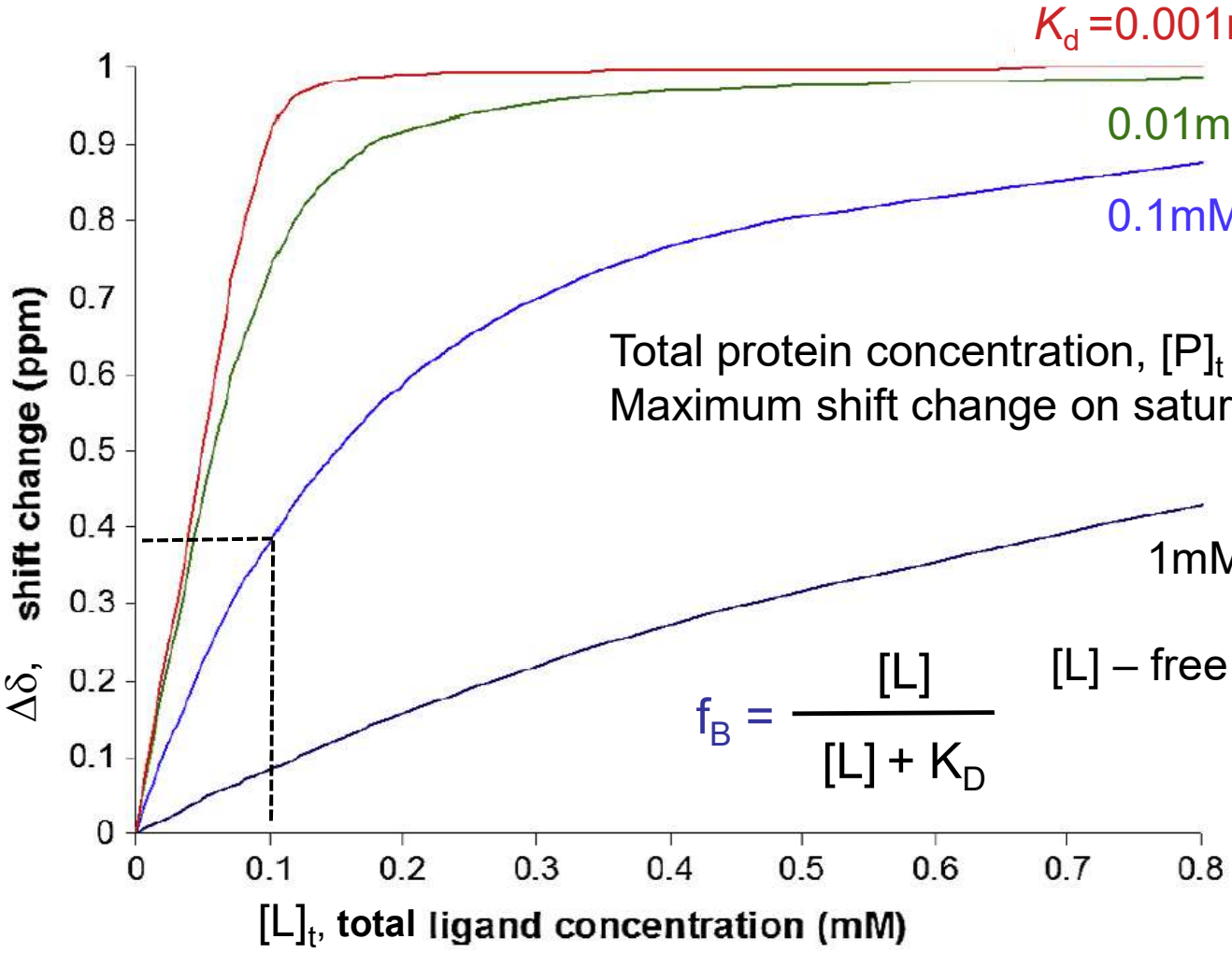


$$\delta_{obs} = (1-f_B)\delta_P + f_B\delta_{PL}$$

$$\Delta\delta_{obs} = \delta_{obs} - \delta_P$$

δ_P – chemical shift of the free protein
δ_{PL} – chemical shift of the protein in the protein-ligand complex
f_B – fraction bound

K_D can be determined in the fast exchange regime



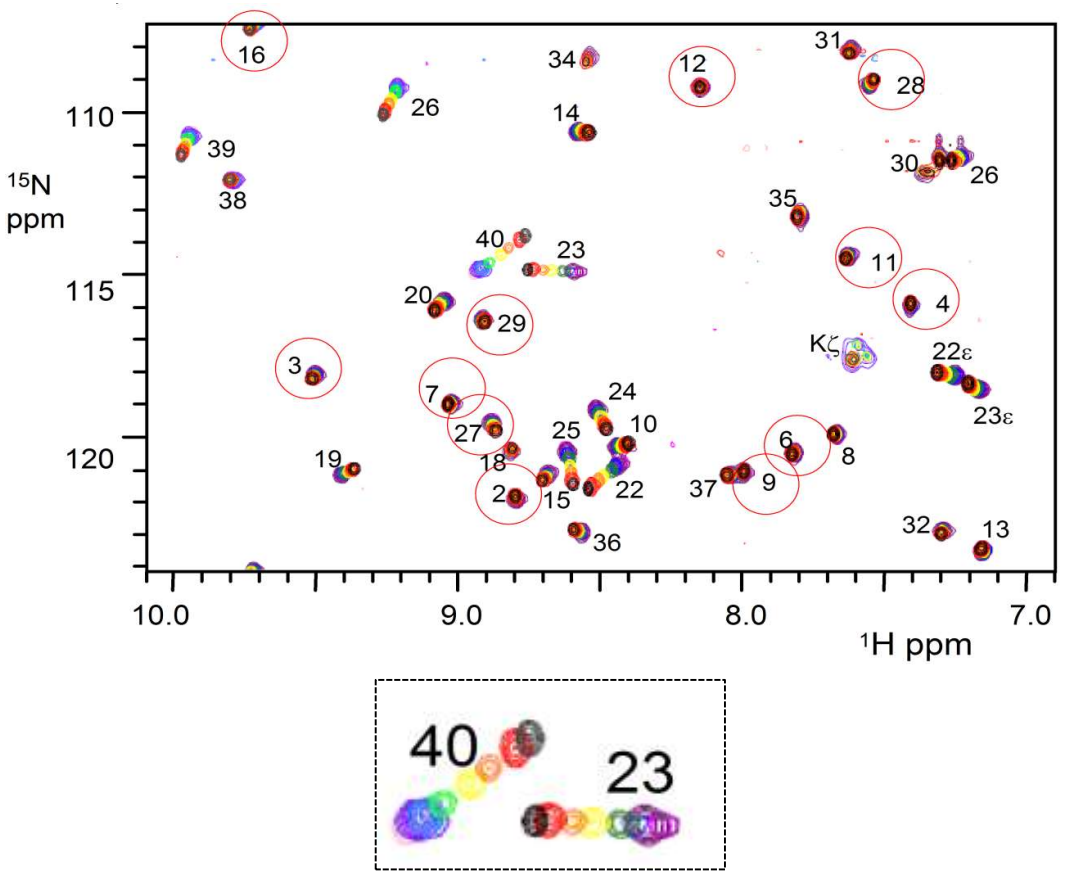
The following equation can be used to determine K_D from saturation curves
(n - number of ligands per one protein molecule)

$$\Delta\delta_{obs} = \Delta\delta_{max}\left\{(n[P]_t + [L]_t + Kd) - [(n[P]_t + [L]_t + K_d)^2 - 4n[P]_t[L]_t]^{1/2}\right\} \Big/ 2n[P]_t$$

An example of K_D determination

K_D values determined from the titration of a hexasaccharide into a protein.
Average K_D= 5.1 ± 5 μM.

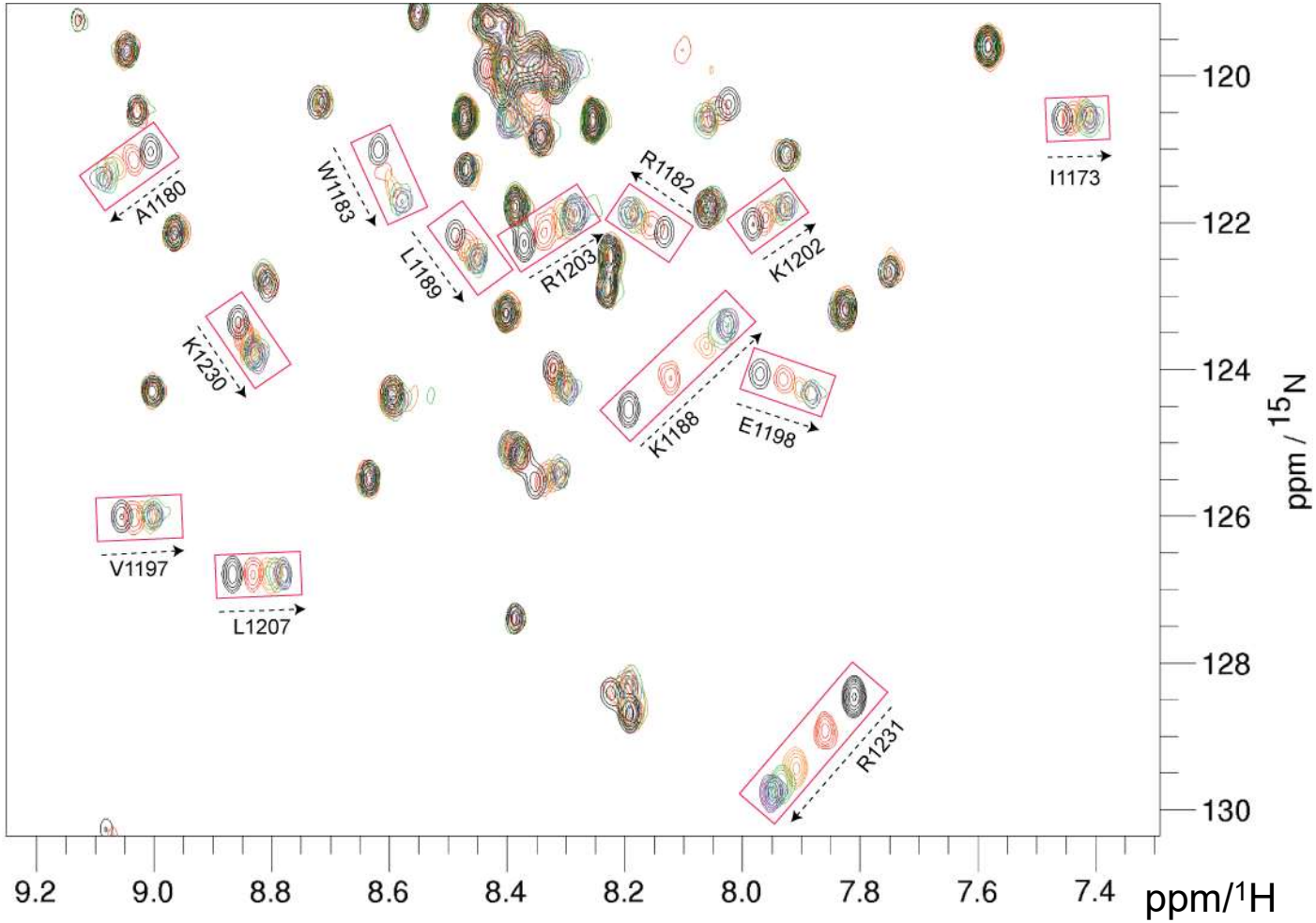
Residues	K_d (μM)	Residues	K_d (μM)
Ile2	17.1	Arg23	4.3
Gly3	8.0	Arg23Hε	5.6
Asp4	4.0	Tyr24	2.4
Val6	1.9	Lys25	2.4
Thr7	10.0	Gln26	1.2
Cys8	3.5	Gln26Hε	1.1
Leu9	12.9	Ile27	1.8
Lys10	1.4	Gly28	1.6
Ser11	5.3	Thr29	11.3
Gly12	3.8	Cys30	8.4
Ala13	1.7	Gly31	N/A
Ile14	1.3	Leu32	2.8
Cys15	4.8	Gly34	9.4
His16	17.9	Thr35	3.7
Val18	14.3	Lys36	2.4
Phe19	2.2	Cys37	4.2
Cys20	2.9	Cys38	1.6
Arg22	3.0	Lys39	1.5
Arg22Hε	5.2	Lys40	2.5



In the limit of *very fast exchange,* peaks have the same shape throughout the titration. As they move out of this limit, peaks may become broader and then sharpen up again close to saturation. *K_D*s can be determined without knowing the identify of residues!

Octasaccharide titrated into a protein

Moving residues are affected by the binding, either because (i) they form the binding site or (ii) they are moving due to allosteric effects.

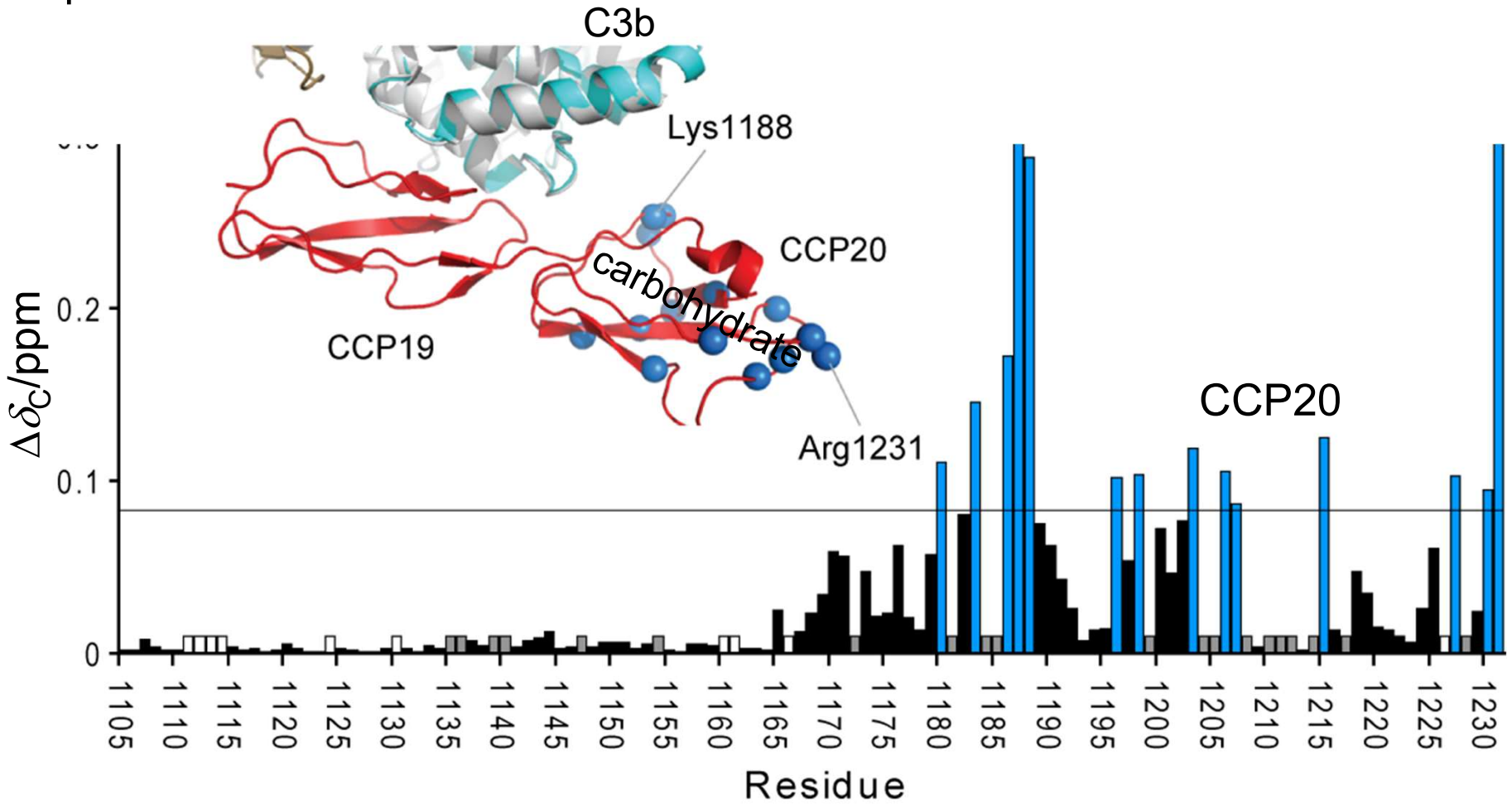


Movement of signals can be quantified and used to map the binding site

Combined chemical shift difference, $\Delta\delta_C$

- $\Delta\delta_C$ between the pure protein and last titration point is calculated
- A carbohydrate molecule binds only to the CCP20 module of this double module CCP protein that is bound to C3b

$$\Delta\delta_C = \sqrt{\Delta\delta_H^2 + \left[\frac{\Delta\delta_N}{5}\right]^2}$$

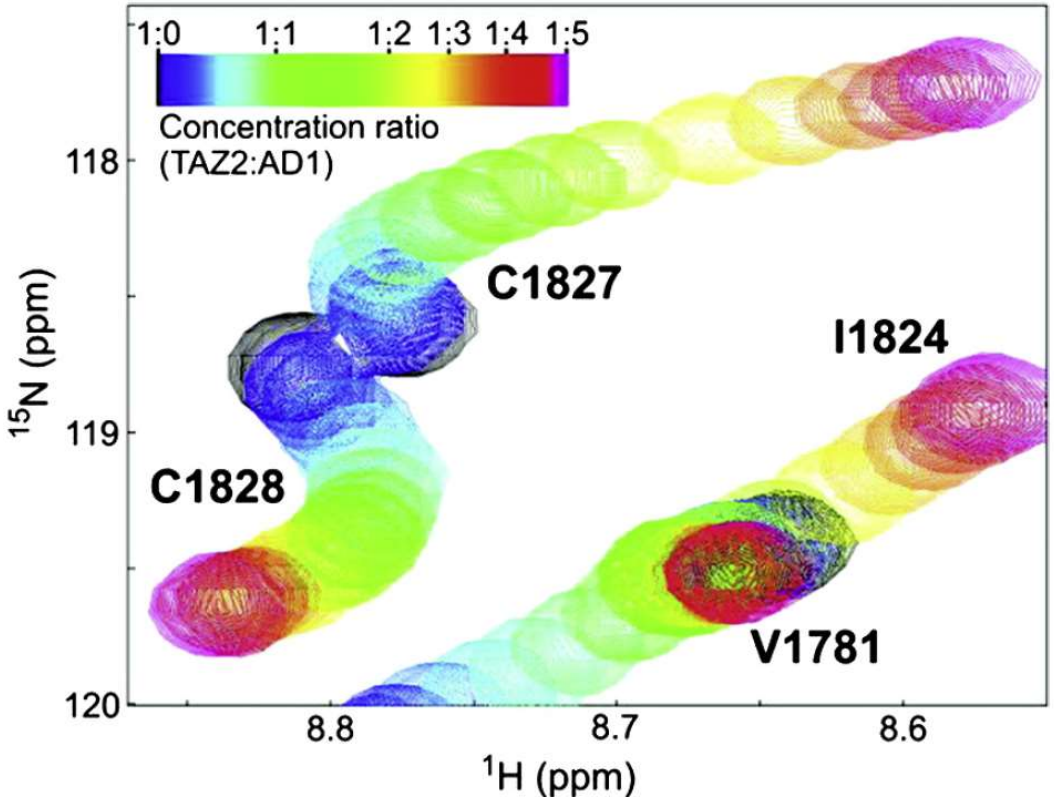


Conformational change induced by the binding

"Curved" movement of signals (and often broadened signals) are usually a sign of conformational changes and can be described by the following equilibria:

1. $P + L \leftrightarrow P^* + L \leftrightarrow PL$ or 2. $P + L \leftrightarrow PL \leftrightarrow PL^*$

Black (free protein) to magenta (1:5 ratio)



1. L binds to a minor state of P, referred to as P*, eventually forming a PL complex

2. The PL complex rearranges upon binding of L into PL* where conformation of P is changed.

SHAPES library and drug discovery

The SHAPES library is a small collection of diverse, low molecular weight, water soluble compounds whose molecular shapes represent those most commonly found in known drug molecules[1].

From an analysis of the comprehensive medicinal chemistry (CMC) database, it was determined that 32 different frameworks describe 50% of all known drugs[2].

When atom type was included, 41 frameworks describe 24% of all known drugs.

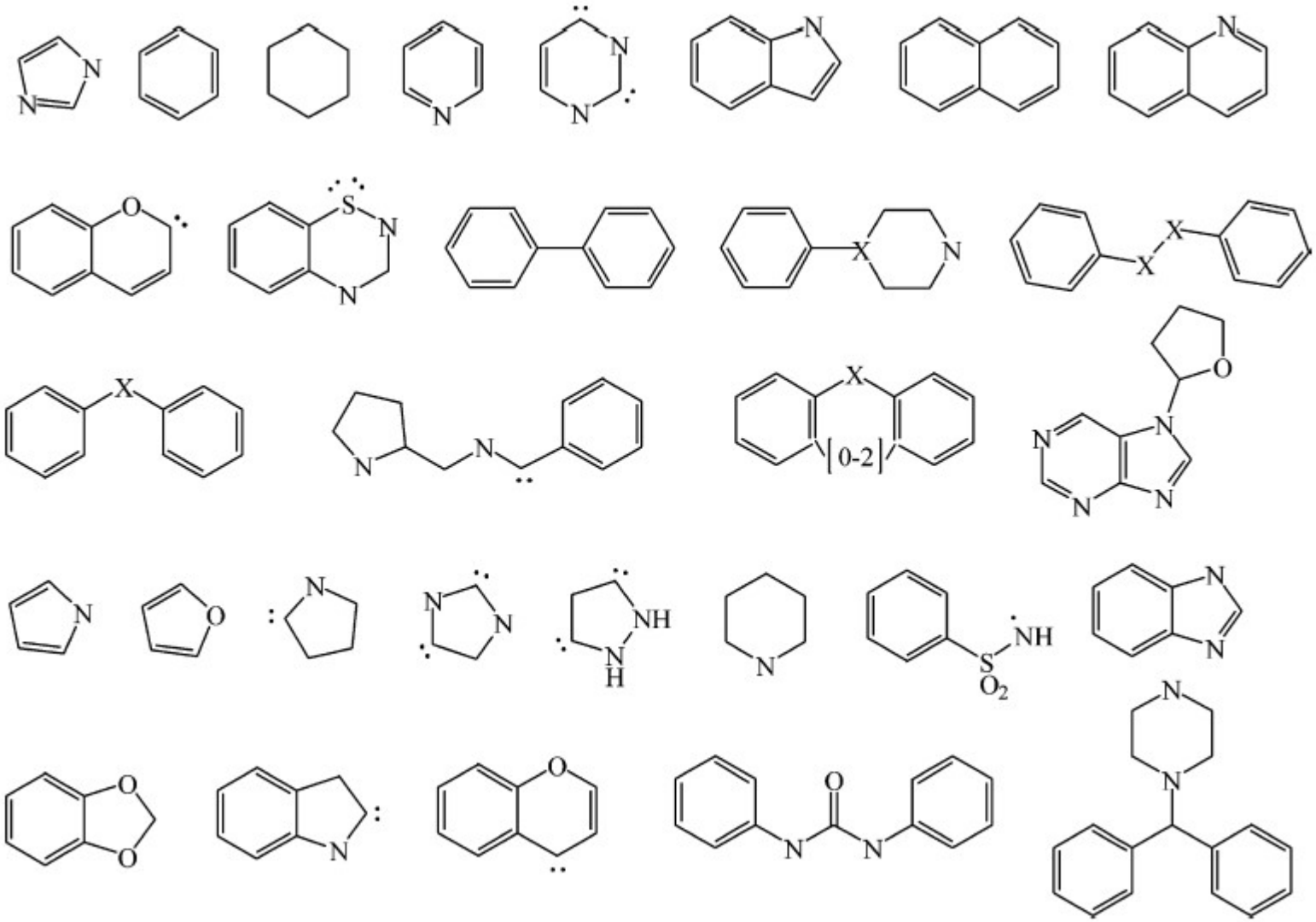
Framework classification was combined with similar data on the most common drug side chains[3] to create the SHAPES library.

The 132 compounds in the original SHAPES library[1] are based on the molecular frameworks shown on the next slide.

A SHAPES library created subsequently contains 500 compounds, each with a combichem accessible linker or side chain[4].

1. J. Fejzo, et al, Chem. Biol. 6 (1999) 755.
2. G.W. Bemis, M.A. Murcko, J. Med. Chem. 39 (1996) 2887.
3. G.W. Bemis, M.A. Murcko, J. Med. Chem. 42 (1999) 5095.
4. C.A. Lepre, Drug Discov. Today 6 (2001) 133.

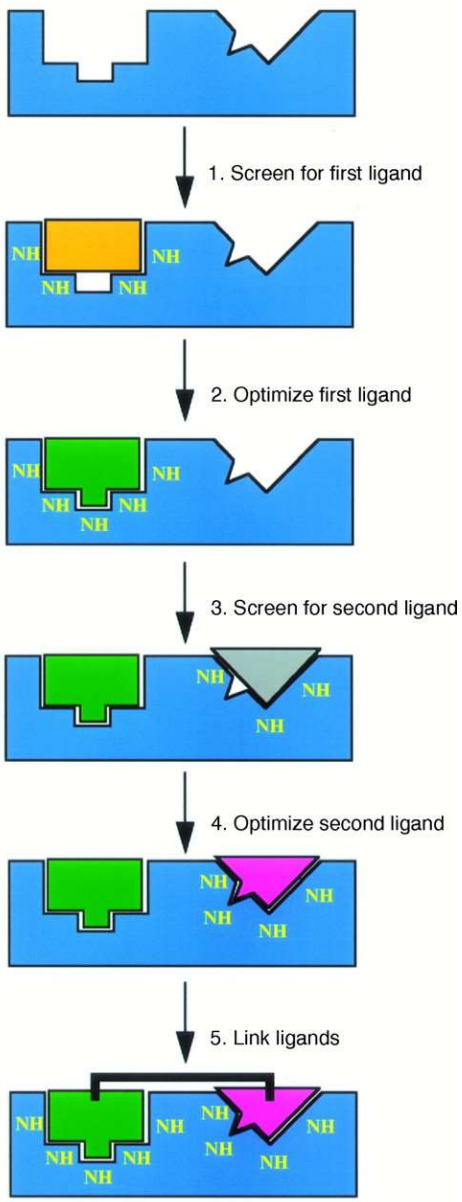
The original SHAPES library

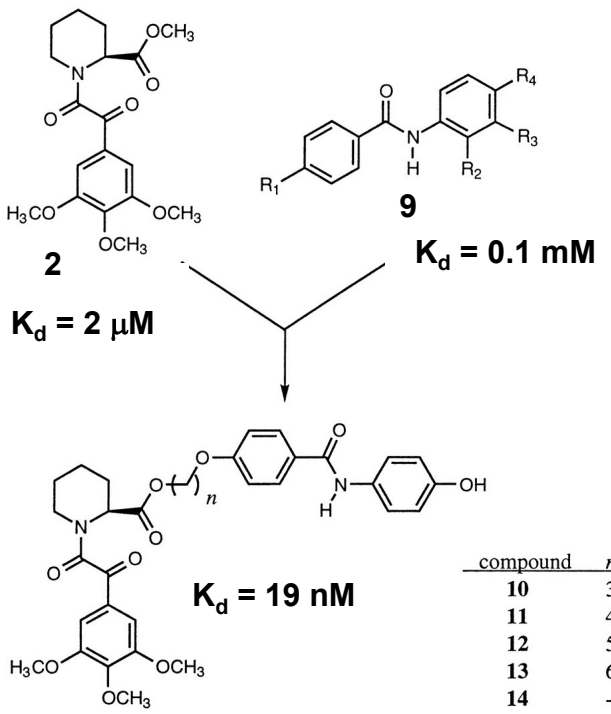


Sidechain attachment points are denoted by single electrons or lone pairs. X represents a C, N, O or S.

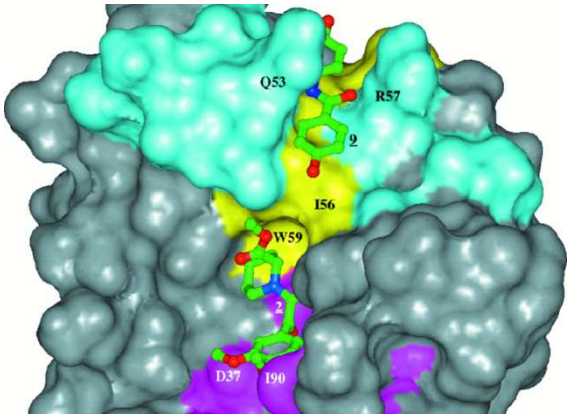
J. Fejzo, et al, Chem. Biol. 6 (1999) 755.

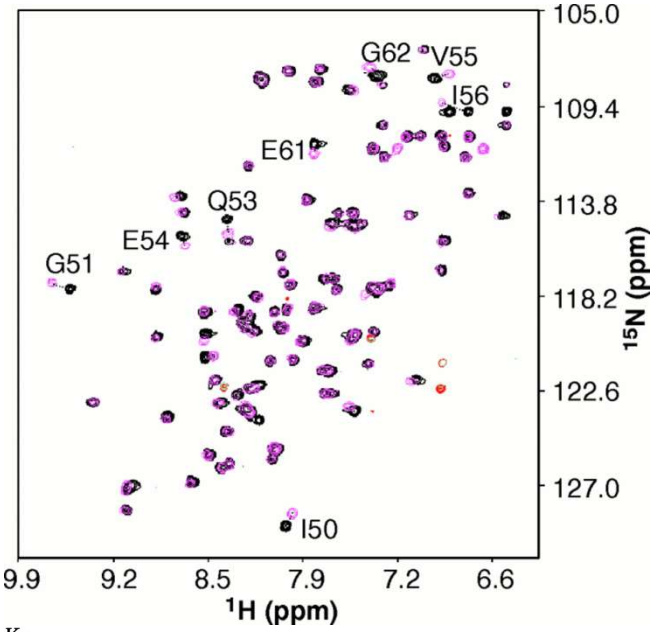
Structure-activity relationships: SAR by NMR





compound	*n*	K_d
10	3	19 nM
11	4	32 nM
12	5	49 nM
13	6	228 nM
14	-	49 nM





Requirements:

- ^{15}N labelled protein
- < ca 40 kD
- preferably assigned NH resonance
- preferably a cryoprobe instrument (sensitivity)

SW. Fesik *Science*, 274, 1531-1534 (1996)