



Core Technologies for Education and Innovation in Life Sciences



Study of Interactions and Protein Structure Determination by NMR

For Application to Protein Characterization

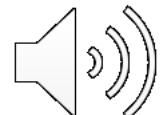
by

Radovan Fiala, Karel Kubicek, and Pavel Kaderavek
CEITEC, Masaryk University

innocore-project@unitn.it

This project received funding by the European Commission under the GA n 2019-1-IT02-KA203-063391

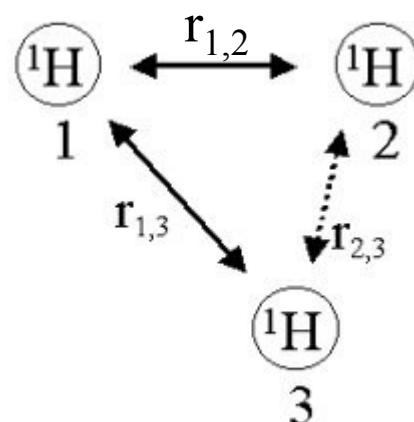
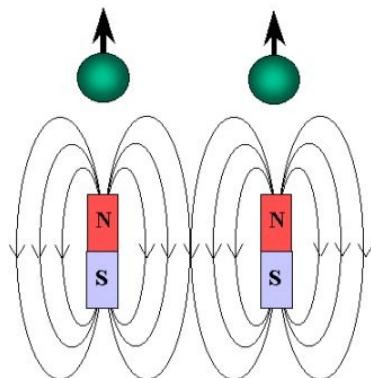
[www.innc
.eu](http://www.innc.eu)



NMR as a tool for study **structure, dynamics and interactions** of biomolecules

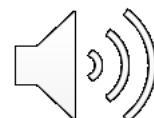
- 0) AA/NA sequence, resonance assignment, standard chemical shifts
- 1) Structure determination of proteins/NAs
- 2) NMR can provide detailed information about the structure at the atomic level resolution relying on the spatial proximity of two interacting protons – nuclear Overhauser enhancement (**NOE**)
- 3) Additional structural information can be obtained (residual dipolar couplings – **RDCs**, **J**-couplings, backbone chemical shifts - **CSI**)

NOE:



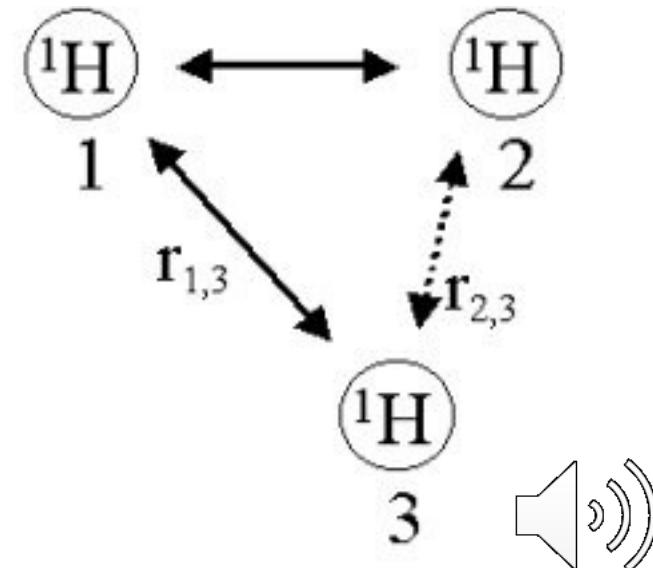
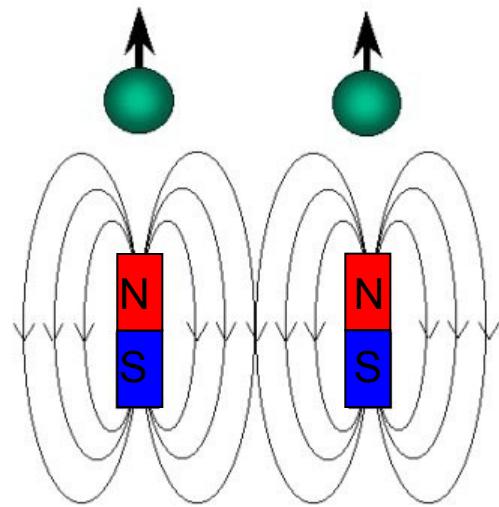
$$r_{1,2}; r_{1,3}; r_{2,3} \leq 6 \text{ \AA}$$

$$1 \text{ \AA} = 1.10^{-10} \text{ m}$$

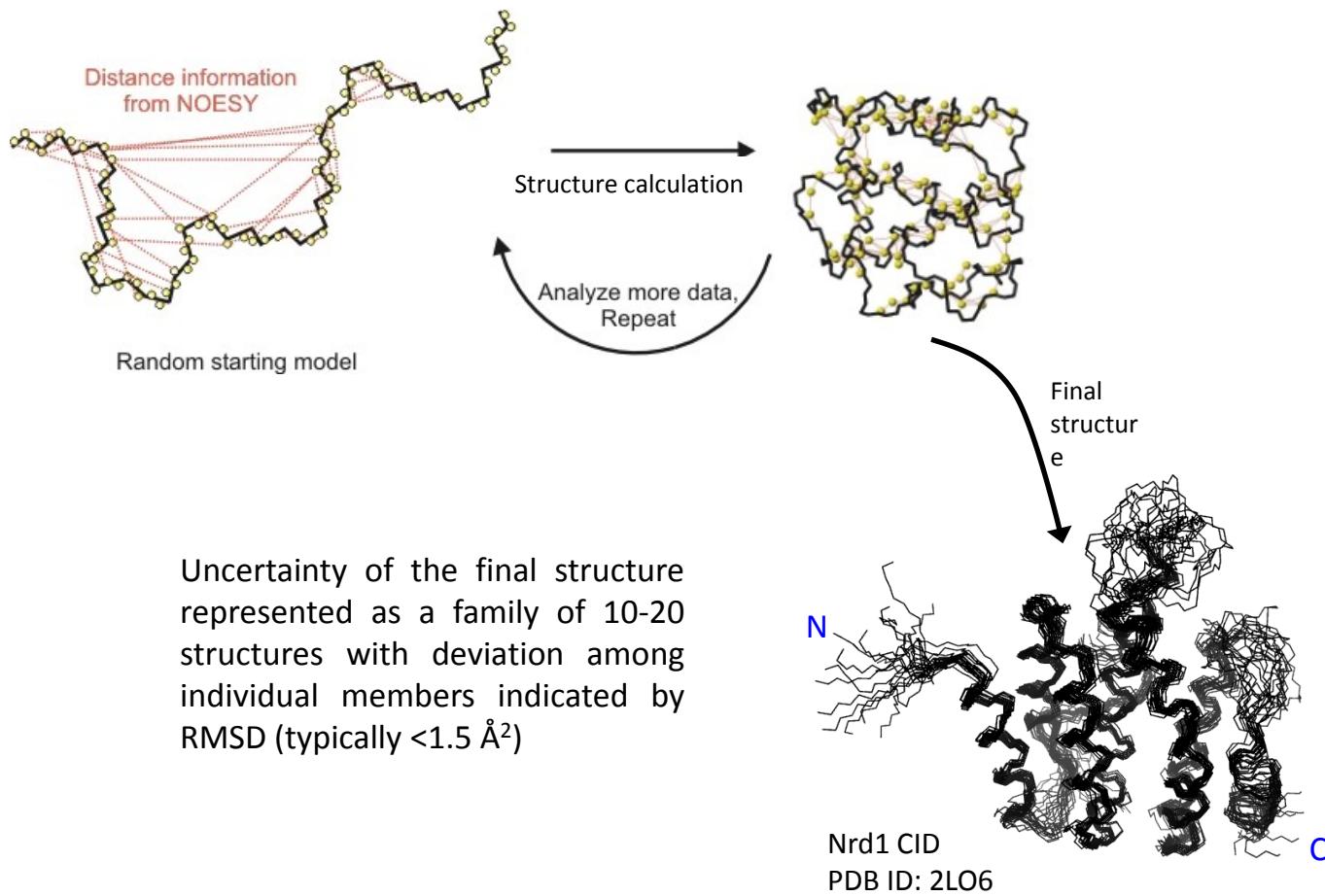


Nuclear Overhauser Effect (SpectroscopY) = NOE(SY)

- i) caused by dipolar coupling between nuclei.
- ii) the local field at one nucleus is affected by the presence of another nucleus.
- iii) the result is a mutual modulation of resonance frequencies.
- iv) the NOE operates through space.
- v) the intensity of the interaction is a function of the distance between the nuclei according to the following equation: $I = A(1/r^6)$, I is the intensity, A is a scaling constant, and r is the distance between the nuclei
- vi) the NOE provides a link between an experimentally measurable quantity, I , and internuclear distance
- vii) NOE is only observed up to $\sim 6\text{\AA}$



Iterative procedure of structure determination by NMR

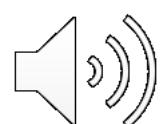
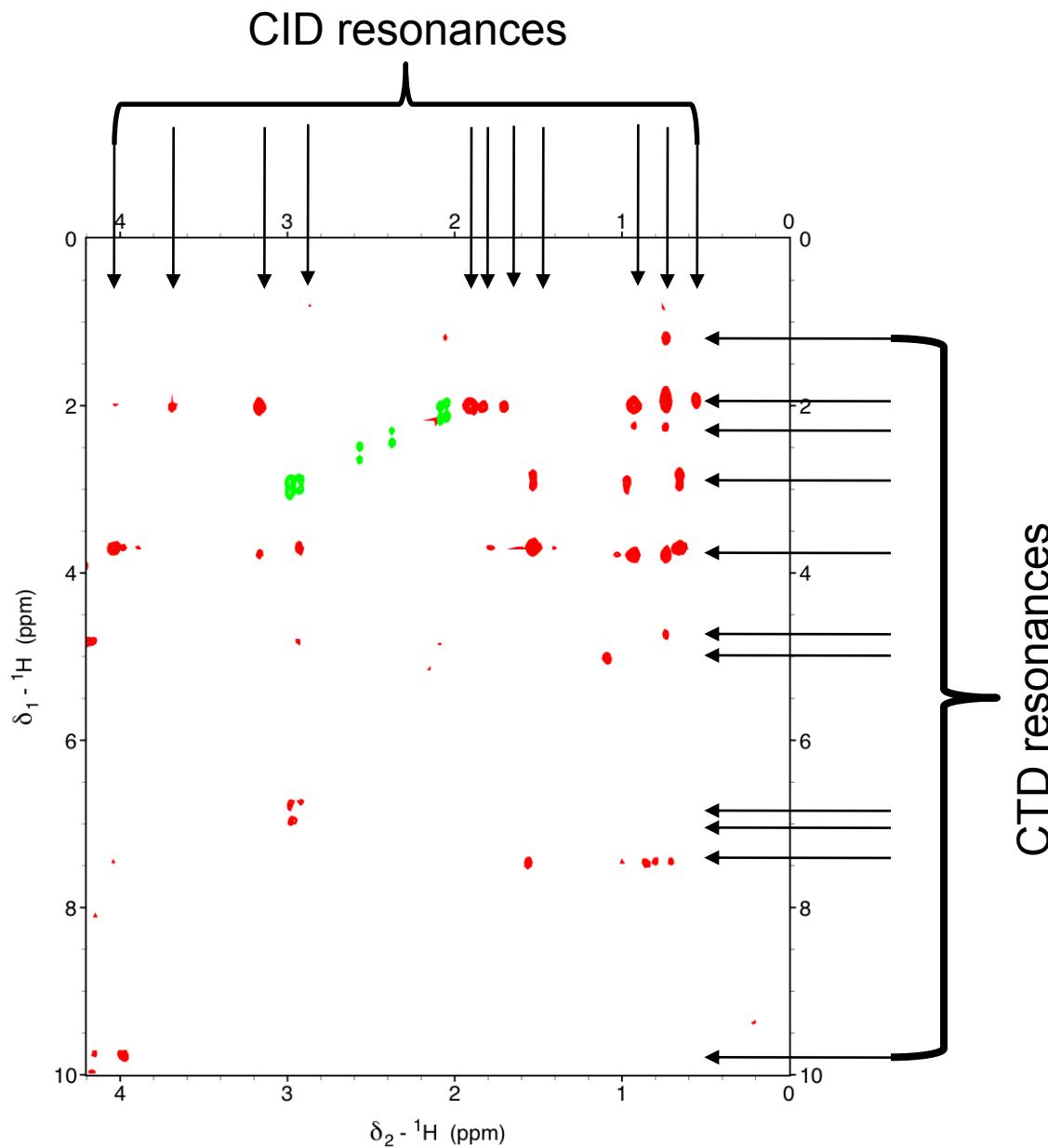


Uncertainty of the final structure represented as a family of 10-20 structures with deviation among individual members indicated by RMSD (typically $<1.5 \text{ \AA}^2$)

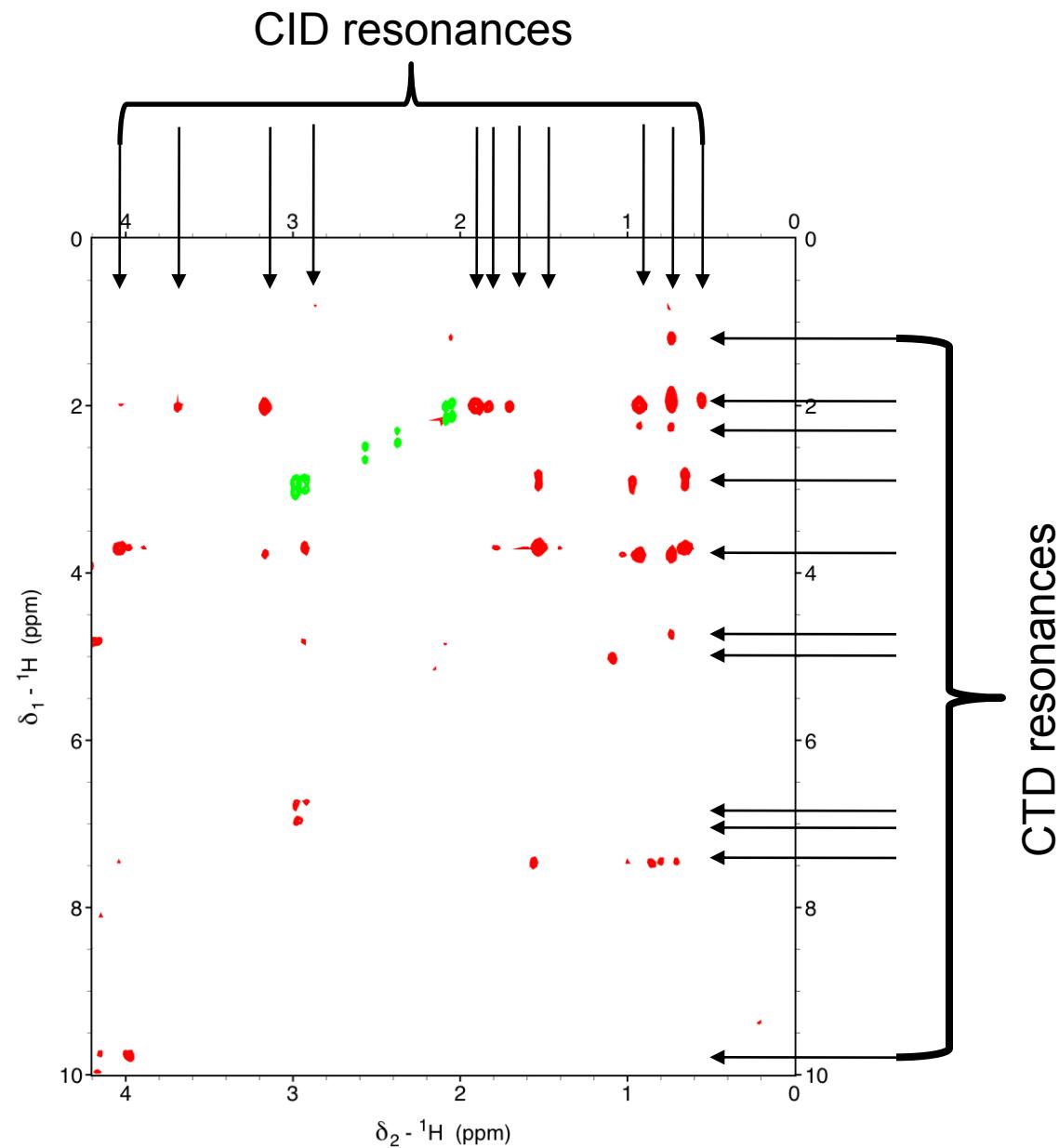
http://www.fbreagents.com/basics_nmr/9proteins.htm



Interligand NOEs between CID and CTD – 900MHz, 150ms, 293K



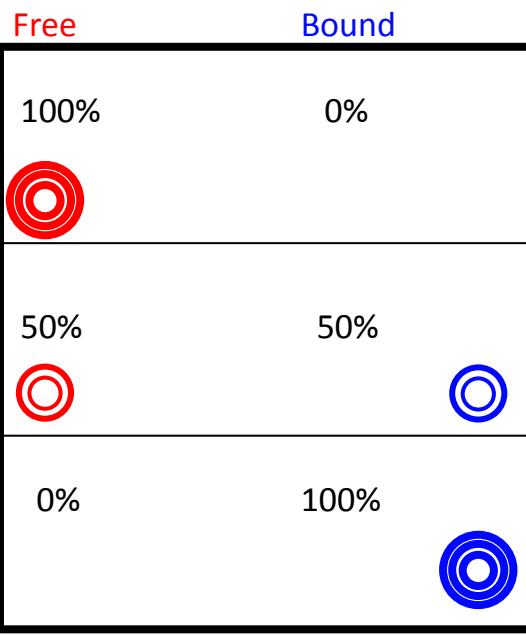
Interligand NOEs between CID and CTD – 900MHz, 150ms, 293K



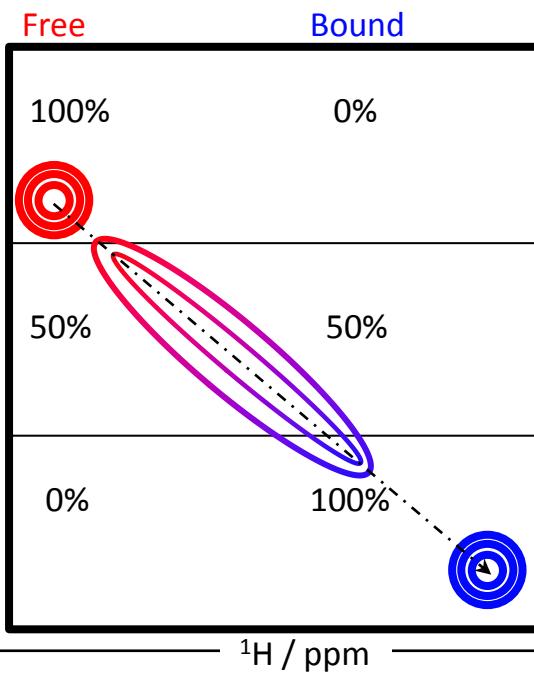
Studying interactions by NMR titration

- 1) **Slow** exch. regime (on the NMR timescale)
 - individual peaks for each of the studied states (e.g. free / complexed forms of a protein), peak intensity representing population of a given state
- 2) **Intermediate** exchange regime
- 3) **Fast** exchange regime
 - single peak whose chemical shift position is given by the molar ratio of the states present in solution

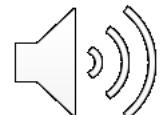
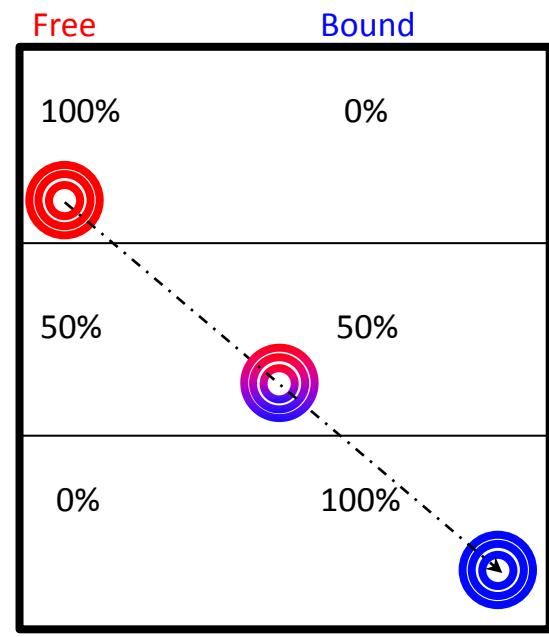
Slow ($K_D < 1 \mu\text{M}$)



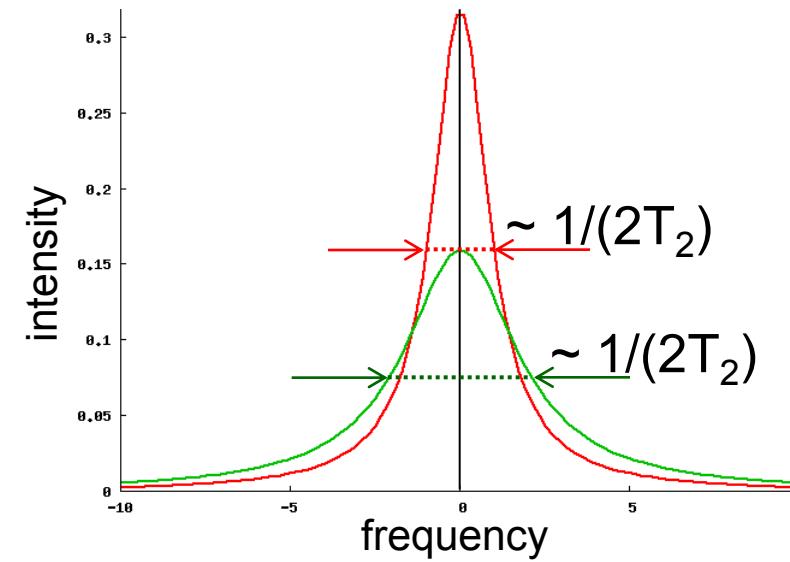
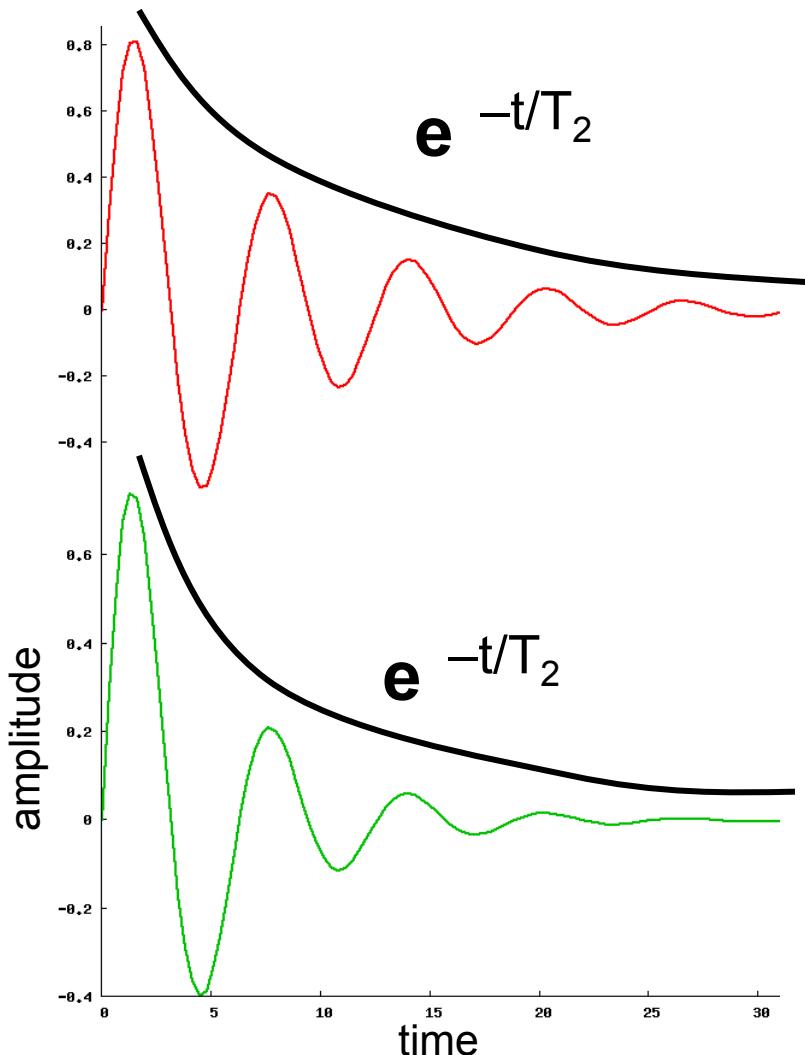
Intermediate



Fast exchange regime ($K_D > 10 \mu\text{M}$)

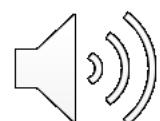


Not all molecules relax (decay) with the same rate



Bigger molecules (higher molecular weight) relax faster \Rightarrow **broad peaks**

Small molecules relax slower \Rightarrow **narrow peaks**



Size



Relaxation

slow (i.e. long t_2 time)

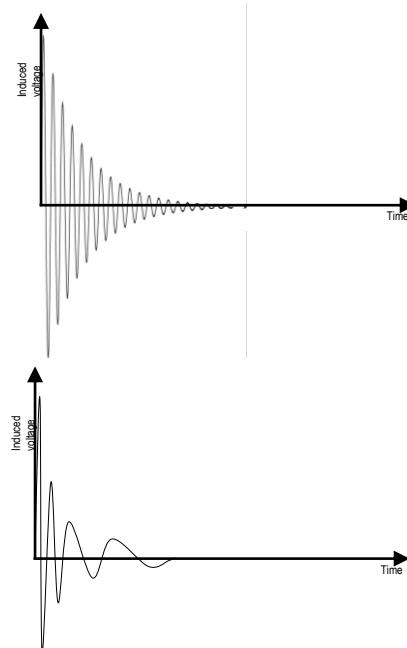
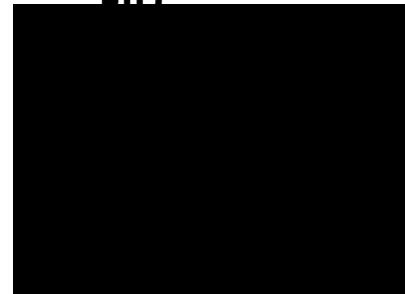


medium

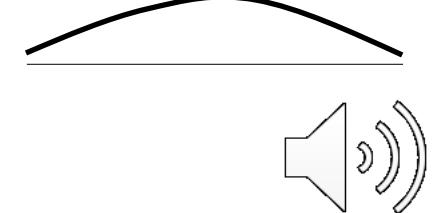
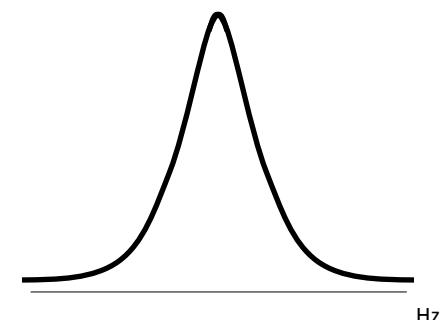
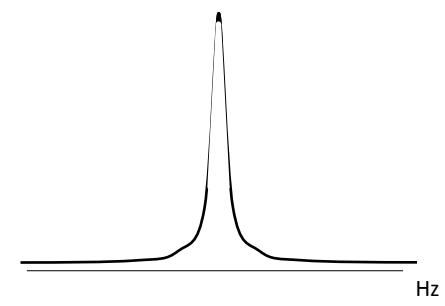


fast

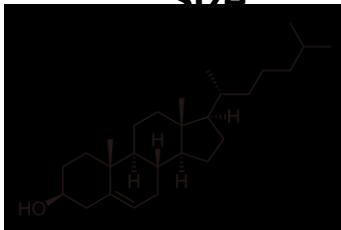
FID



NMR line(width) after FT



Size



e.g. Cholesterol

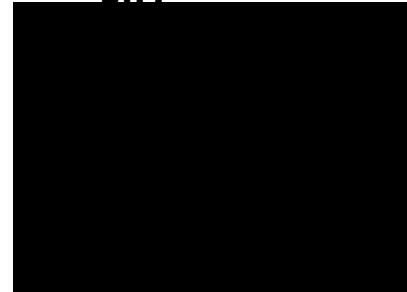
Relaxation

slow (i.e. long t_2 time)

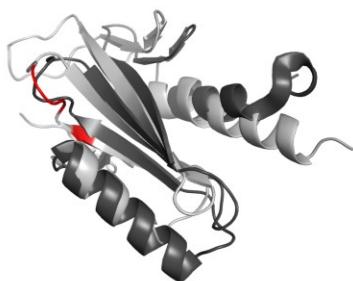
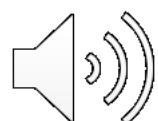
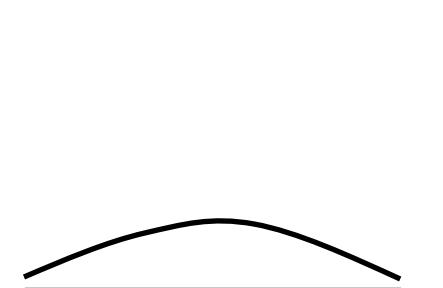
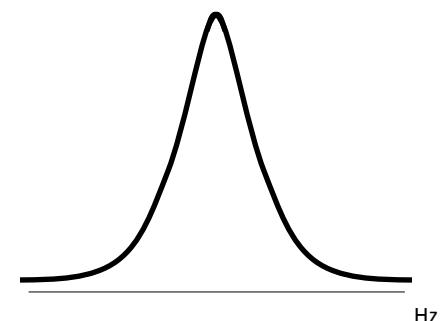
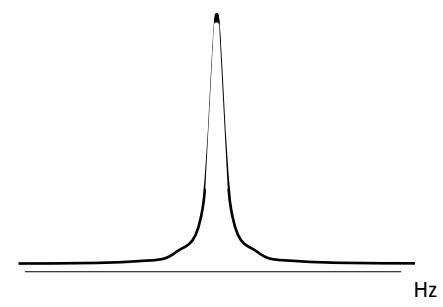
medium

fast

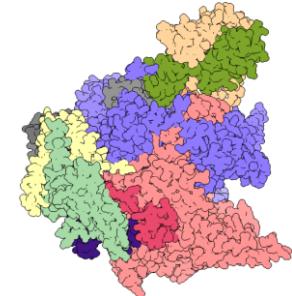
FID



NMR line(width) after FT



Biomolecules 5-30 kDa



Large molecules 50+ kDa

Protein – metal ion interaction

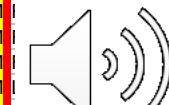
slow exchange case



DR1885 from *deinococcus radiodurans*

1 AAYMTLTNKSP	VGAATPLATSP	LMTT	T HSG-	GMA	C MKMVPW-	LTI	P ARGTL	LTLQ	R D G	H	M	MGLKRP	
2 AAYFTLENHS	GVTTDVANEAQI	M	QEQL	GMV	M RHEVG	F DIP	S HGTL	T L	MPS	H	M	I GLKAP	
3 AFFGEVMN	SASSEVAKVE	D	I KEG	DVM	M RQVPE	F V	I PAQGT	L VL	K PG	H	M	E LKKP	
4 AGYVTLVNA	DIS SNDY	DAML	V SNG	S SQ	M MVVDK	L TV	P AHG	Q VAI	A PG	H	M	E DAHKH	
5 AIFLTIFNN	SAKSDISE	SEL	H T	H KHD	G KMM	M QKIP	I I	IKAHS	STEL	K SG	M	L LKLKKP	
6 GAFMVLTA-	GVSPVAKTVQ	H	H MNG	D VMG	M REVKA	I ELPAG	K AVT	L DP	N GL	H	M	L EDAKHP	
7 AVFVELMNR	AVETPAACKVEL	H	I KEG	D VMK	M RQVDS	I EI	PAKG	T VL	K PGS	H	M	L FDLKAP	
8 AIFMVIENHGA	USAATDAAARVEL	H	HL	AGQDG	M VEVK	G FPVPAH	G SHAL	A RG	G D	H	M	L MGLTKP	
9 AVFVELMNRS	DNVRTIVAAE	T PAAAGK	VEL	H D	I KEG	D VMK	M RQVDS	I EI	PAKG	T VL	K PGS	L FDLKAP	
10 AAYFVVQNHG	KENDTLLGADTP	AA	SAEVHE	H VHKN	G MN	M QKVDS	V DV	APGK	D LR	F AP	G Y	M L MGKQ	
11 GAFMTLHA-	-HQDAKLVGVSS	V GT	-AEL	H EM	N VM	M RQIPS	L DL	P KMQ	D V	Q L	K F	M L MGKQ	
12 AAYFIENKG	DSDDRLLIGVDT	P	IAQQAQL	H E	H V	H AD	G L	M QHVEA	V D	P AGAKVS	F A	M L L D L KDR	
13 AVFVTFANRS	QDDINIVAAE	T PAAAGK	VEL	H D	I KDG	D VMK	M RQIDR	I T	L GAK	T ET	P GSL	M L F D L KTP	
14 MMGMIIVNEG	D EDPYDYLIGA	KTDIAQR	VEL	H K	V IEN	D VA	M VPQER	I EIP	P K	G KVE	F K H H G Y	M I I GLKRP	
15 AVYFTVKNGGR	MADRLTGDT	PNAAKTE	EL	H T	H L	R TG	E VM	M QHIDS	I EV	P AGAEV	K L T P G G H	M I F K P K R P	
16 GAFLTLTN-	TGDADRLLASADV	SET	VEL	H T	H I	MDG	T VM	M RQVDA	I DL	P AKGSVAL	R P G S	H M L I GLKAP	
17 AAFMVLMNHS	MDAVSLIKASSPQF	ERVEL	H R	SM	PVD	G VM	M VEQSR	I PV	P AQGKT	I L K P G D V	H M L MMGKA		
18 AGYLTISNTG	DEDITLT	TEAAT	TSLS	DRTEL	H T	/ETTESG	A CM	M VPVDD	I PI	P AGE	T VEL	A S G G L	
19 AAYMVIVN-	NG	VKAESD	VAA	SVEL	H N	V	M QMRQVEA	I EV	P AN	K P G G F	H M L VLDIADP		
20 GAFMRITA-	D	VDVASP	VAKTVQI	H E	H M	S MS	M QRVNS	V DL	P AG	D S D G Y	H M F		
21 AAFMTIH-	H G	VSAQAEI	SKS	VEL	H T	H V	Y R	M RKVEA	L ALP	E H	K P G G D	H M F	
22 AGYMEIENDA	K	VSVSSPA	AAERVEL	H N	V	A C	M RQVDG	V V	A P V A D	A P G G Y	H M F		
23 GIGEIIMNHS	D	VGFSSS	INAKMNC	H L	H E V	-	V M	K E K G F	V I P A G	E P G K D	H M F		
24 GAFMTVTA-	D	LSVASPA	AKDVQI	H E	H M	S M	M GPVKS	V DL	P AG	D P N G Y	H M F		
25 AAYFTLENKG	DSADRLISVDTPI	AGQQL	H E	H V	H A D	G L	M QHVQA	V D	I P A G A K V V	F A P M A	H M L L D I KDR		
26 GAFMTLQA-	-ADGARL	VGI	STPAA	ARA	E I	H E	M K	M E G	D V	M R	P K V V F	H M L Q D L K Q P	
27 AVFATLVNN	S	GFTAD	VDAAS	F	E V	H E V	M	M QEKPG	G F V I P A	A P G G D	H M L M C L A N D		
28 GAFMRLTA-	-C	VGARS	AI	A E H T	E V	H E M	T M	M RQIAA	A L E L P	A P G G Y	H M L M C L A N D		
29 AGFGQFHNGCA	A	TAANSS	SVFAD	V S L	H E T	T	V S R	M R A V P E	L A L P	A P G G L	H M L M C L A N D		
30 GVYLSLENQGK	V	VASTP	R A E	V E	H T	H V	V M	M R K V E G	G V A V A P	A P G S Y	H M L M C L A N D		
31 GAFMTITS-	-SSDSK	LLSA	QSPVAK	I V Q L	H L	ST MKN	D V	M S	M QPVF	I D L P	A G K P V T L	H M L I D L V K Q	
32 AIFGTLNTT	-DEEISL	TGF	E A S V D	AAA	YEI	H E V	V W	D G -	-	V M	R E K E G	L T I A A G D	
33 GGYVTLRNQS	STPDR	L V A V E	S P A S	A Q V E	L H D	M D R	M D G	G Y M	K M R R L	D D G L A	P P G E T V V L	G P G G T	
34 GAFMHITS-	-STD	SKL	L V T	S P V A	K T V Q I	H Q M	S M K G	D	M S	M Q R V S S	V D L P A G K P V V F	D A N G Y	
35 GGYLT	T	T N T G	E P D R	L T G	G S L E A A	G R G E L	H T	M S	M E G	H V M K	M A P L P N G	L E I A P G K T V T L A P S G N	
36 GGYLT	T	T N N G	S A P D	R L L S	S D I	S D K A E L	H E	M G V K D	-	G M T	M R P V V G	G L E I A P G G K V V L G	G S Y
37 AIYMVL	E N Q T	S S P I	V V N Y	I	T T I A	D R V E	V H Q	H I	H E D	G M M K	M R Q A K H	-L K I P A K Q K Q L F E	R G G Y
38 GAFMKI	H N D E A	K Q D F L L	G G S	S P V A	D R V E	V H T	H I	N D N	G V M R	M R E V E	G G V P L E A K S	S V T E L K P G S Y	H M F M G M L K Q P
39 GAFMKI	H N D E A	K Q D F L L	G G S	S P V A	D R V E	V H T	H I	N D N	G V M R	M R E V E	G G V P L E A K S	S V T E L K P G S Y	H M F M G M L K Q P
40 AAYFI	I H N G	G K T A D	R L L S	V D S	M I A P	T A E L	H E	H V M	M Q Q G	D L M K	M Q Q V P N	V A I P A G G N V T F A P M A Y	H M F M N P T D R
41 AGFLT	I	T N E G	D S A D E	L T S	V T	S E A G E	-V T V	H E	I D G T	-	M K E V D R	I E V P A H G Q L V F	K P G Q D H L
42 GCYVTLT	-	-A G V D	D R L V A	V E	T T A A	A R G E I	H T	M S	M D G	G V M R	M R K L A D	G L A P G K A V A L K P	G Q D H L
43 GGGFVVRNGG	S ADD	R L L A V E	S P A	G R V E L	H E	M T	M E N	D V	M K	M R K L E D	G I A V P A G G	T V E L K S C O G L	H M F M E V K K P
44 GGFFQLTNHG	D T E	D A L I A A A	E S P I A	G R V E I	H T	H T	N E D	G V M	M K	K I D E	-V R V A A H E	T V V F K P G S	H M F M E V K K P
45 GGFLTI	A N D G	K K A D	K L V S	V S A P	G V R	V E I	H E	M T	M Q D	-Q I M K	M R K L E G	G L D L P A G K T M Q L K S G	G S Y
46 GGFLTI	A N D G	K K A D	K L V S	V S A P	G V K R	V E I	H E	M T	M Q D	-Q I M K	M R K L E G	G L D L P A G K T M Q L K S G	G S Y
47 GGYLT	E N R G	H A P E R L	Q T A S	A A H A	L R T E I	H E	H M	A V N N	-G V M T	M R P L I D G L V I A P G Q I	V K L A P G G C	H M F M E V K K P	
48 AAYFV	V H N N G	Q A D D	R L L S	V D S P I	S D D A	Q L	H E	H A M S A T	G A M	M Q Q V P S	-V V V P A G K D L T F A P G A Y	H M F M E V K K P	
49 GGYVT	I K N T G	D S D D	K L V G I E	S S A A G	R A F E I	H E	H M	A V N N	-D V M H	M R K L D E G I V I P A G Q	T V E L K P G C	H M F M E V K K P	
50 GGFLTI	E N K G	G S A D	R L V S	G T A D	I A G K	V E I	H E	H M	M D N	-G V M K	M R P L D K G L P I E P G K T	V T L M P G S Y	H M F M E V K K P

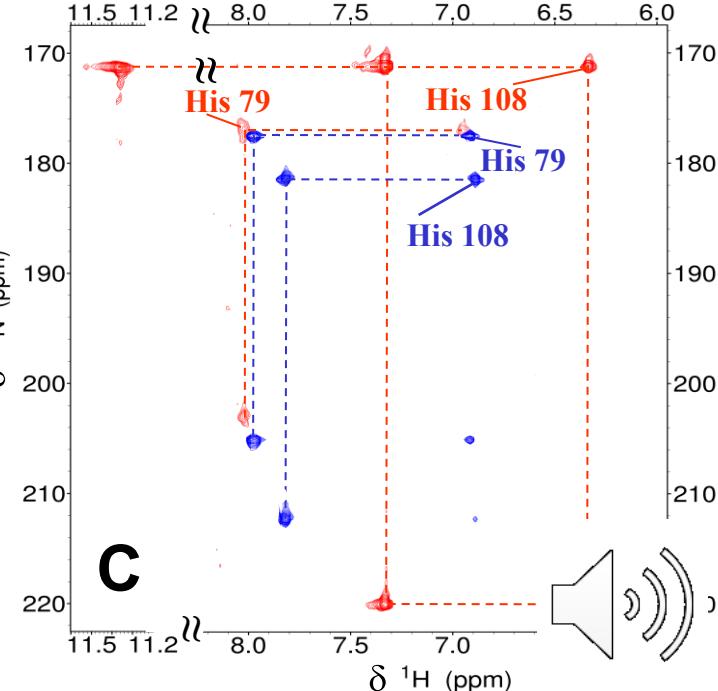
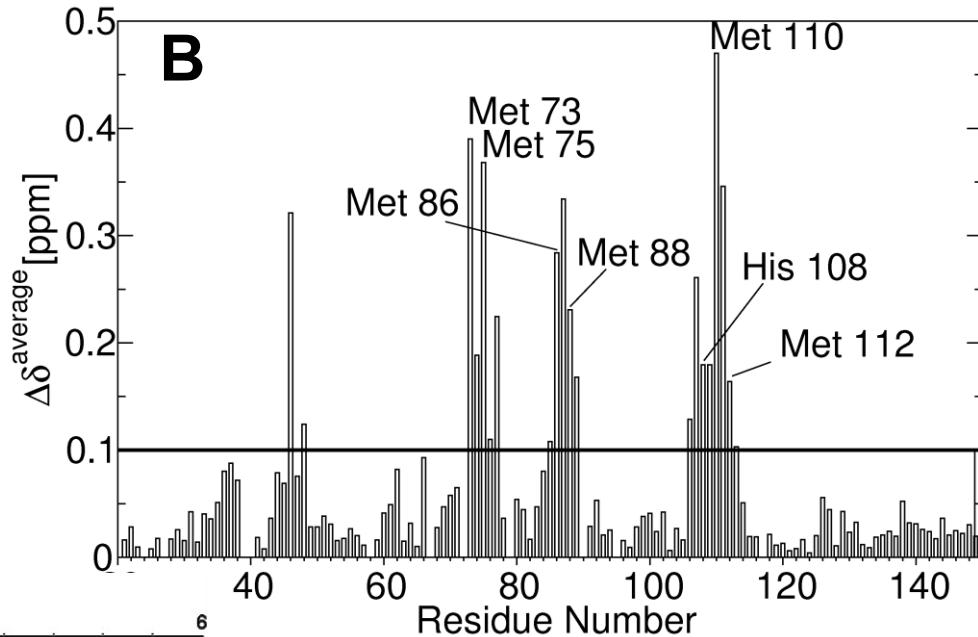
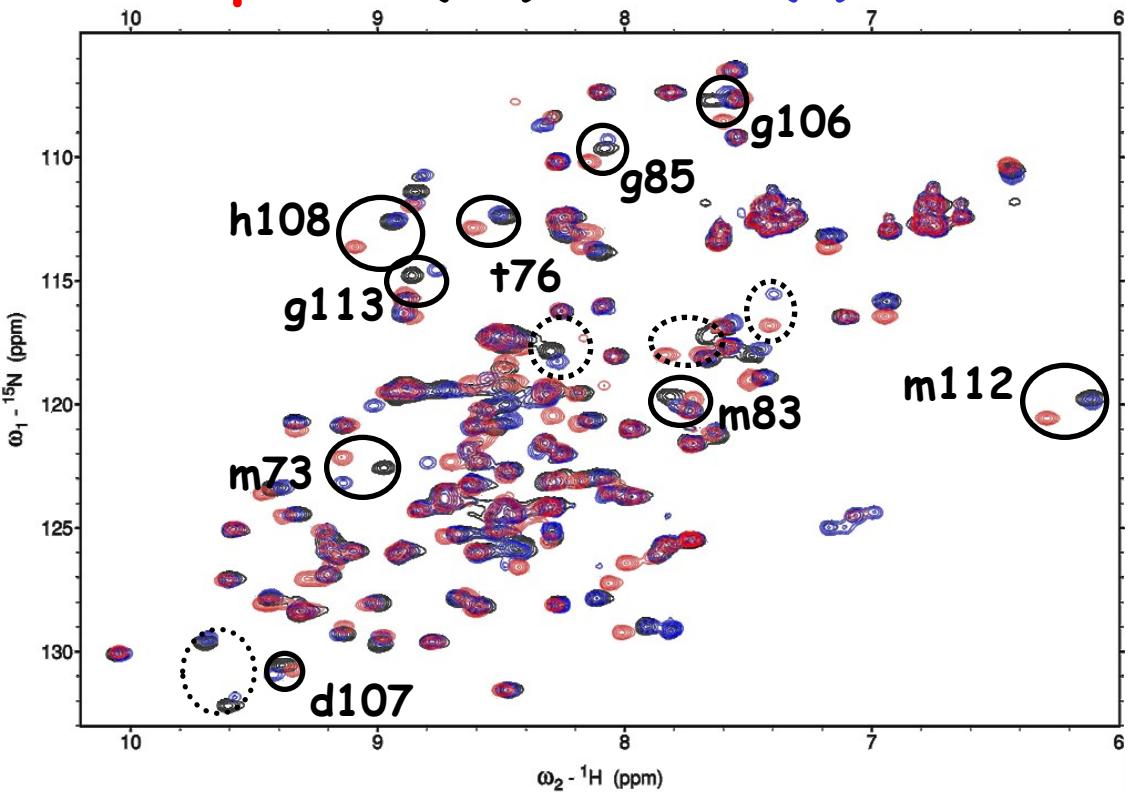
*

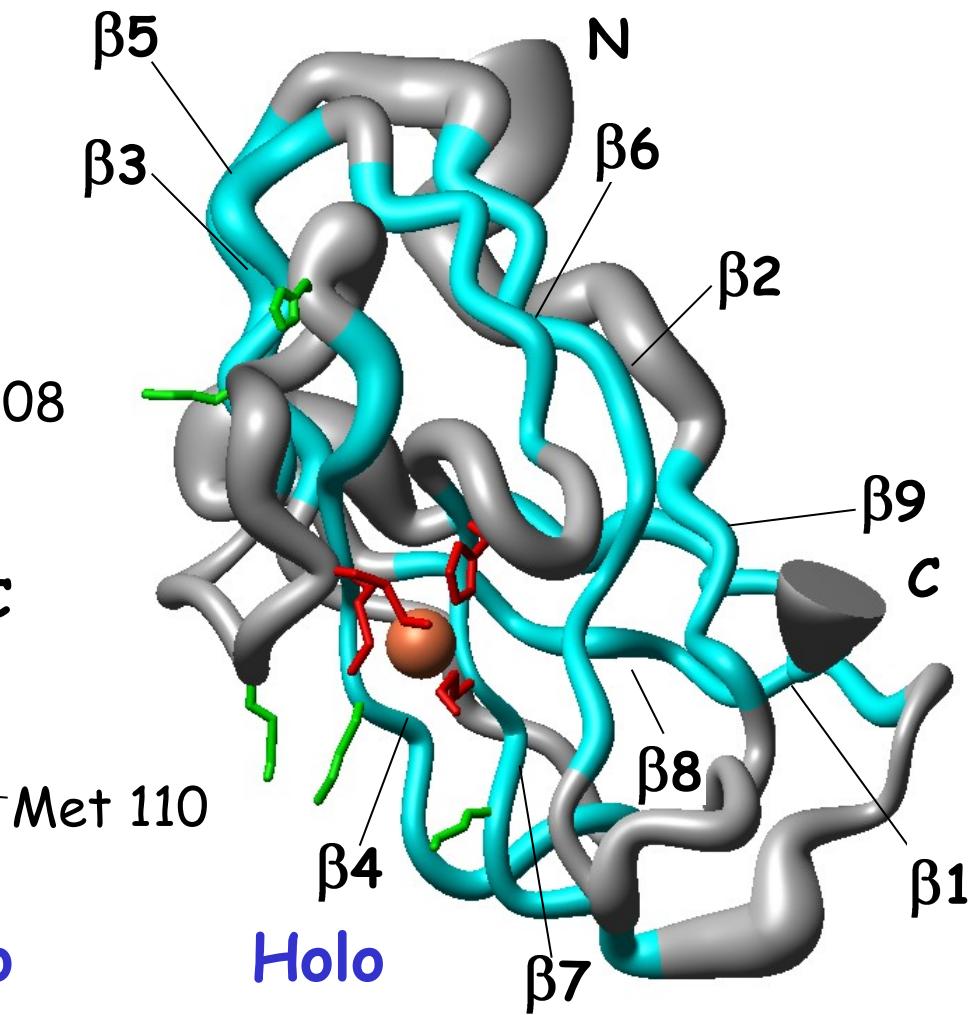
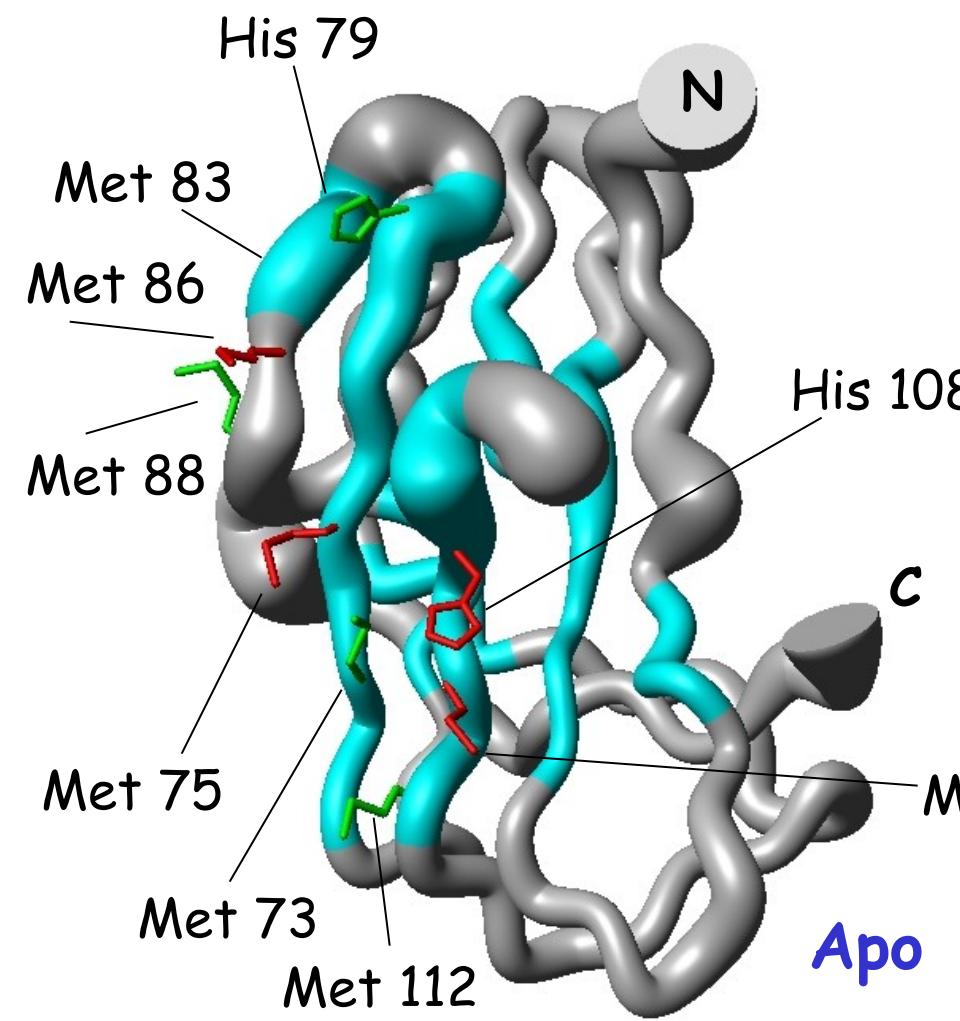


Interaction of DR1885 with copper

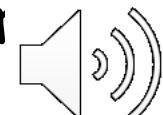
-titration (A,B)
 2J HSQC (C)

Apo, **Cu(II)** and **Cu(I)**



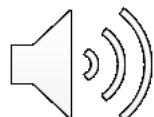


cyan - β -sheets
grey - random

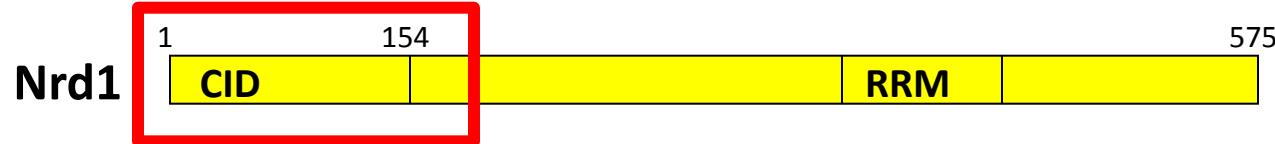
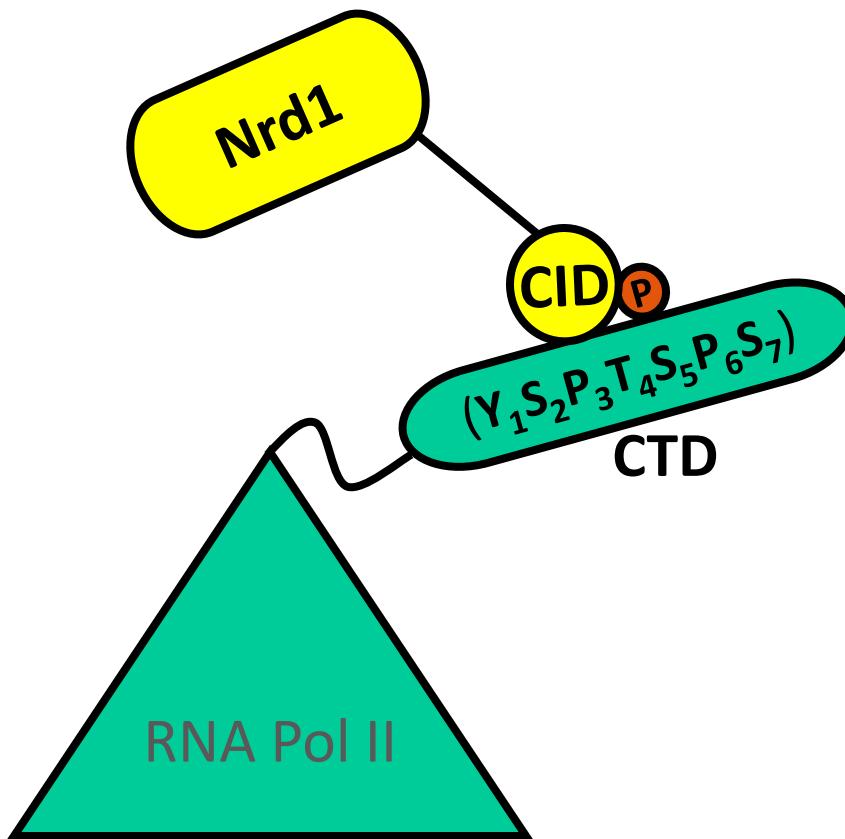


Protein – peptide interaction

fast exchange case



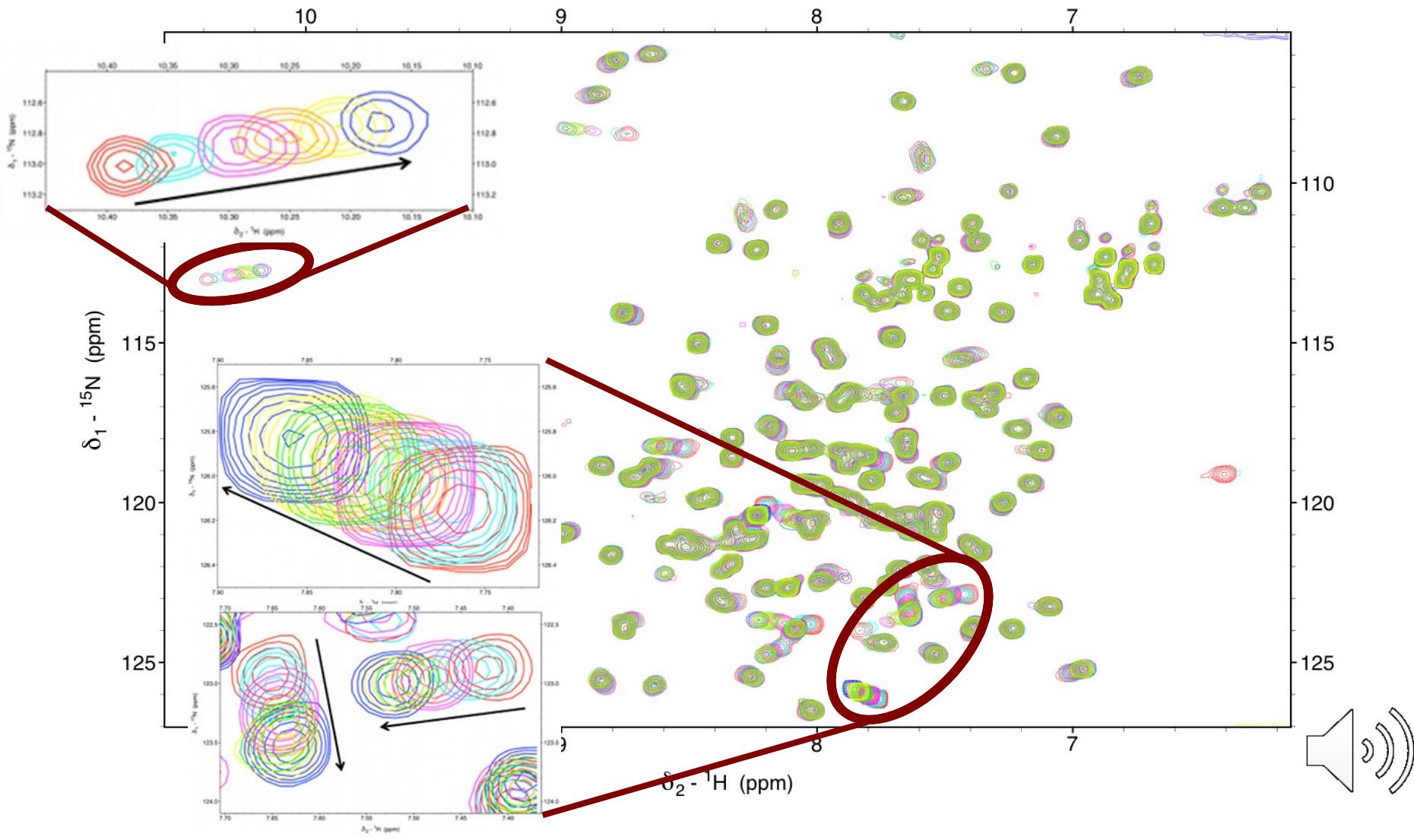
Interaction of Nrd1-CID with CTD



RRM: RNA recognition motif; **CID:** CTD interaction domain; **CTD:** C-terminal domain



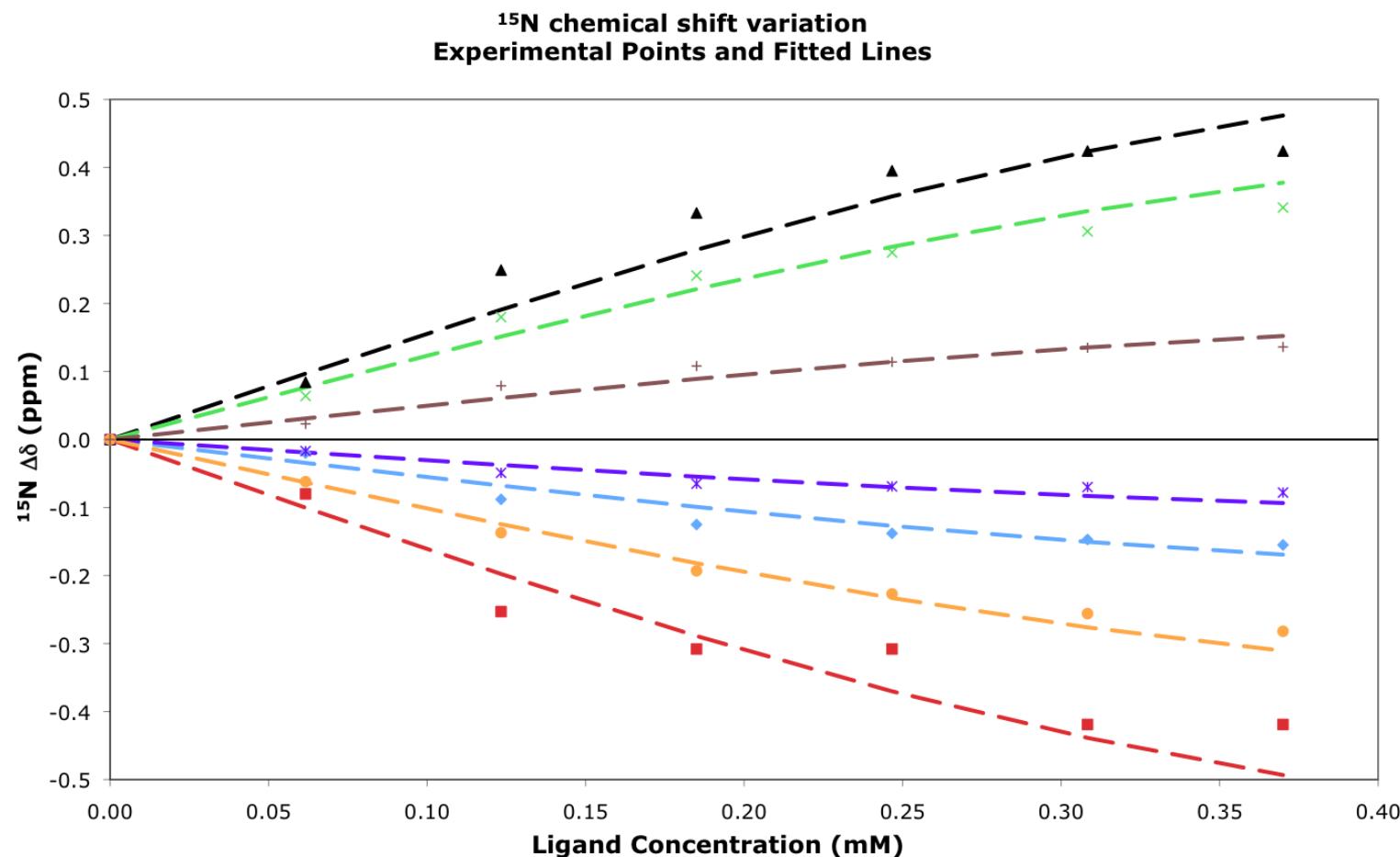
Interaction of ^{15}N enriched CID with unlabeled CTD-Ser5P in n -steps, $n=6$ in our case
 - peaks corresponding to the interacting residues of CID change their chemical shift (position in the spectrum) => interaction surface, binding constant, stoichiometry



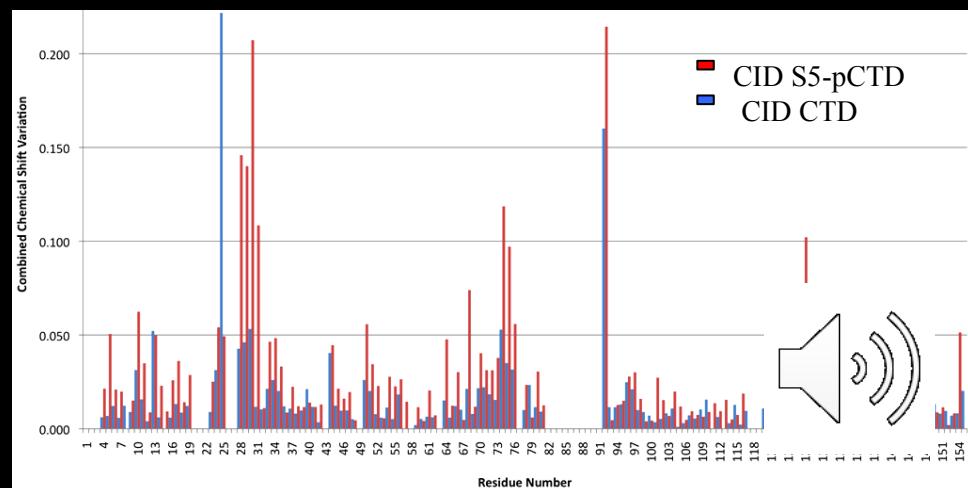
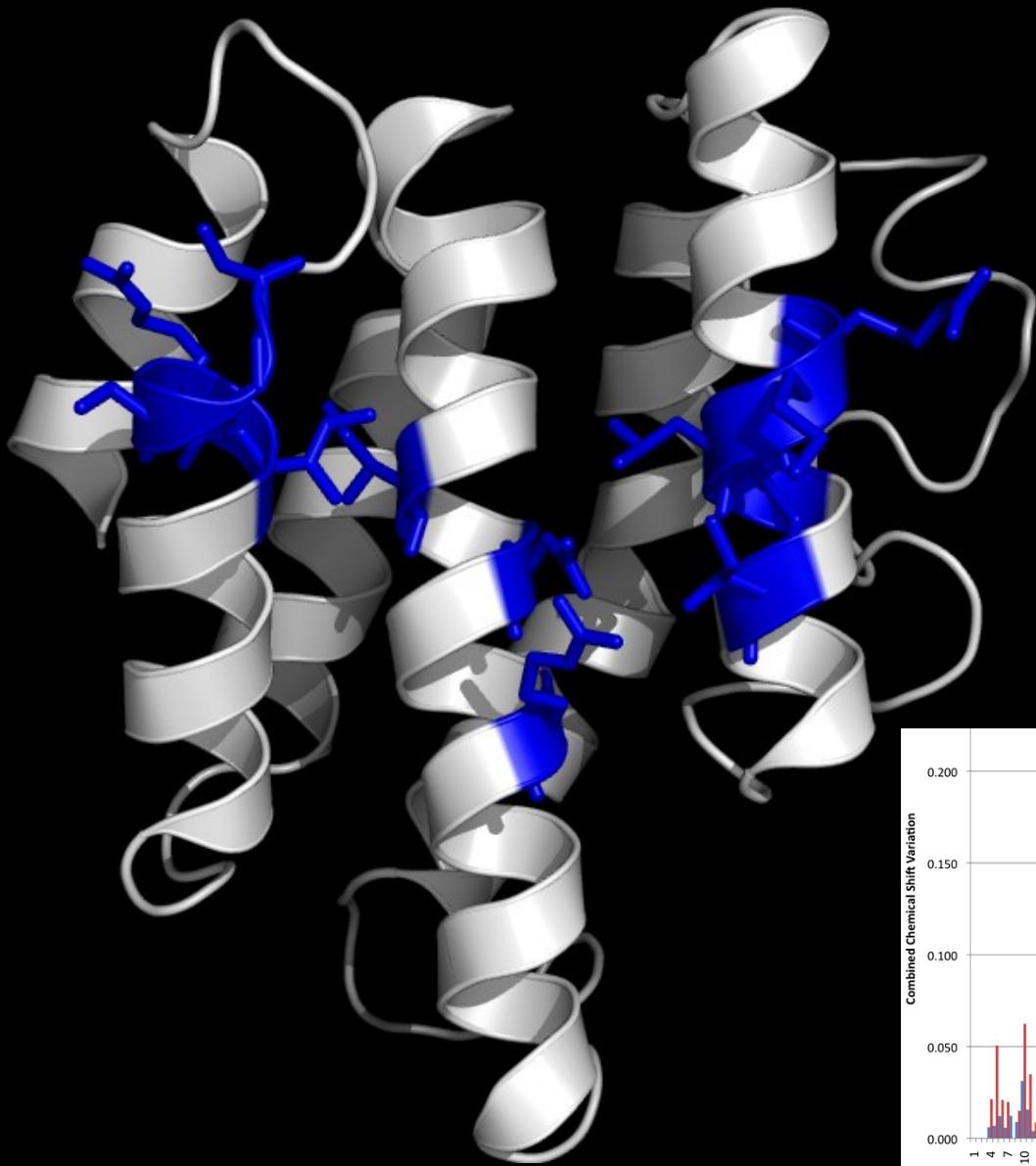
Interaction of Nrd1-CID with CTD

NMR Titration:

- ~ 0.6 mM ^{15}N enriched CID + ~ 0.8mM (YSPTpSPS)₂
- ~ 0.6 mM ^{15}N enriched CID + ~ 0.8mM (YSPTSPS)₂
- $\mu\text{M-mM}$ range of interaction ->
 - > fast exchange regime on NMR time-scale
- NMR-derived $K_d = 0.080\text{mM}$ and 35mM

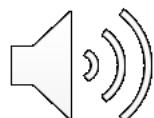


Nrd1 CID interaction surface — CID residues experiencing the largest chemical shift variations upon the interaction with 5-phospho-Ser CTD shown in blue with side-chains in stick representation

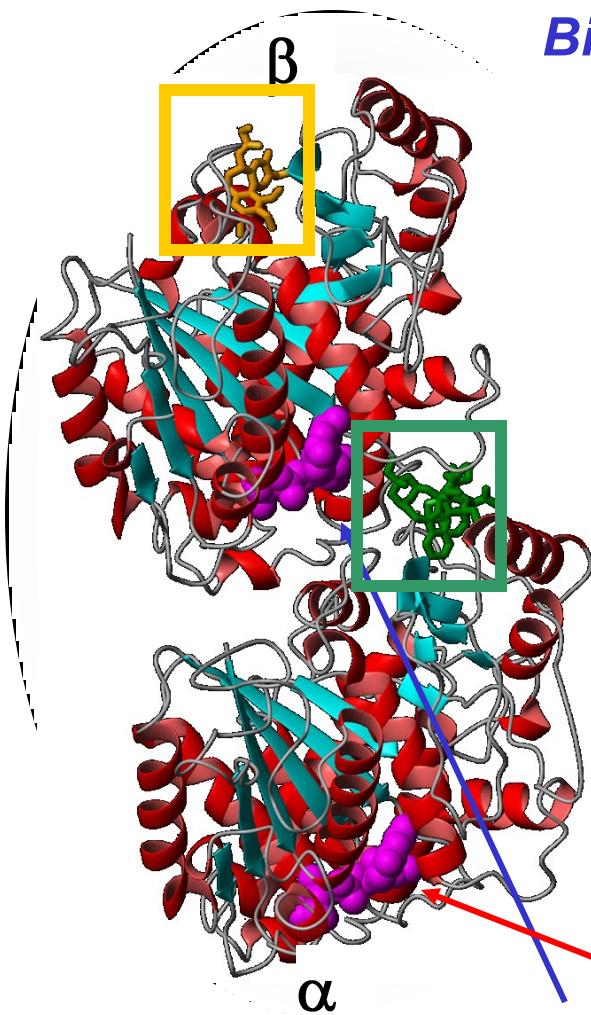


Protein – peptide interaction

drug-receptor case



Tubulin - successful target for anticancer therapy



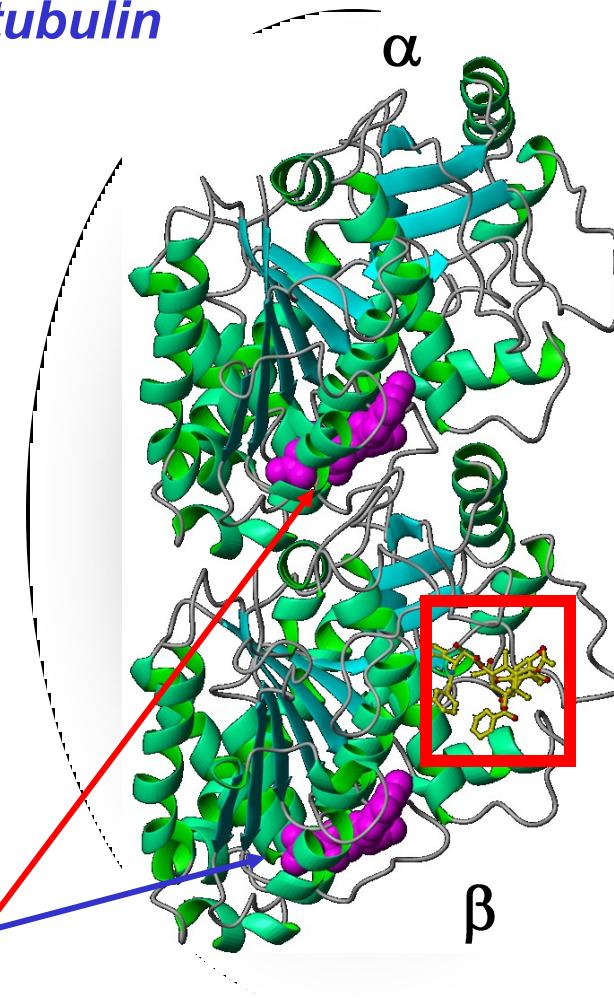
Binding sites on tubulin

Colchicine

Vinca

Taxane

GTP

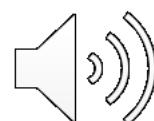
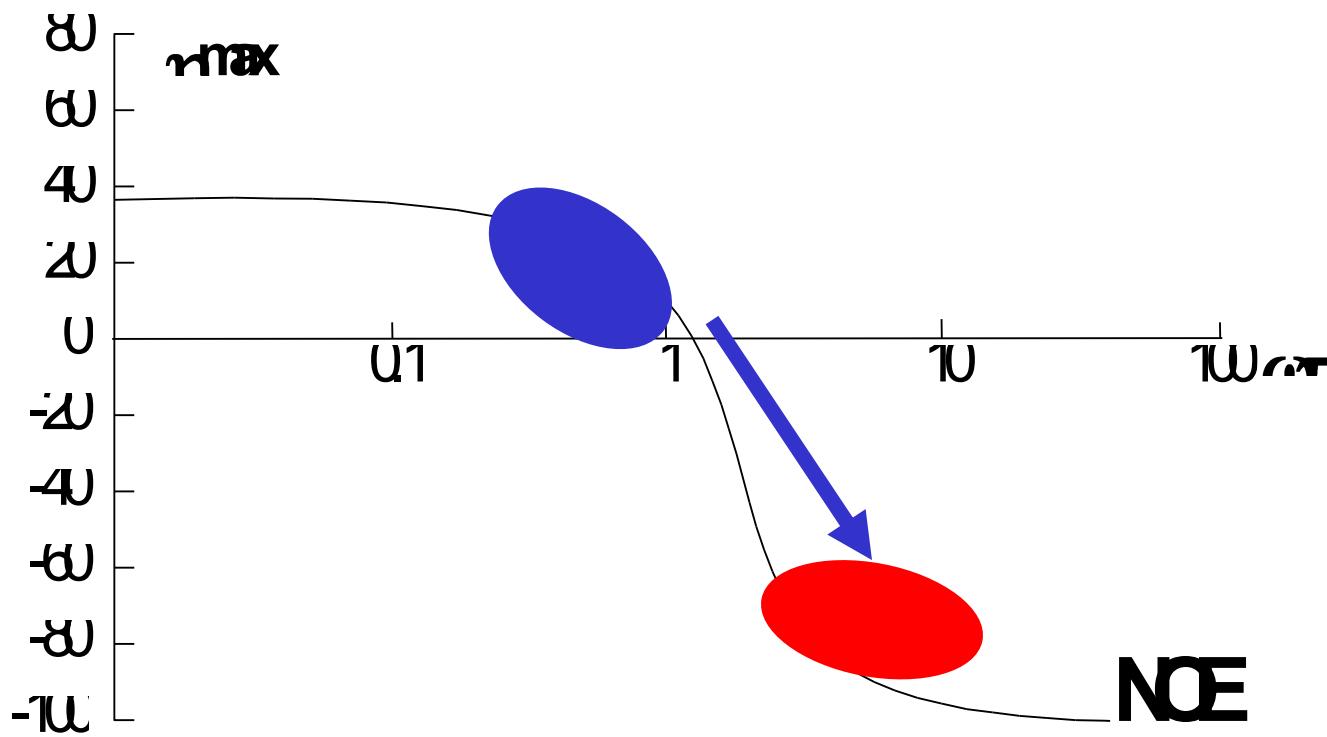


Transferred-NOE

$$\text{NOE} = p_{\text{bound}} \cdot \text{NOE}_{\text{bound}} + p_{\text{free}} \cdot \text{NOE}_{\text{free}}$$

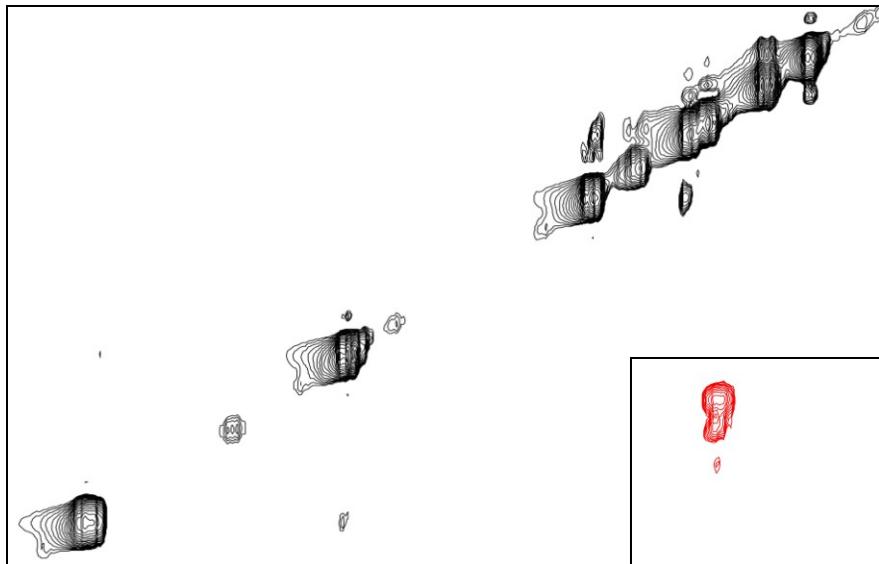
$\tau_{c,\text{bound}} \gg \tau_{c,\text{free}}$ (and $p_{L,\text{free}} \gg p_{L,\text{bound}}$)

$$\text{NOE}_{\text{bound}} > \text{NOE}_{\text{free}}$$



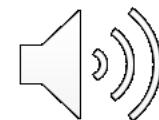
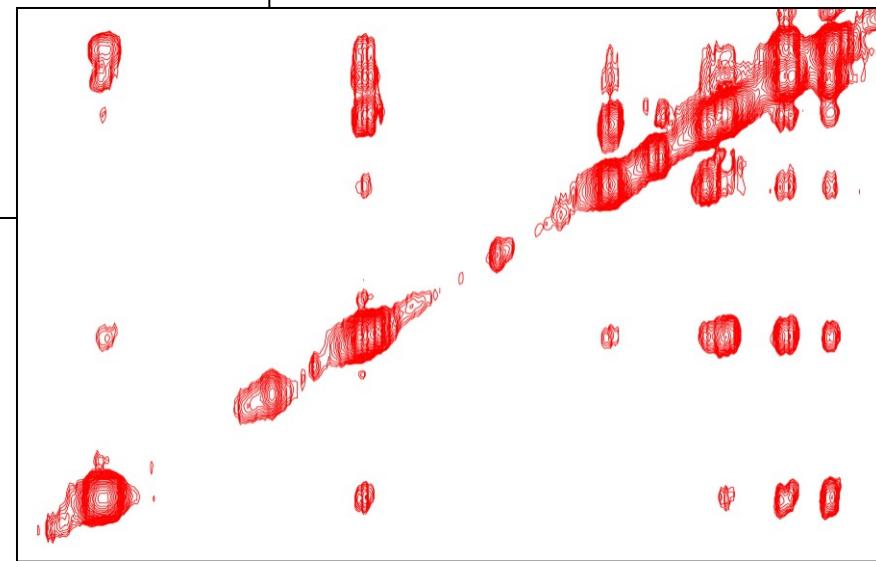
Transferred NOE Experiments

tr-NOESY~500 μ M tubulysin (TBS) **without** and **with** ~10 μ M tubulin

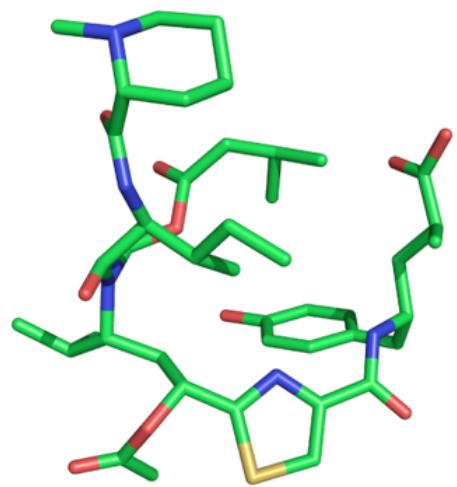
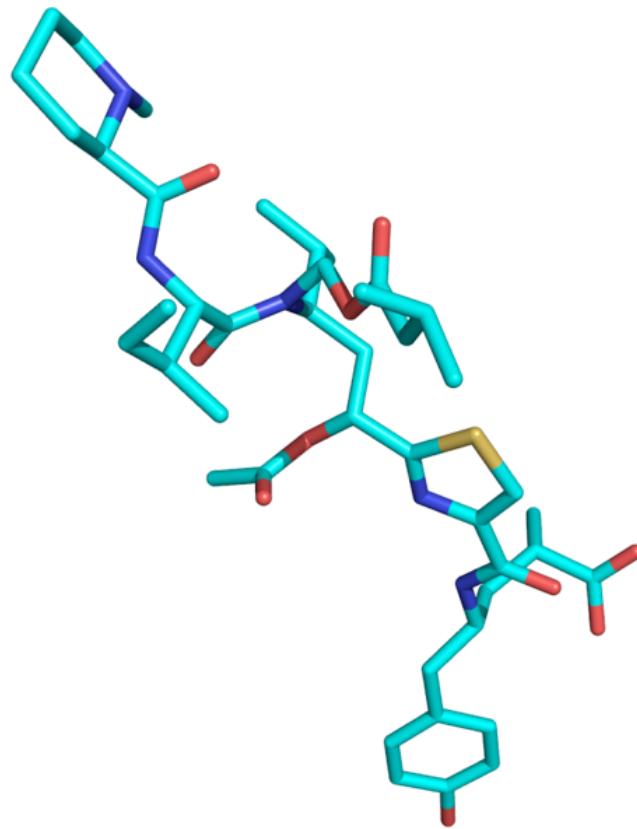


↑
tbs w/o tub

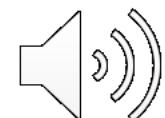
tbs:tub 50:1



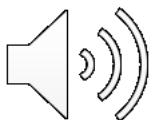
900MHz,
mixing time=100ms

A**B**

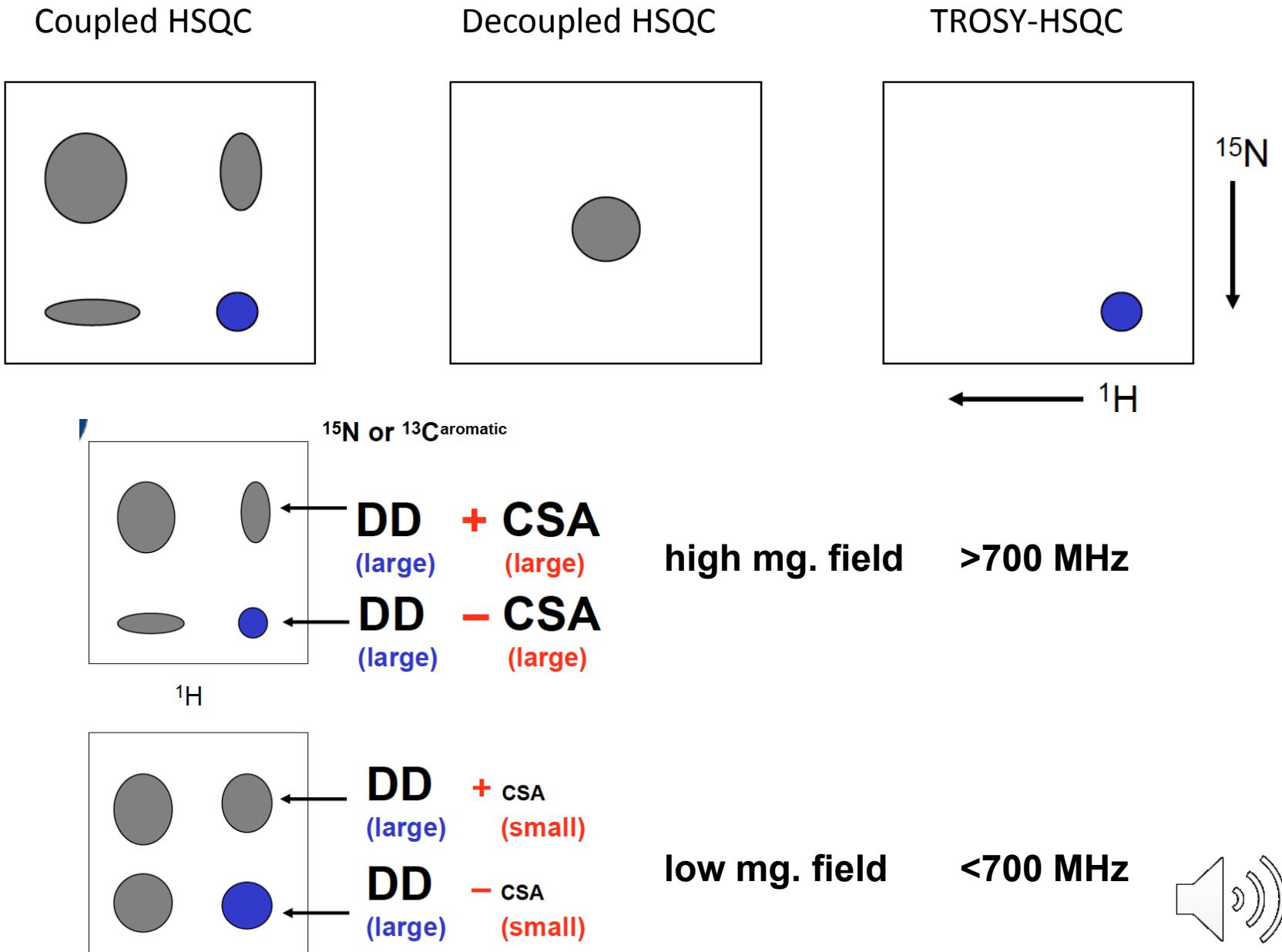
Conformation of the tubulin-bound - NMR (A) and free – X-Ray T²⁰⁰ (B)



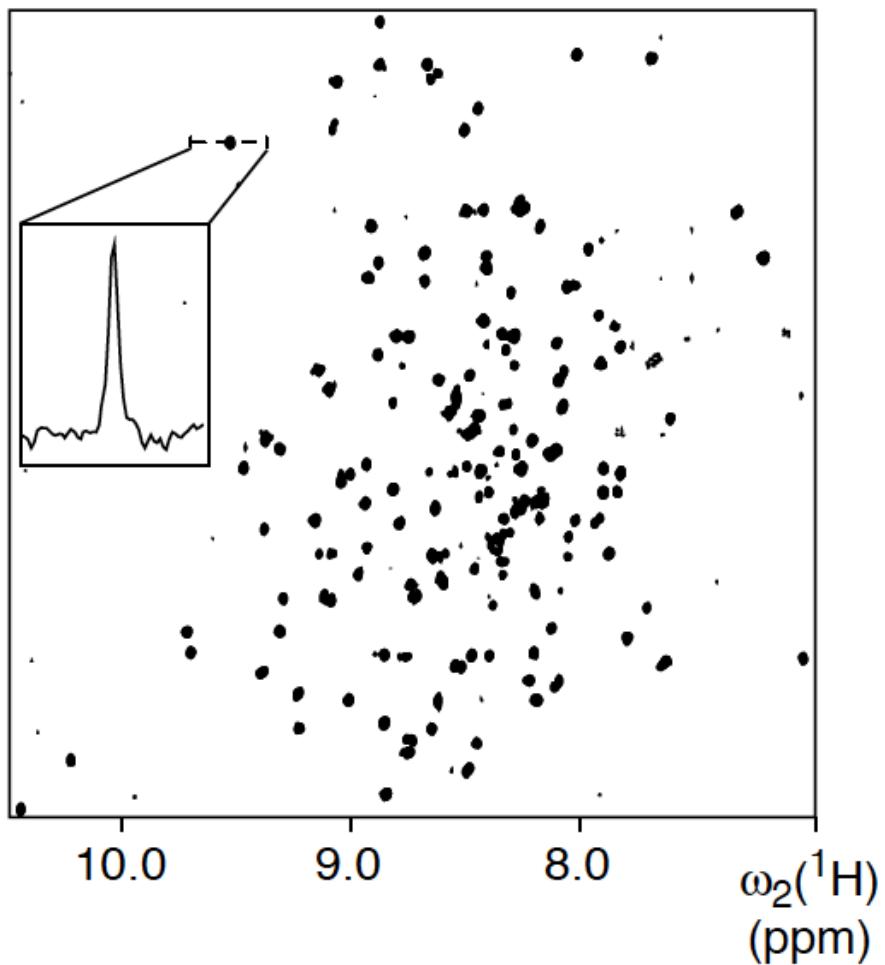
Large biomolecules and their interactions



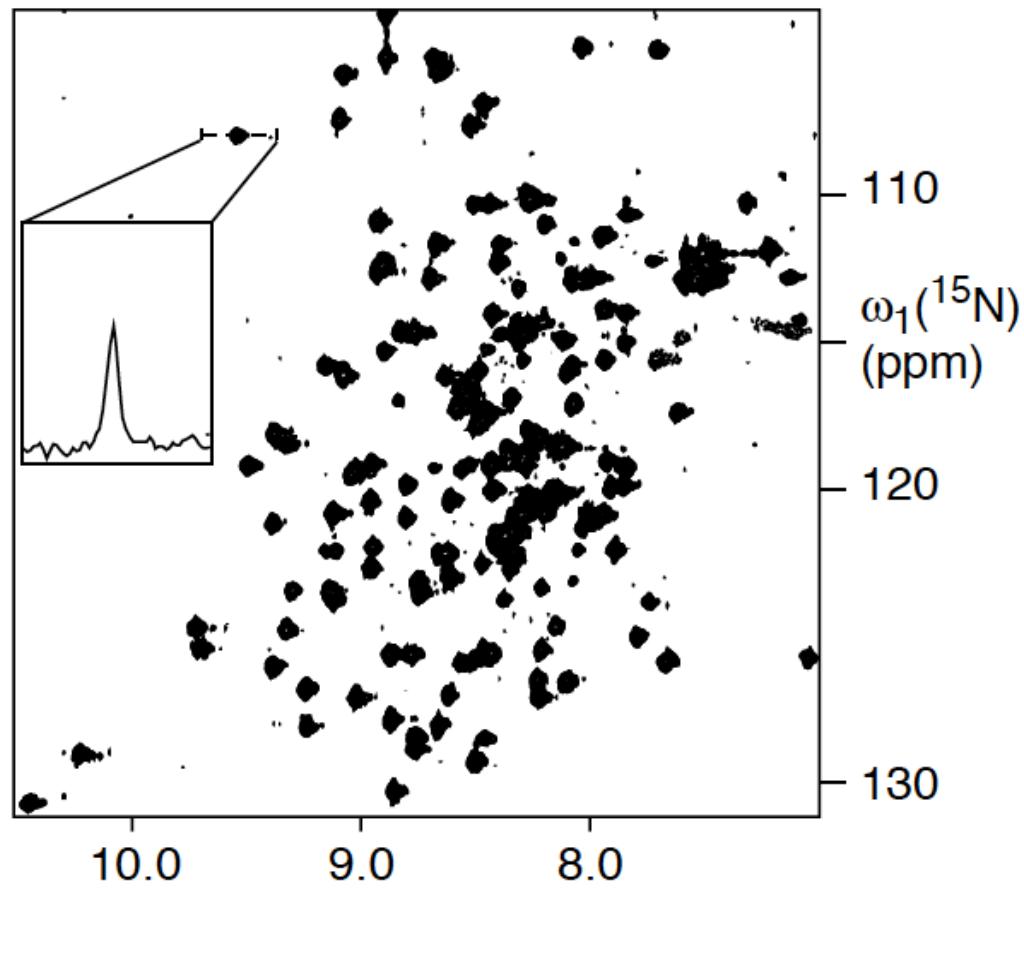
TROSY-based NMR experiments to study complexes up to 900kDa



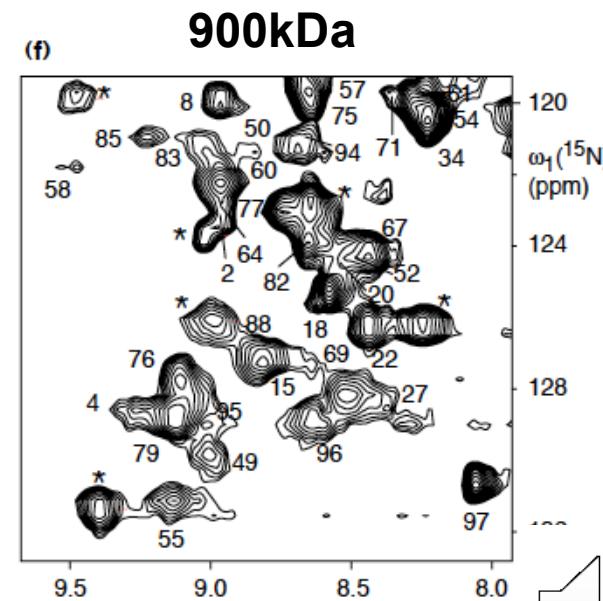
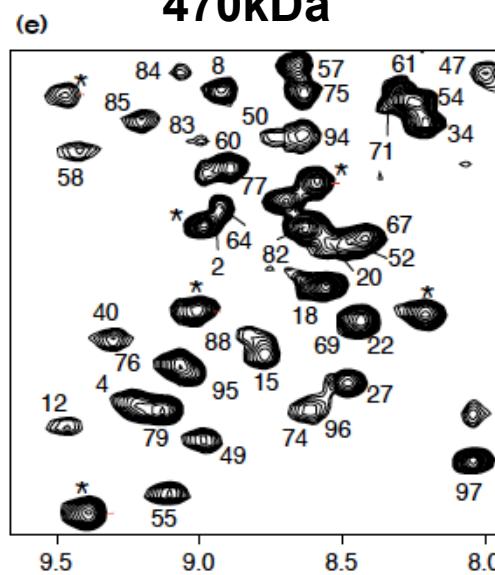
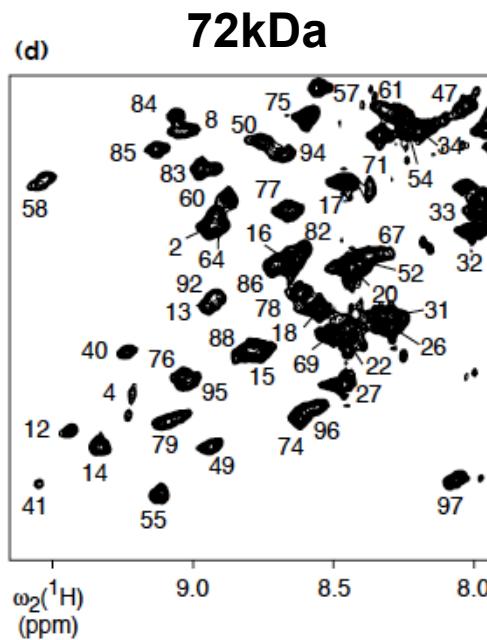
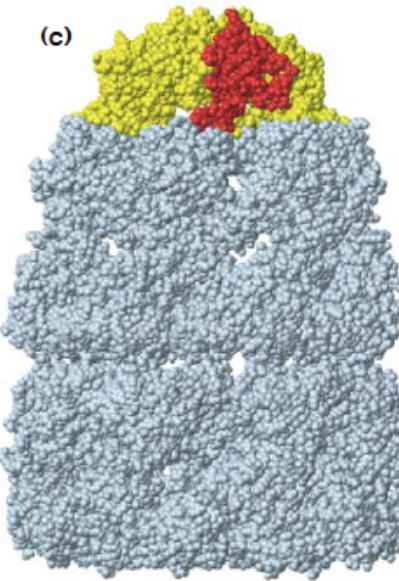
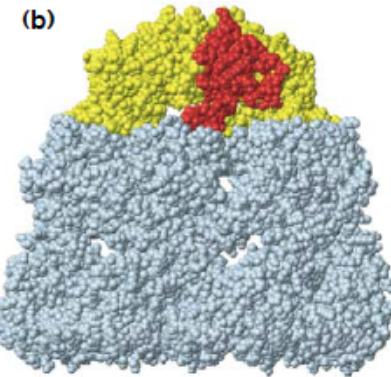
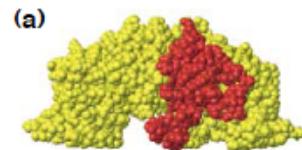
(a) 2D [$^{15}\text{N}, ^1\text{H}$]-TROSY



(b) 2D [$^{15}\text{N}, ^1\text{H}$]-COSY

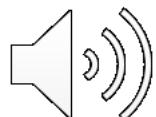


70% ^2H , ^{13}C , ^{15}N sample in H_2O , 700MHz
CRINEPT-TROSY



Summary

NMR is a robust tool for studying structural properties and interaction properties of biomolecules of variable molecular size at various levels of resolution.





Core Technologies for Education and Innovation in Life Sciences



Co-funded by the
Erasmus+ Programme
of the European Union

Thank you for your attention

For Application to Protein Characterization

by
Radovan Fiala, Karel Kubicek, and Pavel Kaderavek
CEITEC, Masaryk University

innocore-project@unitn.it

This project received funding by the European Commission under the GA n° 2019-1-IT02-KA203-063391

[www.innc
EU](http://www.innc.eu)

