Workshop:

Scientific Writing Skills

Lecture title:

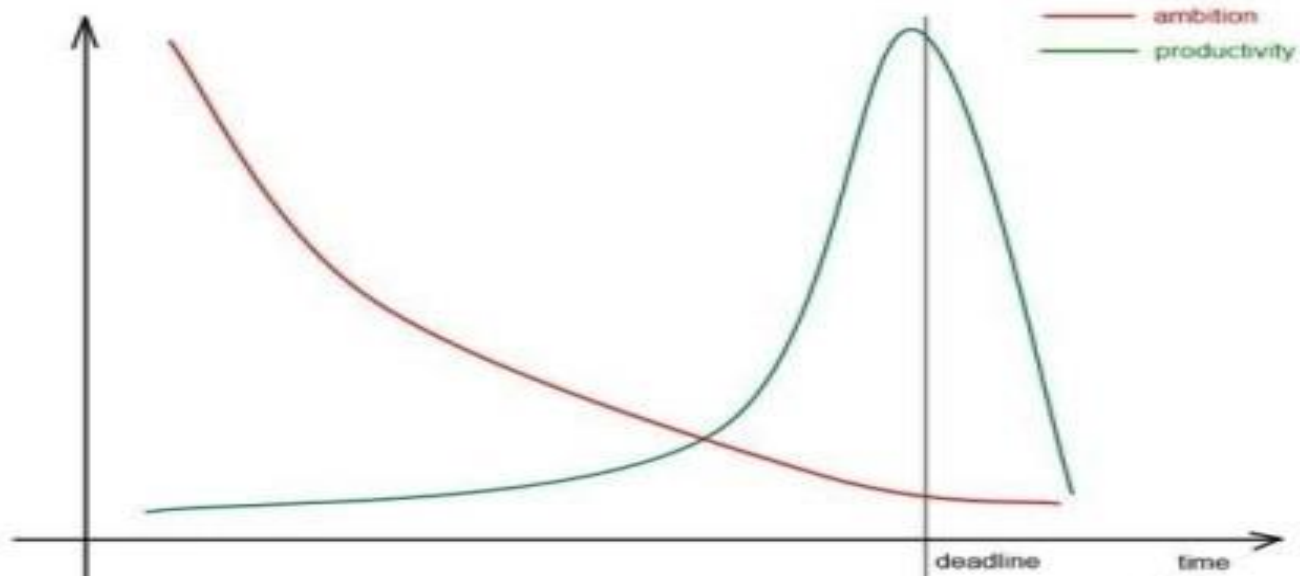
Success of your project depends on you (and so does its failure)!

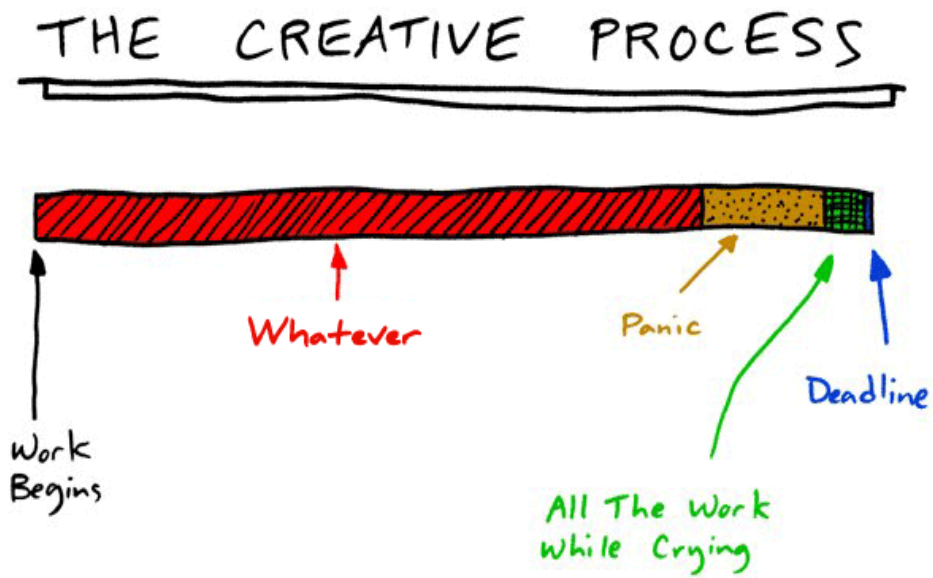
Presenter:

Karel Kubíček

Date:

September 19, 2016





Toothpaste For Dinner.com

THE LAWS OF HERMAN

1. Your vacation begins after you defend your thesis.

2. In research, what matters is what is right, and not who is right.

3. In research and other matters, your adviser is always right, most of the time.

4. Act as if your adviser is always right, almost all the time.

5. If you think you are right and you are able to convince your adviser, your adviser will be very happy.

6. Your productivity varies as (effective productive time spent per day)^{1,000}.

7. Your productivity also varies as 1/(your delay in analysing acquired data)^{1,000}.

8. Take data today as if you know that your equipment will break tomorrow.

9. If you would be unhappy to lose your data, make a permanent back-up copy of them within five minutes of acquiring them.

10. Your adviser expects your productivity to be low initially and then to be above threshold after a year or so.

11. You must become a bigger expert in your thesis area than your adviser.

12. When you cooperate, your adviser's blood pressure will go down a bit.

13. When you don't cooperate, your adviser's blood pressure either goes up a bit or it goes down to zero.

14. Usually, only when you can publish your results are they good enough to be part of your thesis.

15. The higher the quality, first, and quantity, second, of your publishable work, the better your thesis.

16. Remember, it's your thesis. You (!) need to do it.

17. Your adviser wants you to become famous, so that he/she can finally become famous.

18. Your adviser wants to write the best letter of recommendation for you that is possible.

19. Whatever is best for you is best for your adviser.

20. Whatever is best for your adviser is best for you.

These laws were inspired by the 'Laws of the House of God' from *The House of God* by Samuel Shem (Richard Marek, 1978), which provided a somewhat different brand of advice to medical interns. The author thanks Jonathan Spanier, Yigal Komem and other colleagues for suggestions.

Irving P. Herman, *Nature*, 455, **2007**, 228

What makes a good (PhD) student

a) Choose a supervisor whose work you admire and who is well supported by grants and departmental infrastructure.

b) Take responsibility for your project.

c) Work hard — long days all week and part of most weekends. If research is your passion this should be easy, and if it isn't, you are probably in the wrong field. Note who goes home with a full briefcase to work on at the end of the day. **This is a cause of success, not a consequence.**

d) Take some weekends off, and decent holidays, so you don't burn out.

e) Read the literature in your immediate area, both current and past, and around it. You can't possibly make an original contribution to the literature unless you know what is already there.

f) Plan your days and weeks carefully to dovetail experiments so that you have a minimum amount of downtime.

g) Keep a good lab book and write it up every day.

h) Be creative. Think about what you are doing and why, and look for better ways to go. Don't see your PhD as just a road map laid out by your supervisor.

i) Develop good writing skills: they will make your scientific career immeasurably easier.

j) To be successful you must be at least four of the following: ***smart***, ***motivated***, ***creative***, ***hard-working***, ***skillful*** and ***lucky***. **You can't depend on luck, so you had better focus on the others!**

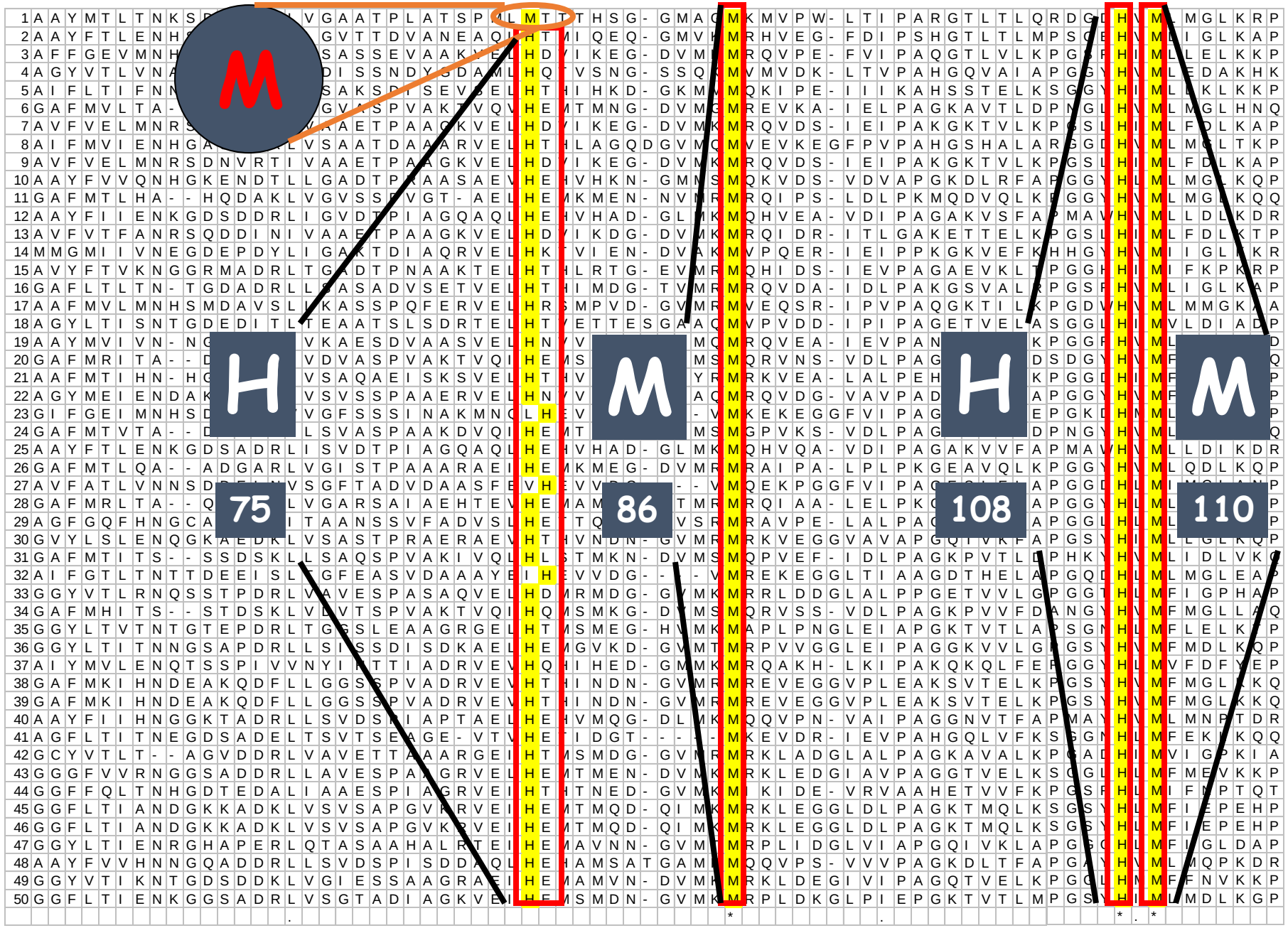
Nature 441, 252 (10 May 2006) | 10.1038/nj7090-252b

Example #1

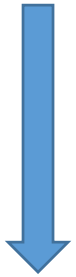
Two issues

1) Messy Colleague
2) Tough supervisor

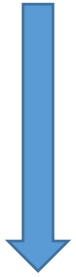
A BLAST search over all non-redundant GenBank genomes



BLAST – 52 sequences



Selected 4



For NMR the one with best expression and 15N HSQC peaks distribution

What do you need to perform NMR resonance assignment

1) Primary sequence

MQQDDDFQNF VATLESFKDL KSGISGSRIK KLTTYALDHI DIESKIISLI
IDYSRLCPDS HKLGSLYIID SIGRAYLDET RSNSNSSSNK PGTCAHAINT
LGEVIQELLS DAIAKSNQDH KEKIRMLLDI WDRSGLFQKS YLNAIRSKCF
AMDLEHHHHHH

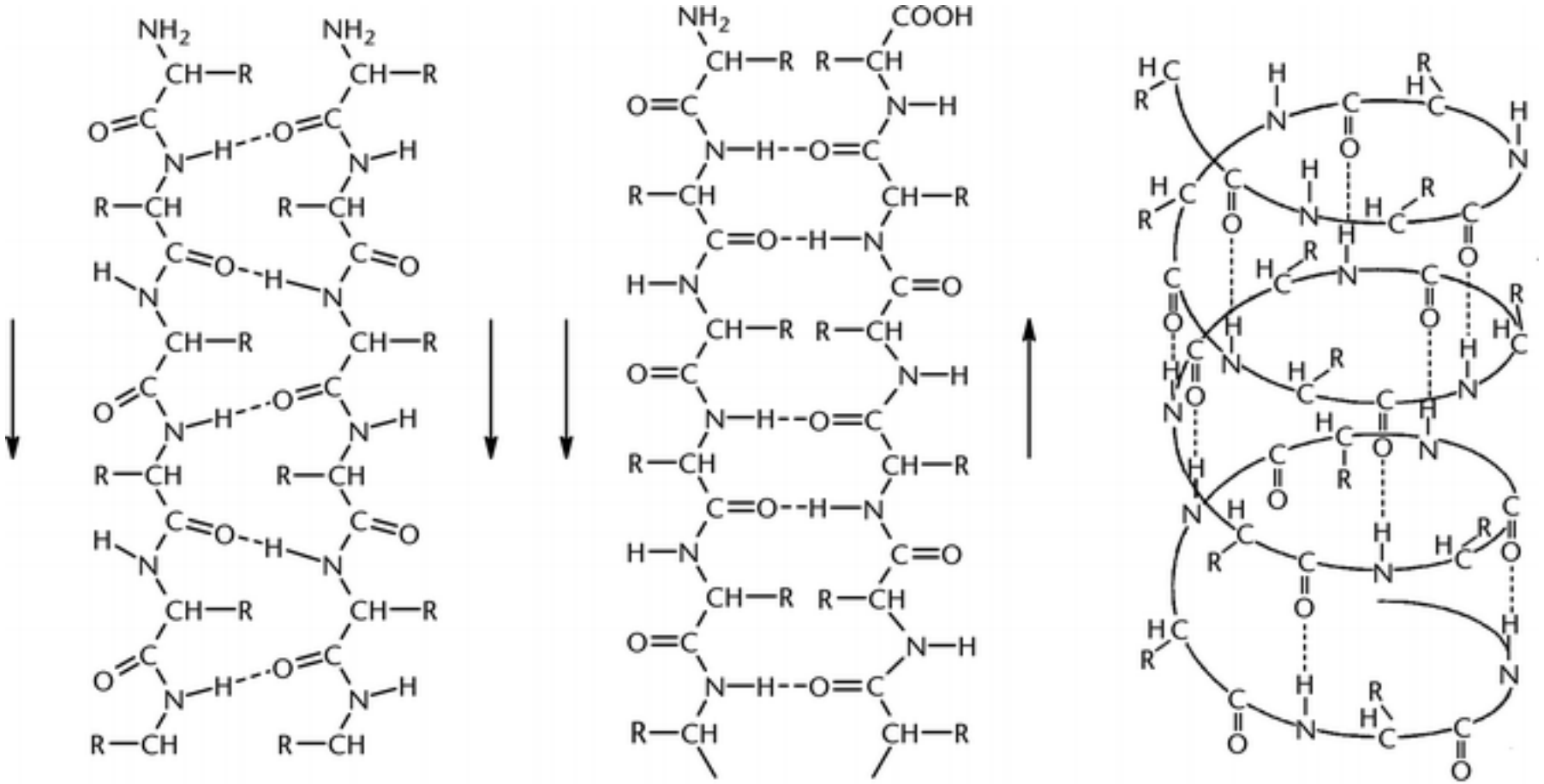
2) Chemical shifts of Cα/Cβ (*vide infra*)

3) Secondary structure

4) Exact peak positions

5) ^{13}C + ^{15}N isotopically enriched protein ~60-200 aa (<35 kDa)

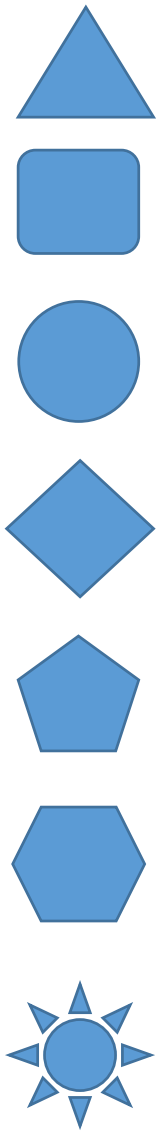
Secondary structure organization in proteins



N- ⟶ C-terminus

H-bond

Some AAs have unique chem. shift, some resonances are degenerated



A

G

S / T

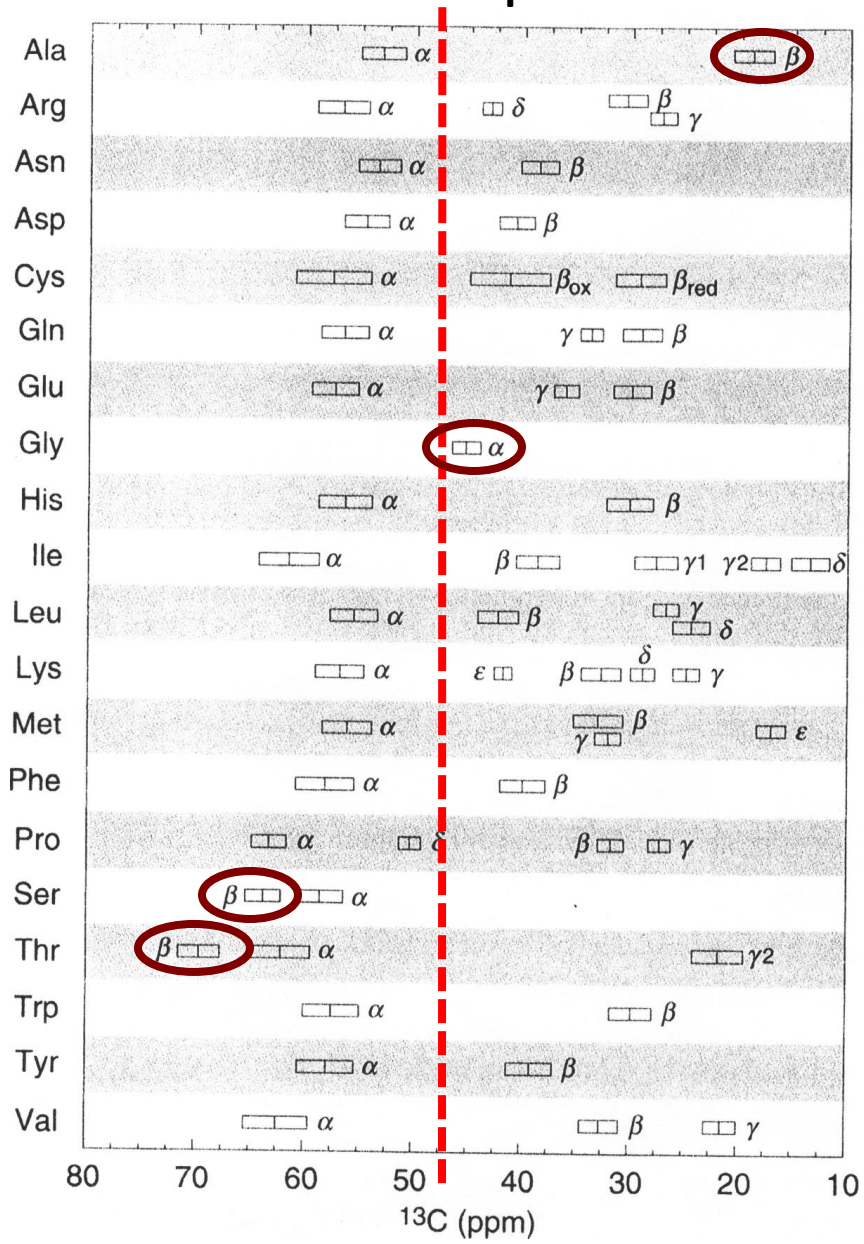
L / D / N / C / F / Y

V / P

I

Q / R / K / M / H / W / E

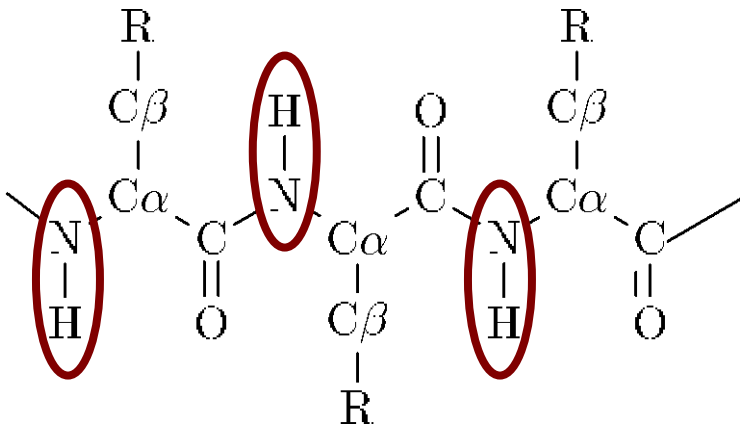
^{13}C chem. shifts in proteins

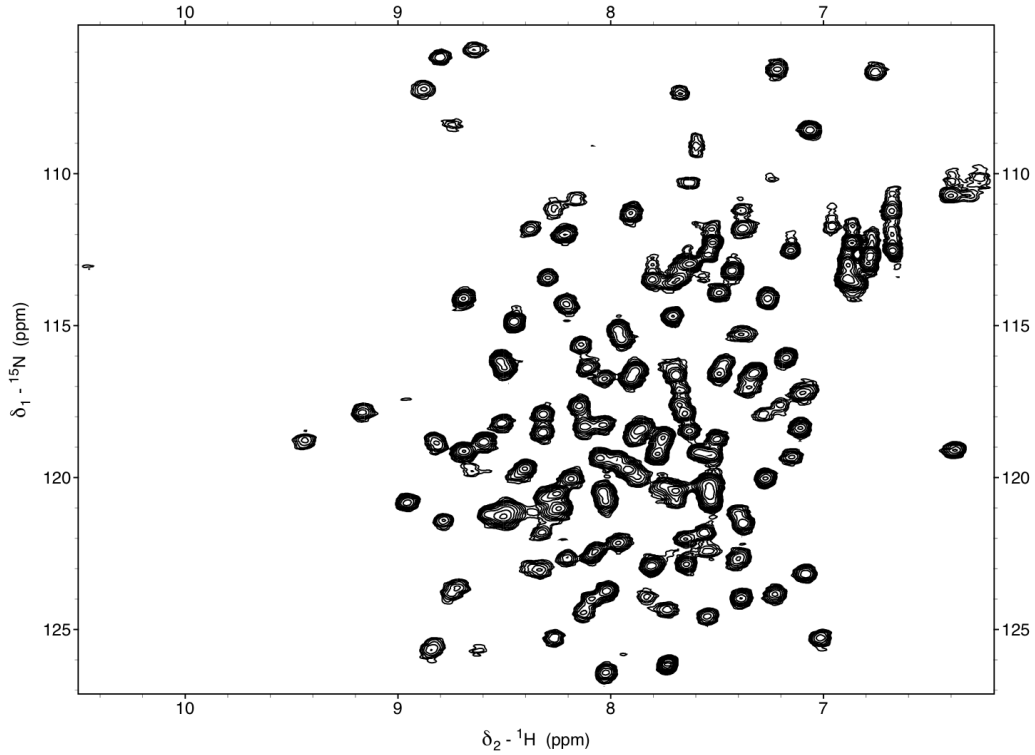


Cavanagh et al. Protein NMR Spectroscopy 2nd ed.

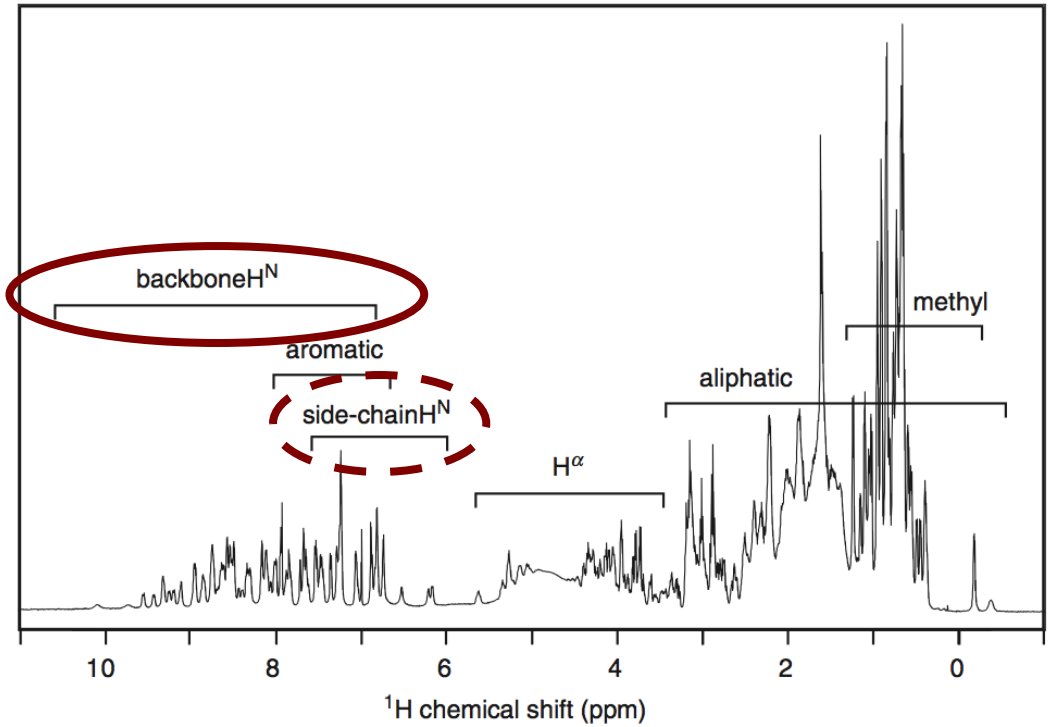
^{15}N-^{1}H HSQC

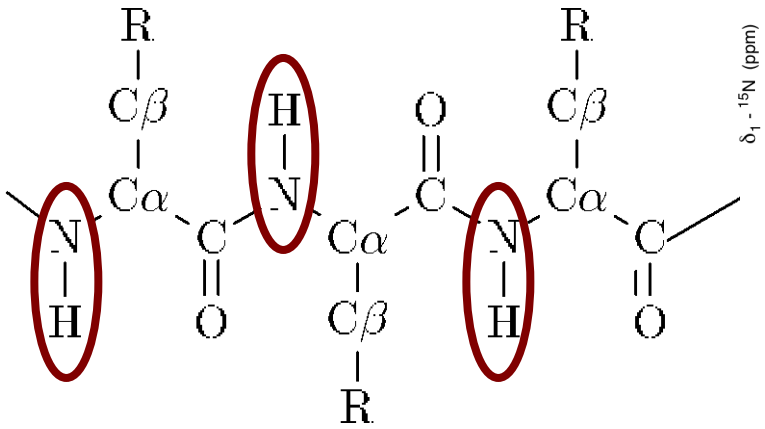
1) 1 peak ≅ 1 aa
2) Excellent info about protein folding state
3) No sequential info
4) For sequential assignment (to know which peak is which aa), 3rd dimension needed

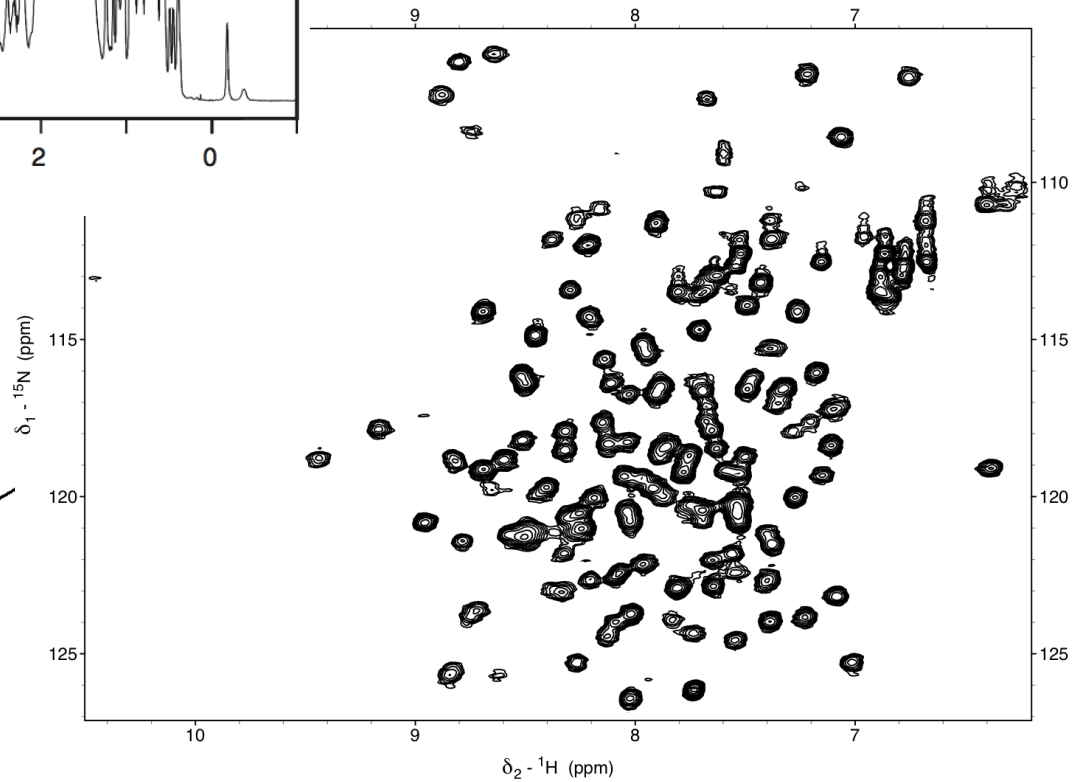




^{1}H 1D, Cavanagh et al., 2007



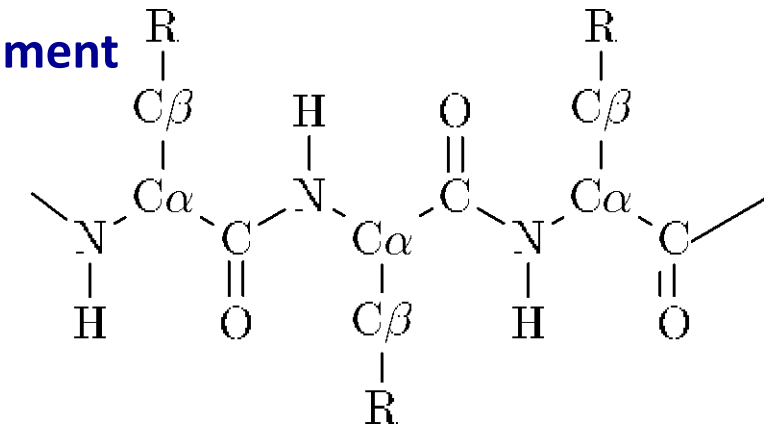




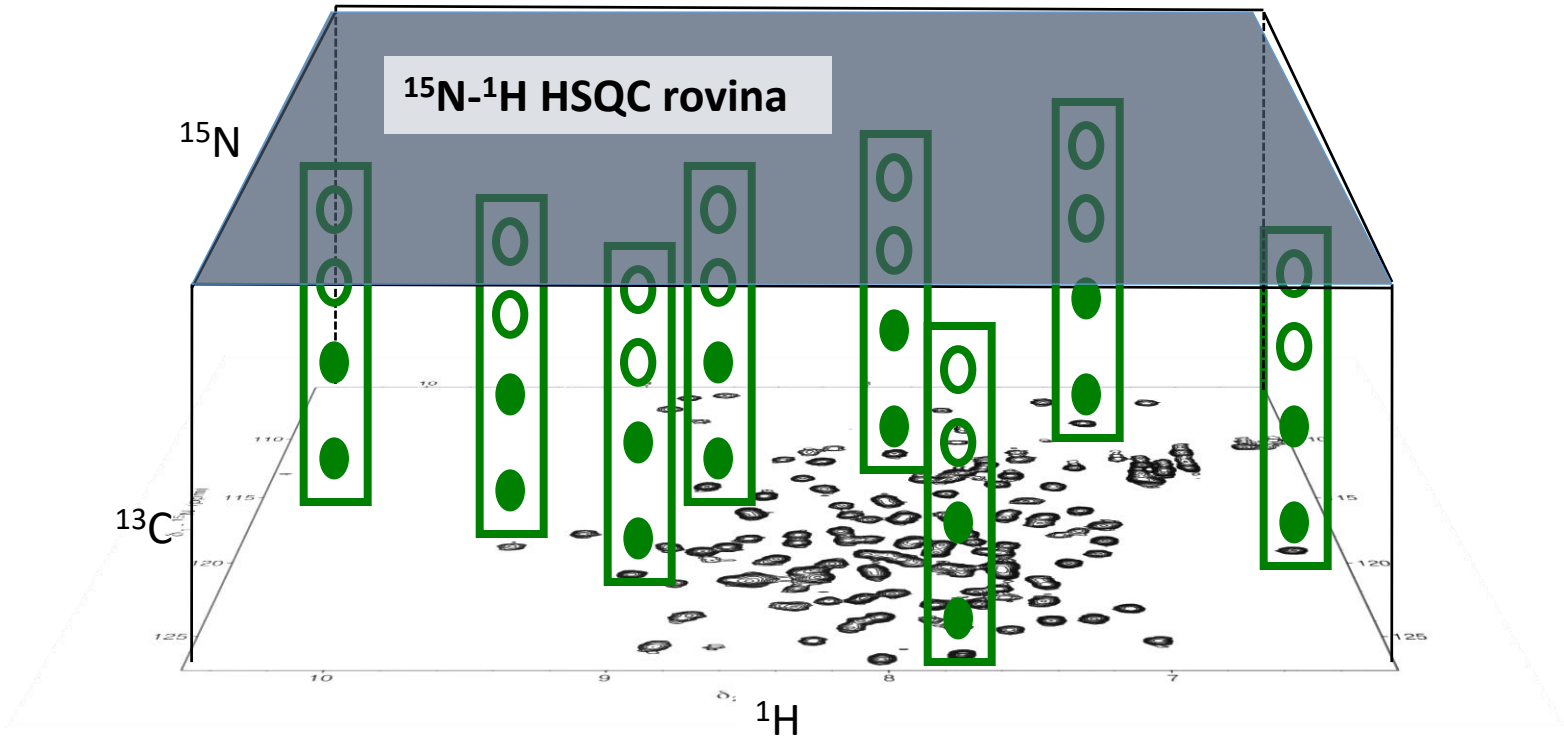
^{1}H-^{15}N HSQC, cca 155 aa, well folded, 600MHz, 293K

So far so good

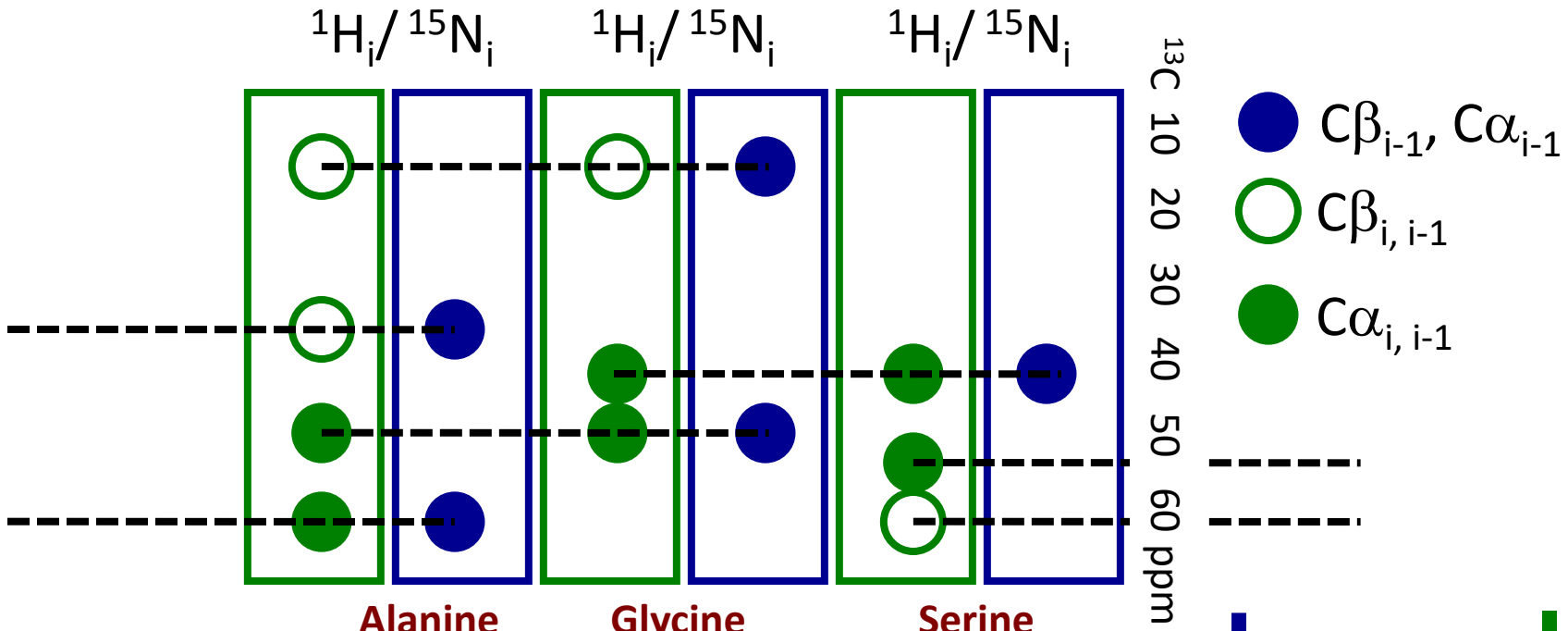
Std. set of 3D NMR spectra for sequential assignment

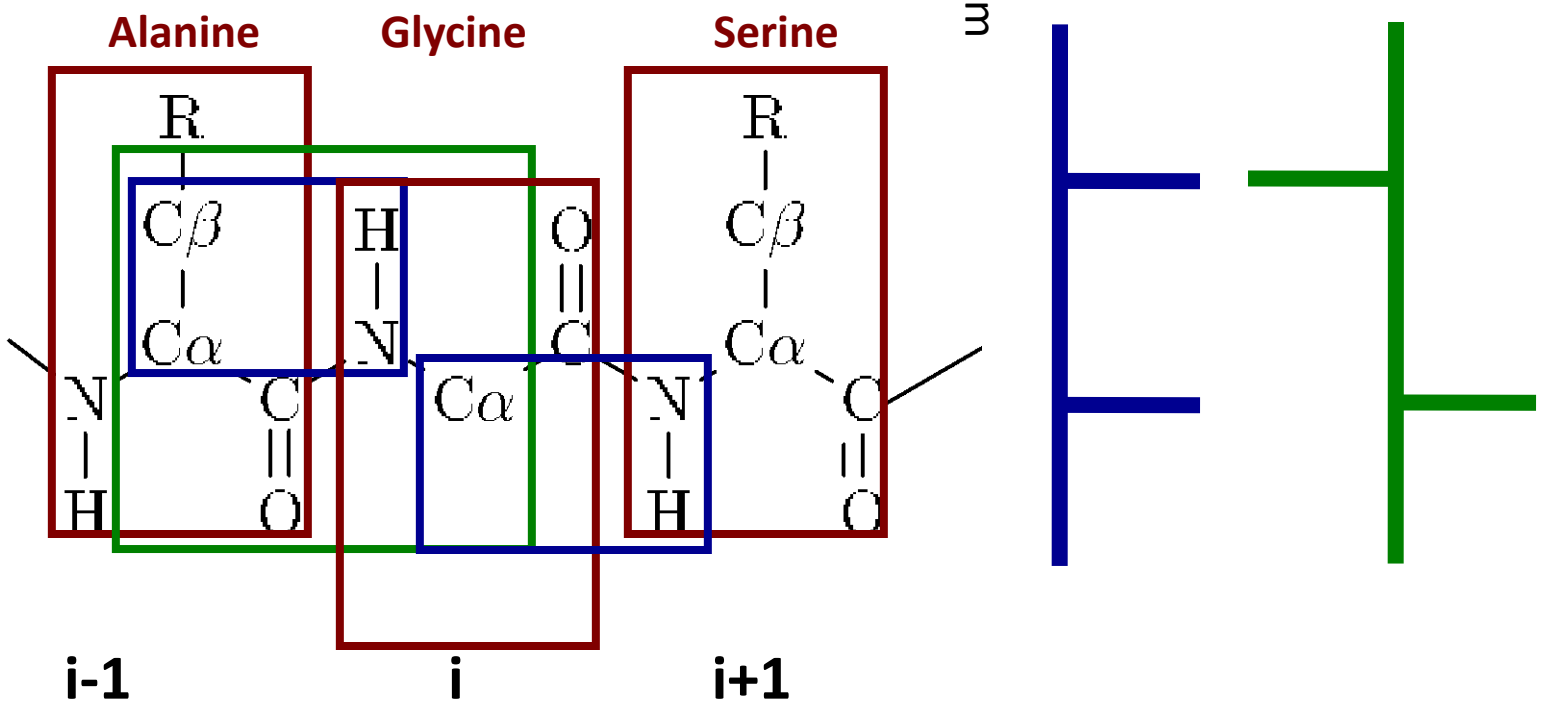


1) HNCO – C=O
2) HNCA – $C_{\alpha,i}$, $C_{\alpha,i-1}$
3) HNCOCA – $C_{\alpha,i-1}$
4) **HNCACB** – $C_{\alpha,i}$, $C_{\alpha,i-1}$, $C_{\beta,i}$, $C_{\beta,i-1}$
5) **HNCOCACB** – $C_{\alpha,i-1}$, $C_{\beta,i-1}$



Vždy nutná minimálně dvojice spekter


$^{1}H_i/^{15}N_i$
$^{1}H_i/^{15}N_i$
$^{1}H_i/^{15}N_i$
^{13}C 10 20 30 40 50 60 ppm
$C\beta_{i-1}$, $C\alpha_{i-1}$
$C\beta_{i,\ i-1}$
$C\alpha_{i,\ i-1}$


Alanine
Glycine
Serine
R
Cβ
Cα
N
H
C
O
i-1
i
i+1

Problem #1

The protein sequence doesn’t match the peaks

⇒Expected  (S / T) residues - **6**

⇒Observed in spectra - **22**

After desperation of own incompetence and misinterpreting and misprocessing of the spectra, I dared to ask my colleague to show me the four sequences used in expression

$\Rightarrow$ one of the four proteins had 22 S /T

$\Rightarrow$ colleague **messed up** the sequences and worked with different one she thought and provided

$\Rightarrow$ luckyly enough S / T are very specific and I could recognize them

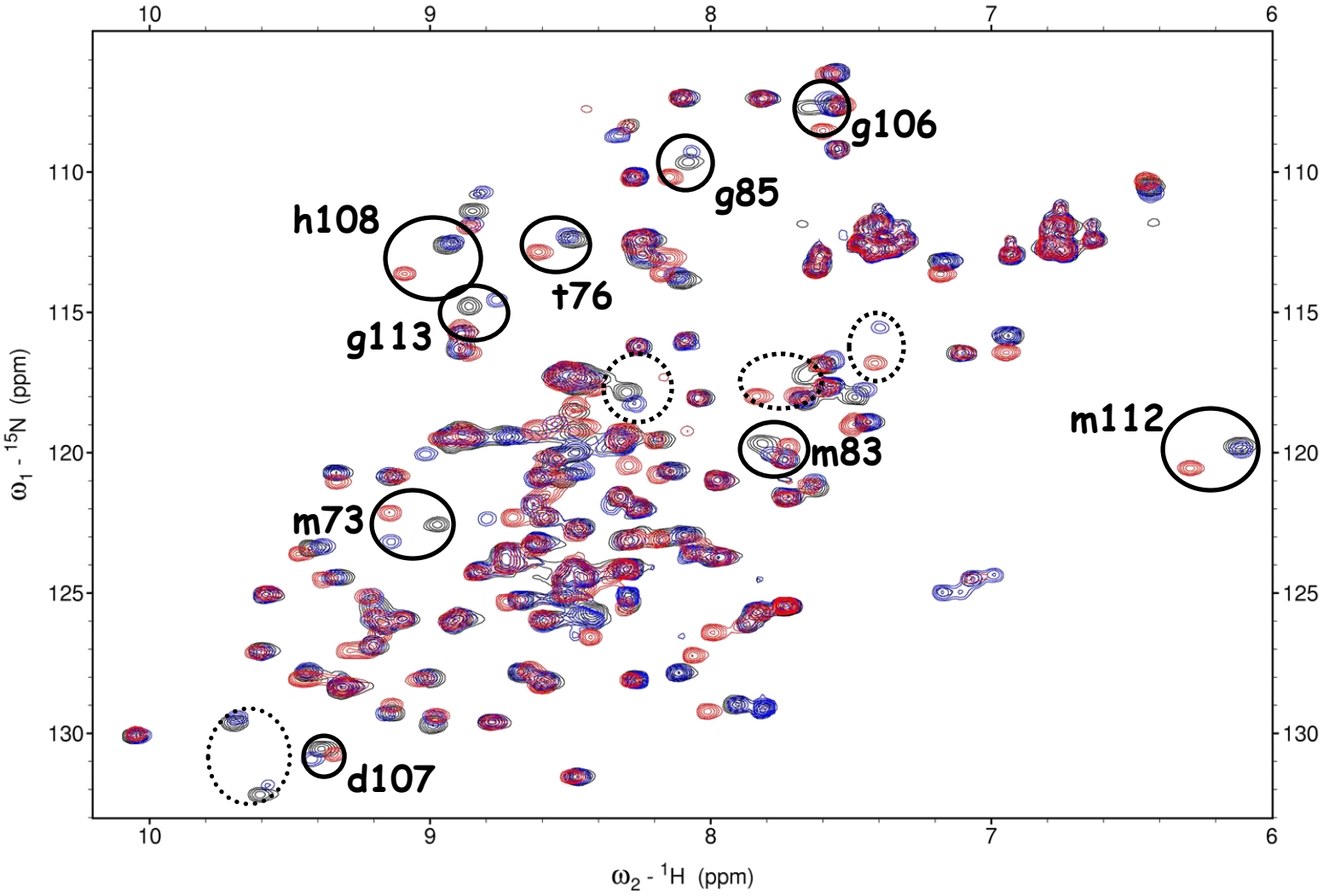
Problem #2

Supervisor didn't support the project anymore and didn’t trust the protein can bind copper as EPR didn’t show any signal

⇒ EPR had receiver contaminated with copper ☺ or ☹

⇒ Despite obtaining ”red-light” from supervisor, I moved on

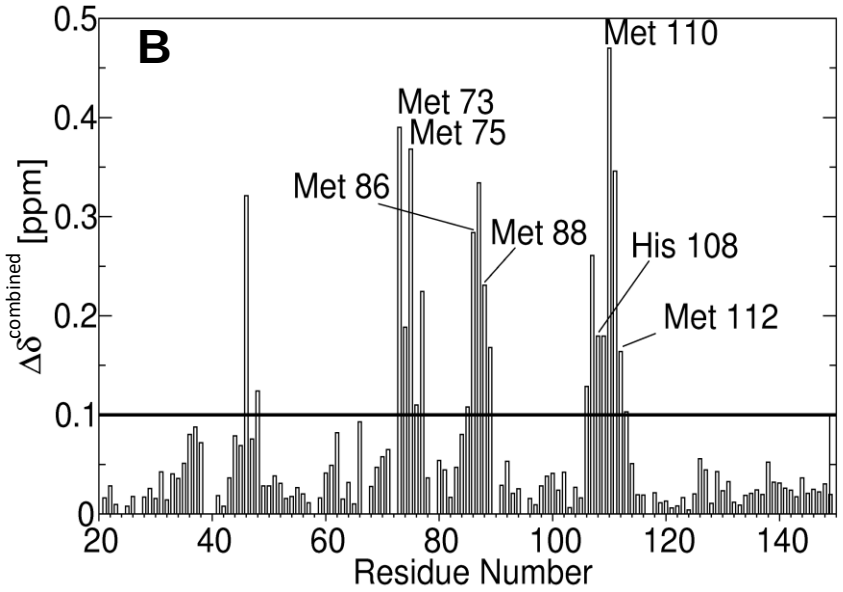
Titration by Cu(II) and Cu(I)

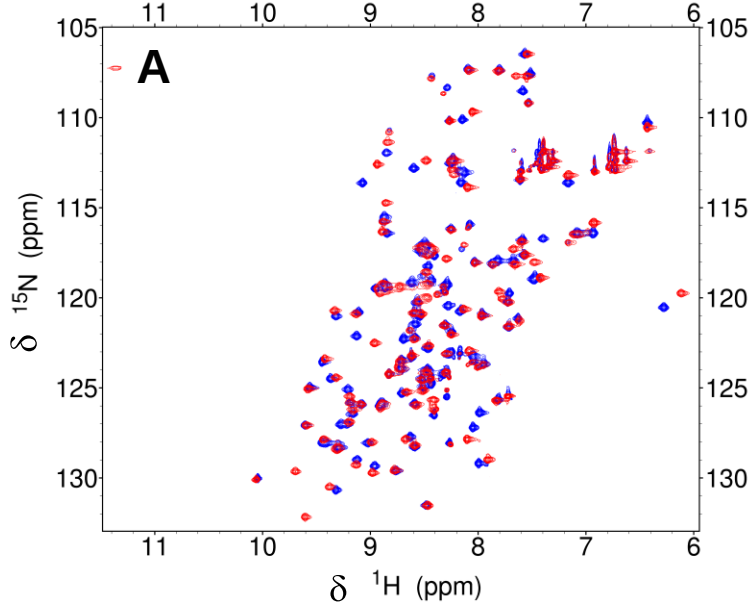


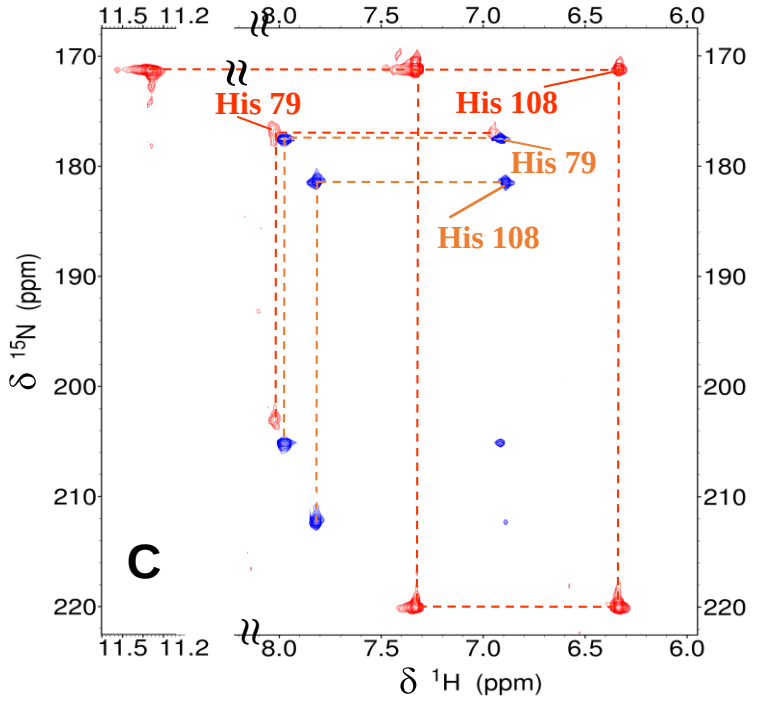
Apo, **Cu(II)** and **Cu(I)**

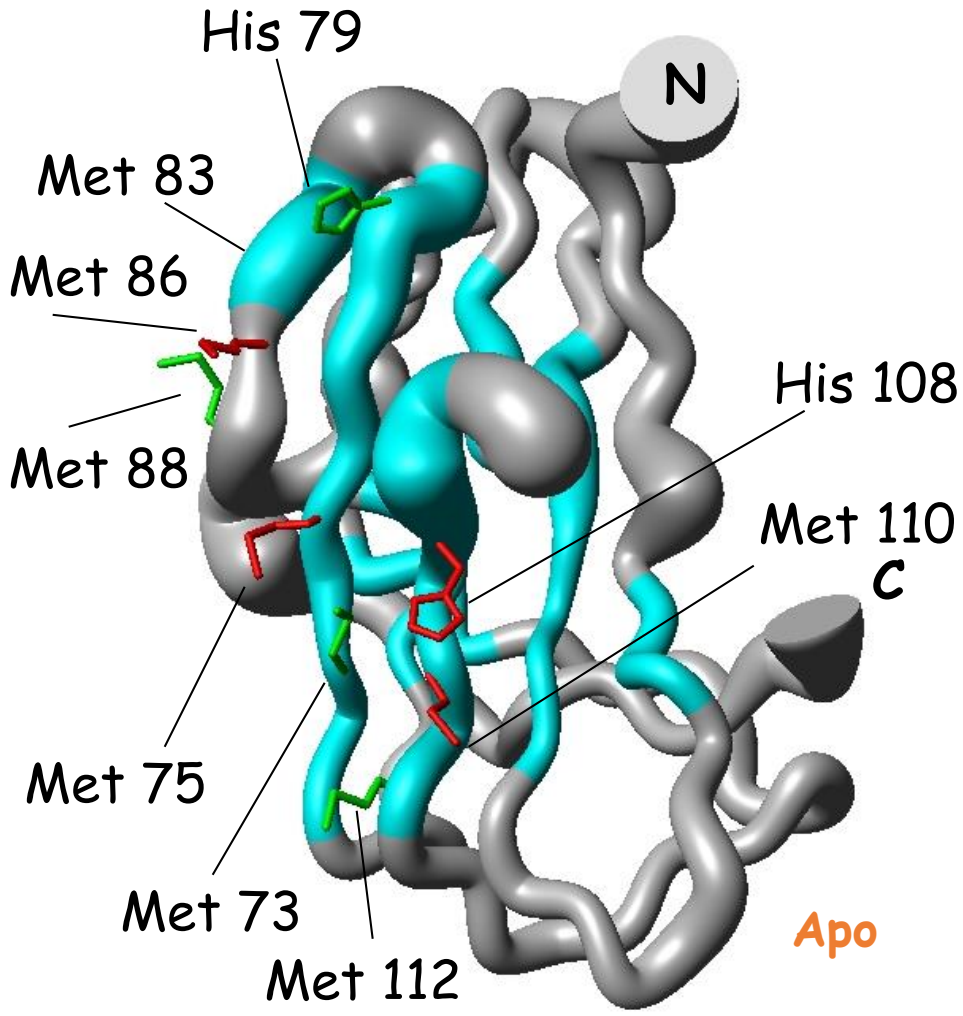
Interaction of DR1885 with copper

-titration (A,B)
-^{2}J HSQC (C)







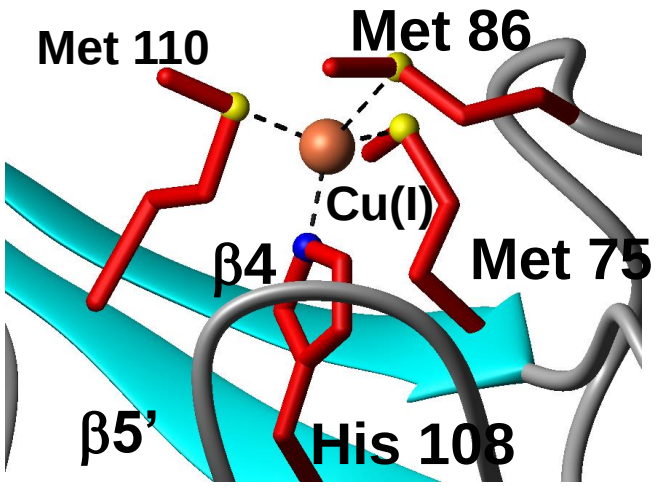

His 79
N
Met 83
Met 86
His 108
Met 88
Met 110
C
Met 75
Met 73
Met 112
Apo
Conserved **Met&His**
Other **Met&His** residues

“Problem” #3

Alphabetical order of authors: **Banci** L, Bertini I, Ciofi-Baffoni S, Katsari E, Katsaros N, **Kubicek** K

i) Being last on your paper is not bad
ii) In case the story is complete and makes sence
iii) Our wasn’t ☹
iv) EXAFS measurement could bring precious info
v) EXAFS expert is prof. Mangani ☹

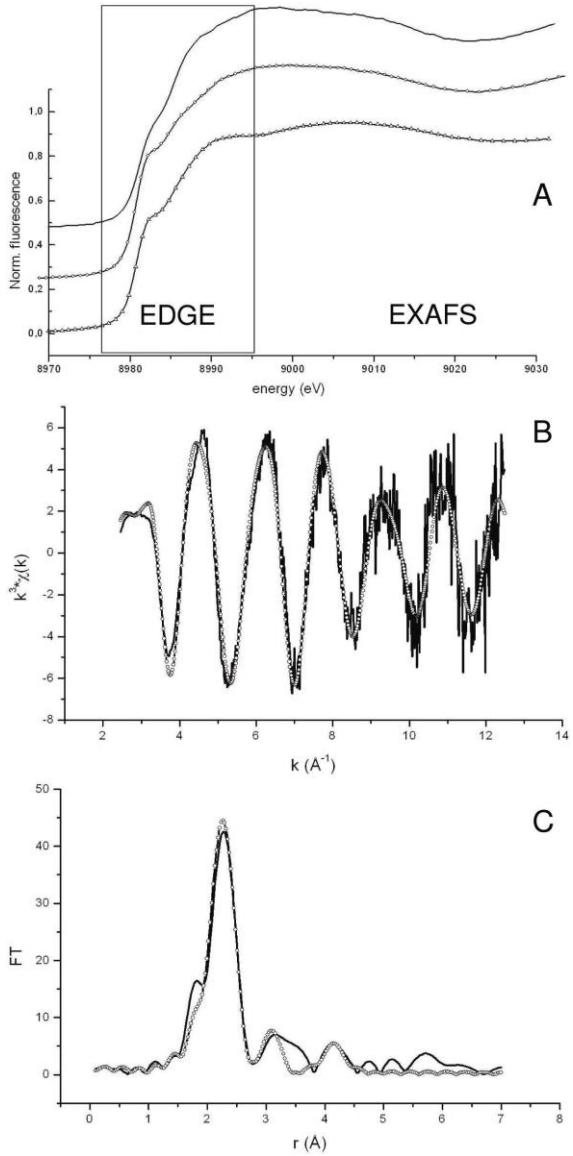
X-Ray Absorption Spectroscopy

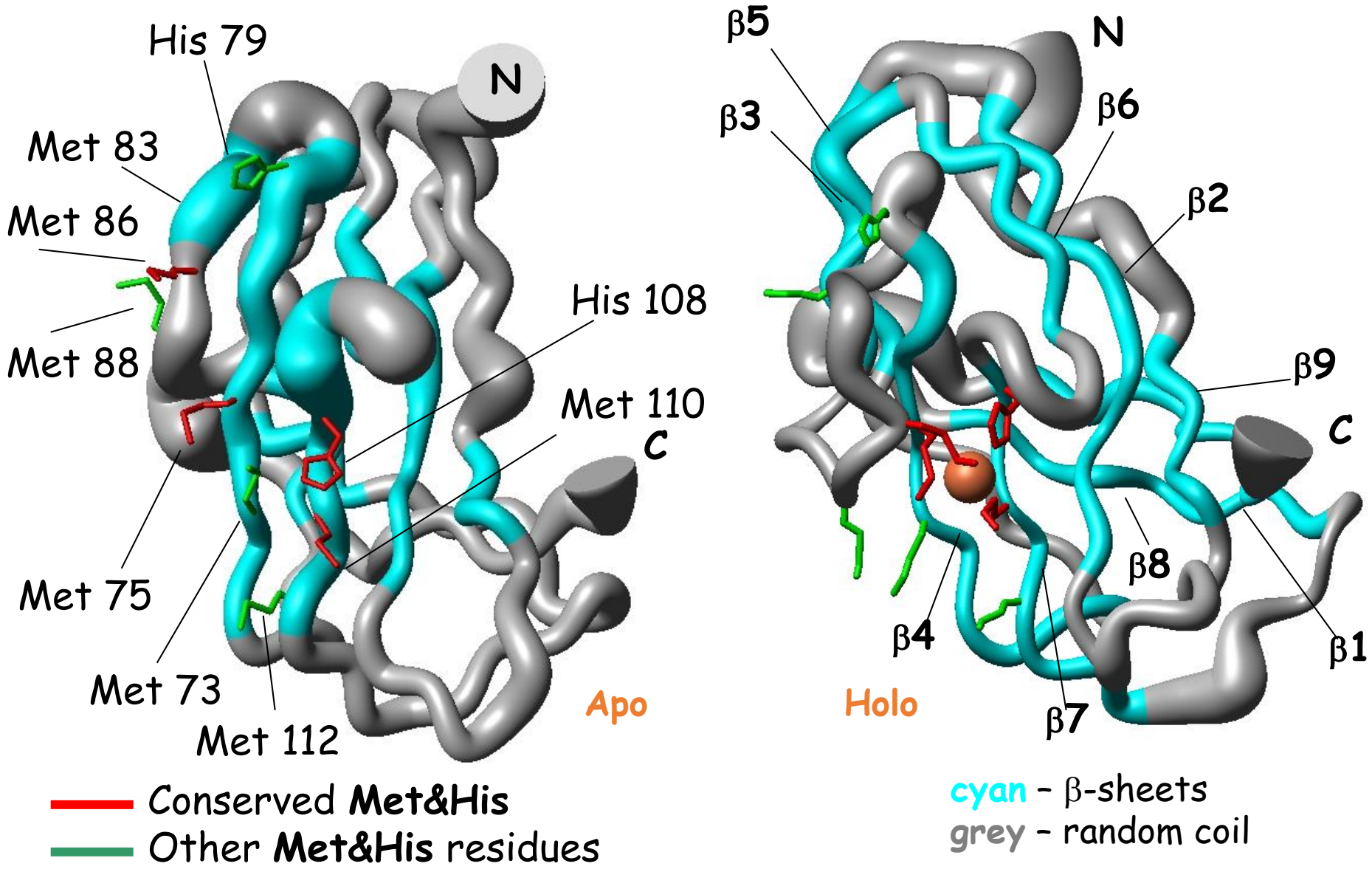


Cu(I)DR1885 ΔE=-10.3 eV

	Ligand	r(Å)	$2\sigma^2.10^3(Å^2)$	R-exafs	ε(fit index)
Fit1 (1shell)	2S	2.299	4(1)	0.446	0.49
Fit2 (1shell)	3S	2.301	9(1)	0.403	0.41
Fit3 (2shells)	3S	2.300	8(1)	0.334	0.29
	1N§	1.982	4(1)		
Fit4 (2shells)	3S	2.303	8(1)	0.305	0.27
	1N*	1.999	7(2)		

§ no MS *His, MS




His 79
N
Met 83
Met 86
His 108
Met 88
Met 110
C
Met 75
Met 73
Met 112
Apo
Conserved Met&His
Other Met&His residues
β5
N
β3
β6
β2
β9
C
β8
β4
β1
β7
Holo
cyan – β-sheets
grey – random coil

A copper(I) protein possibly involved in the assembly of Cu_A center of bacterial cytochrome *c* oxidase

Lucia Banci*, Ivano Bertini*†, Simone Ciofi-Baffoni*, Efthalia Katsari*, Nikolaos Katsaros‡, Karel Kubicek*, and Stefano Mangani*§

*Magnetic Resonance Center and Department of Chemistry, University of Florence, Via Luigi Sacconi 6, 50019, Sesto Fiorentino, Florence, Italy; ‡Institute of Physical Chemistry, National Centre for Scientific Research Demokritos, GR-15310 Agia Paraskevi Attikis, Greece; and §Department of Chemistry, University of Siena, Via Aldo Moro, 53100, Siena, Italy

Edited by Gregory A. Petsko, Brandeis University, Waltham, MA, and approved January 25, 2005 (received for review August 20, 2004)

Sco1 and Cox17 are accessory proteins required for the correct assembly of eukaryotic cytochrome *c* oxidase. At variance with Sco1, Cox17 orthologs are found only in eukaryotes. We browsed bacterial genomes to search proteins functionally equivalent to Cox17, and we identified a class of proteins of unknown function displaying a conserved gene neighborhood to bacterial Sco1 genes, all sharing a potential metal binding motif $H(M)X_{10}MX_{21}HXM$. Two members of this group, DR1885 from *Deinococcus radiodurans* and CC3502 from *Caulobacter crescentus*, were expressed, and their interaction with copper was investigated. The solution structure and extended x-ray absorption fine structure data on the former protein reveal that the protein binds copper(I) through a histidine and three Mets in a cupredoxin-like fold. The surface location of the copper-binding site as well as the type of coordination are well poised for metal transfer chemistry, suggesting that DR1885 might transfer copper, taking the role of Cox17 in bacteria. On the basis of our results, a possible pathway for copper delivery to the Cu_A center in bacteria is proposed.

structure (EXAFS) and NMR data, indicates that DR1885 is a copper protein, possibly involved in the assembly of CcO. In particular, we propose that it can take the role of the mitochondrial Cu(I) chaperone Cox17 in the extracytoplasmic environment of bacteria.

Materials and Methods

Sequence Analysis. The STRING program (Search Tool for the Retrieval of Interacting Genes/Proteins, www.bork.embl-heidelberg.de/STRING) was used to identify the bacterial Sco1 neighboring genes. The BLAST program was used to search over all nonredundant GenBank database genomes for the DR1885 homolog sequences. Sequence alignments were performed with CLUSTALW (11). Prediction of transmembrane helices and membrane topology of all sequences was obtained by using the HMMTOP and TMPRED programs (12, 13).

Protein Cloning and Purification. The genes from *D. radiodurans*,

Note the alphabetical order of authors

Example #2

One issue

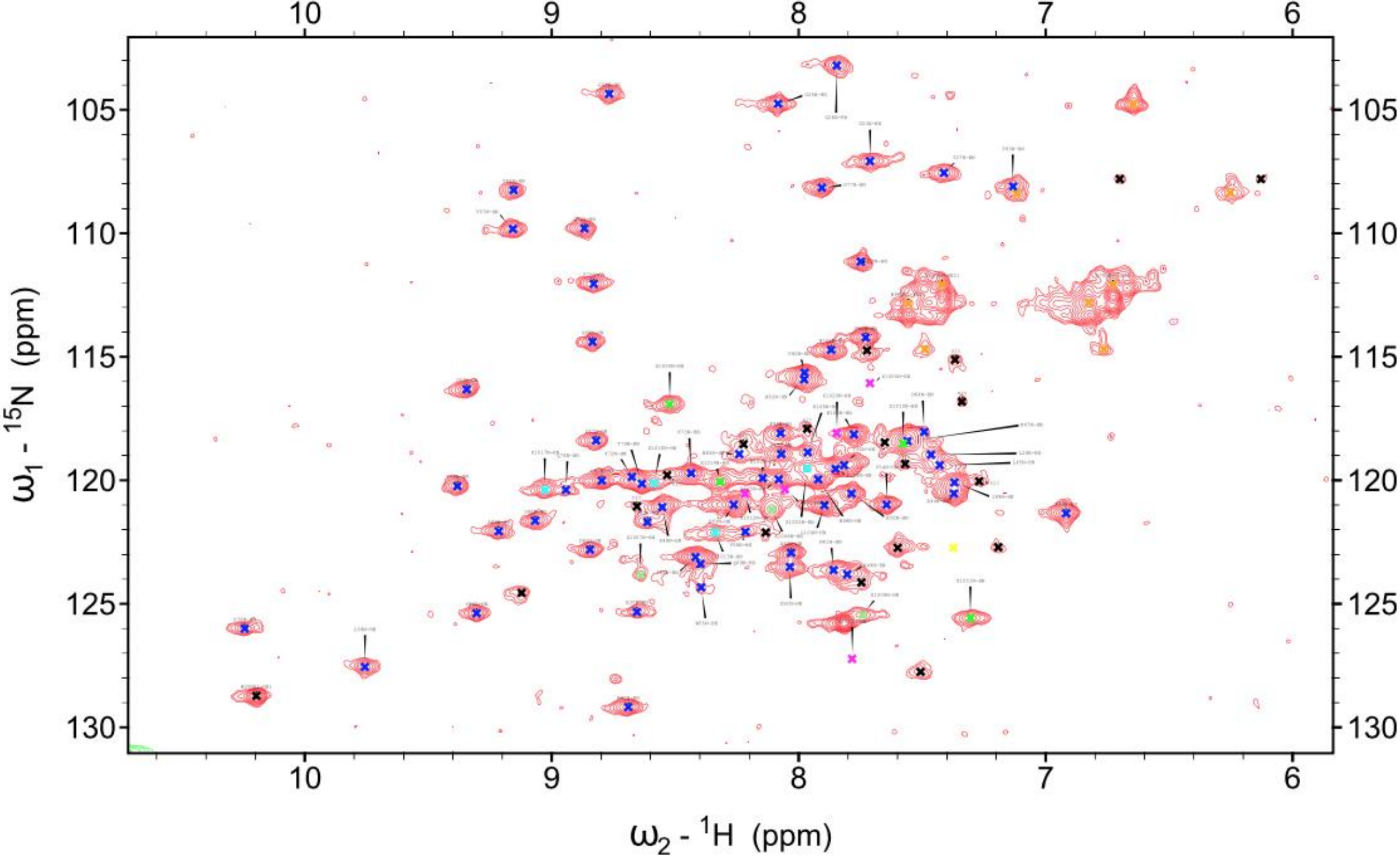
1) Things are not as easy as they seem to be

1) Protein expresses well

2) ^{15}N HSQC looks nice

3) Protein is stable for about 7-14 days

4) Something, however, still doesn’t fit (concentration reachable only to .5mM)


10
9
8
7
6
105
110
115
120
125
130
ω_1 - ^{15}N (ppm)
ω_2 - ^{1}H (ppm)

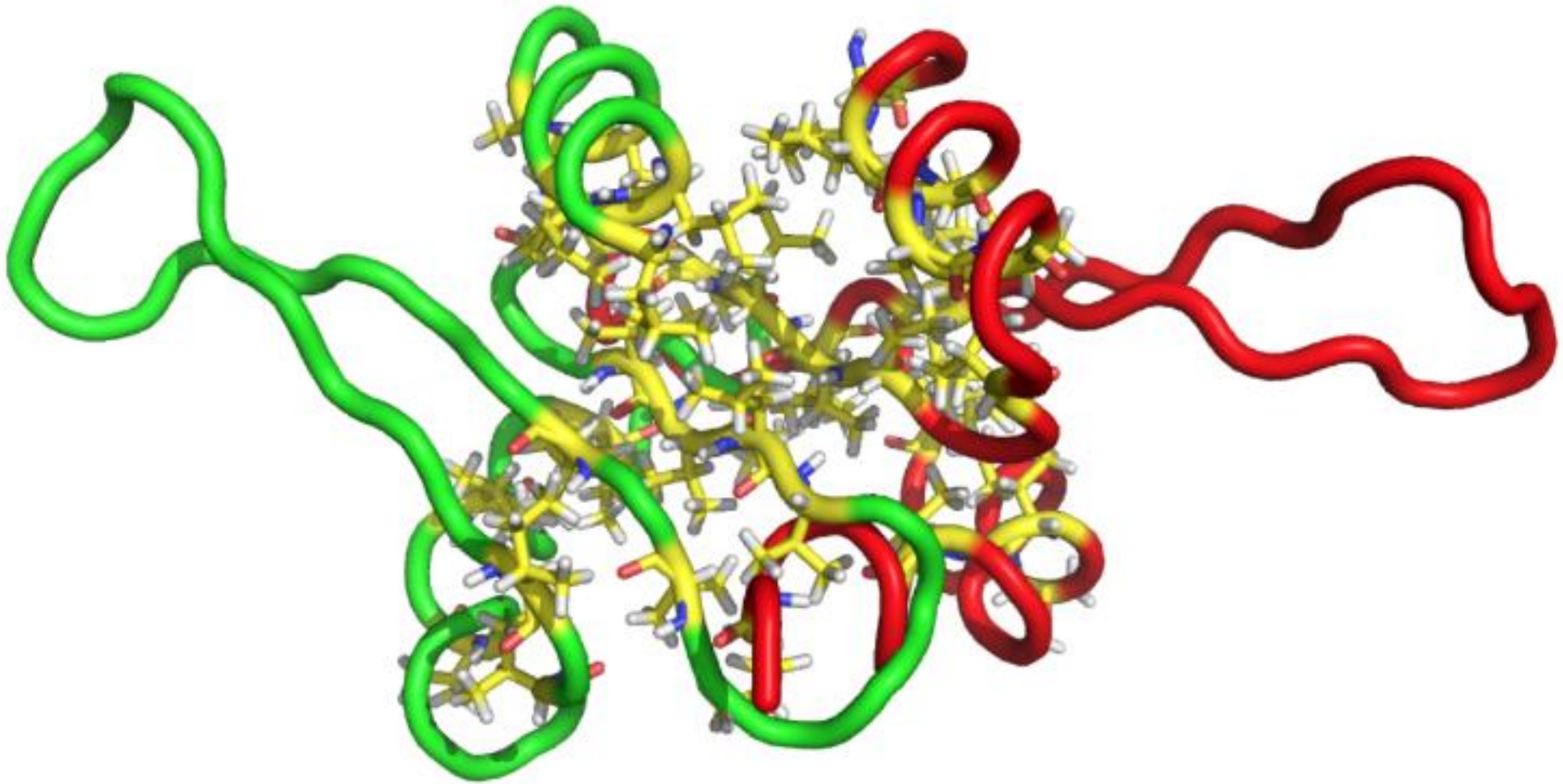
Something doesn't fit

=> During the structure calculation, lot of hydrophobic residues exposed to the solvent ☹

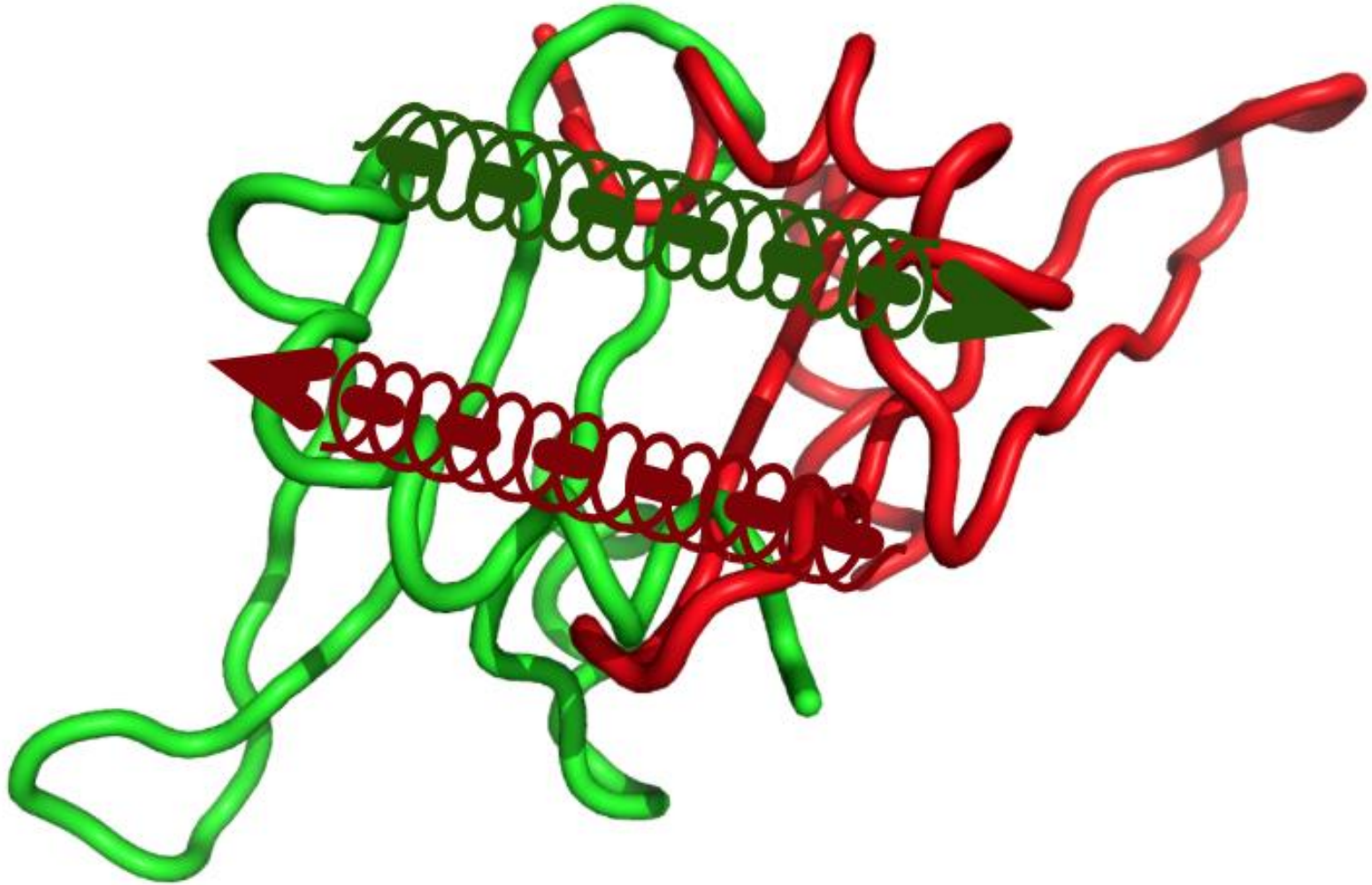
=>SAXS, NMR and AUC point to dimer while we expect monomer

T_1 [s]	T_2 [s]	T_1/T_2	ν_N [MHz]	τ_c [ns]	MW [kDa]	MW monomer [kDa]
1.4163	0.0385	36.7870	70.964	16.394	26.801	$\approx 12.8kDa$

Is this a real structure or is it an artefact of experimental conditions?



As it is domain of RNA binding protein, may be this is the correct arrangement



But fluorescence anisotropy shows no significant binding to RNA ☹

2 - 3 years of work and no plausible result☹

Back to roots!

i) Check the protein sequence once again
ii) Read literature
iii) Push!
iv) And find some diligent student(s) and co-workers to help with the job(s)

Comparison of **original** (don't call it old, it's not kind) and **extended** constructs

1) Protein expresses well
    - Expressions even better
2) ^{15}N HSQC looks nice
    - Spectra even nicer
3) Protein is stable for about 7-14 days
    - Protein holds for even longer
4) Something, however, still doesn't fit (concentration reachable only to .5mM)
    - Is that enough?

Time to move to 21st century

1) From traditional 3D NMR spectra for assignment

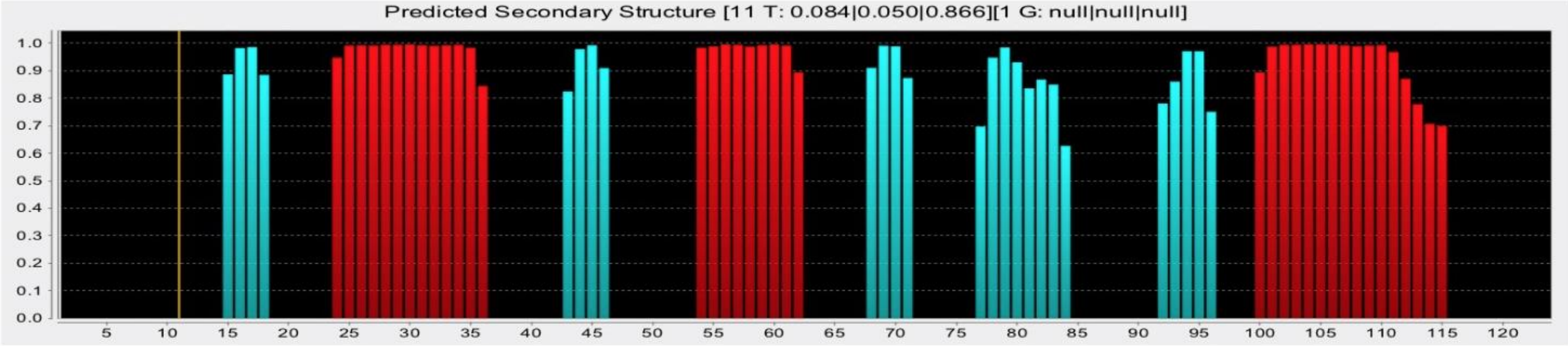
2) New approaches => new challenges!

Expt. Name	Dimensionality	Correlated Nuclei
^{15}N-HSQC	2D	H^N-N
^{13}C-HSQC	2D	H-C
HNCACB	3D	H^N-N-(Cα_i-Cβ_i,-Cα_{i-1}-Cβ_{i-1})
CCCONH	3D	H^N-N-C$_i$
HNCO	3D	H^N-N-CO$_{i-1}$
HNCACO	3D	H^N-N-(CO$_{i-1}$,CO$_i$)
HBHACONH	3D	H^N-N-Hα_{i-1}-Hβ_{i-1}
HCCCONH	4D	H^N-N-C$_i$-H$_i$
HCCH-TOCSY	4D	C-H-C$_i$-H$_i$
HNCH-NOESY	4D	H^N-N-C$_{ij}$-H$_{ij}$
HCCH-NOESY	4D	C-H-C$_{ij}$-H$_{ij}$

Backbone and side-chain assignment achieved in 3+3 weeks:

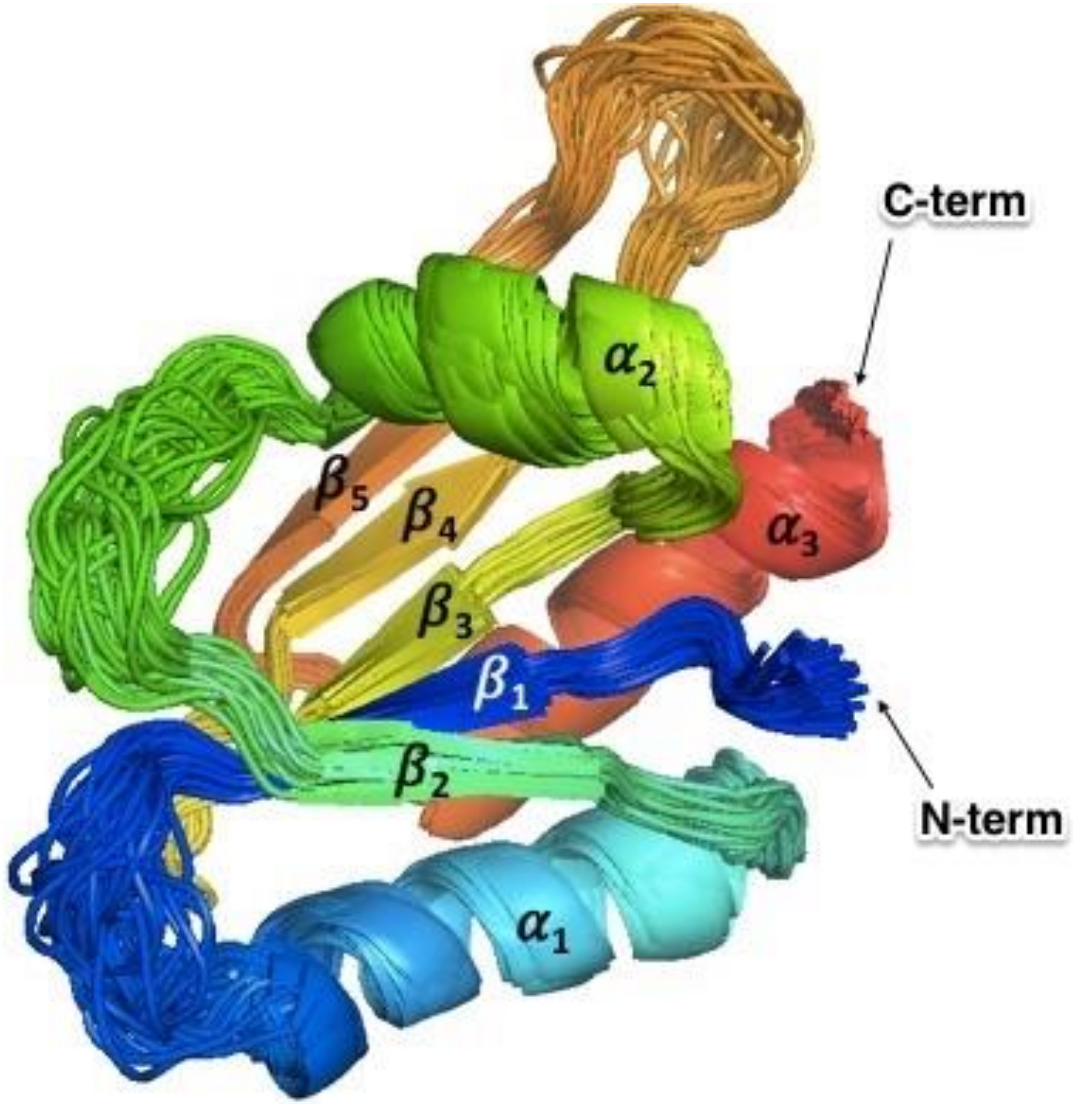
acquisition+processing, respectively

Secondary structure estimation from expt. data looks fantastic



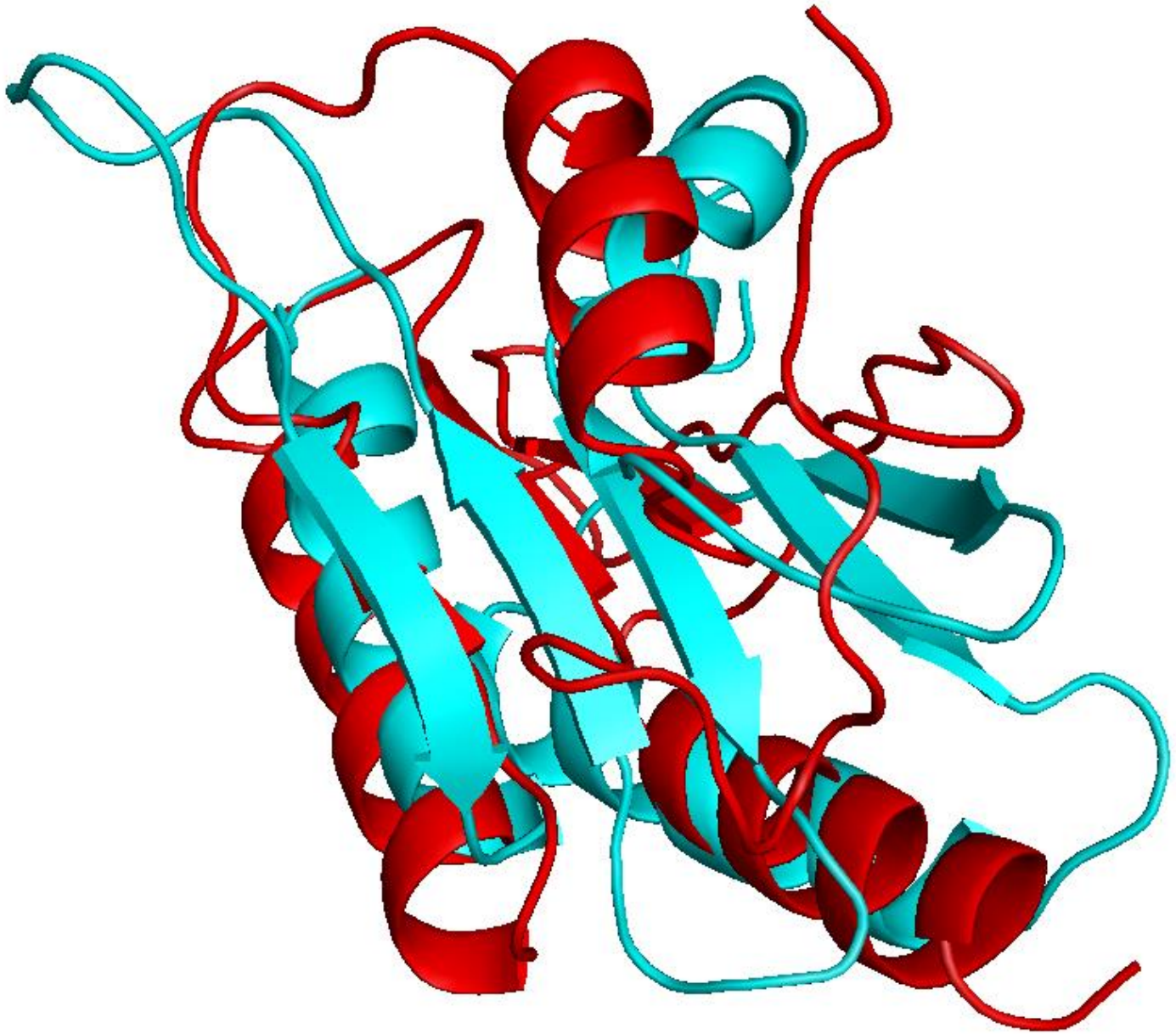
Structure calculated and refined within autumn semester of bachelor study

CYANA structure (50 structures)
(topology -$\beta_1\alpha_1\beta_2\alpha_2\beta_3\beta_4\beta_5\alpha_3$)

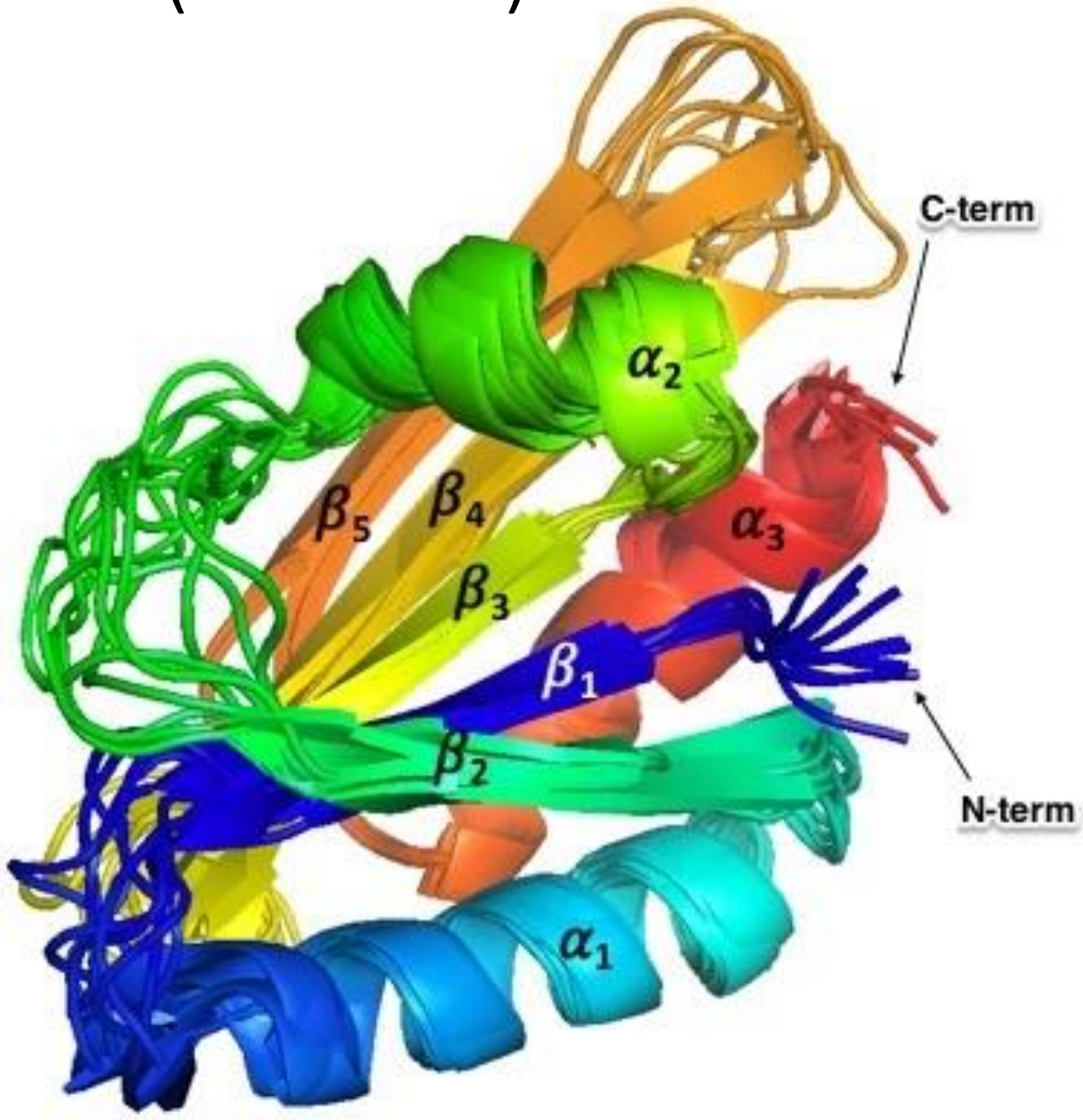


But! Does it make sense? Where would the RNA bind?

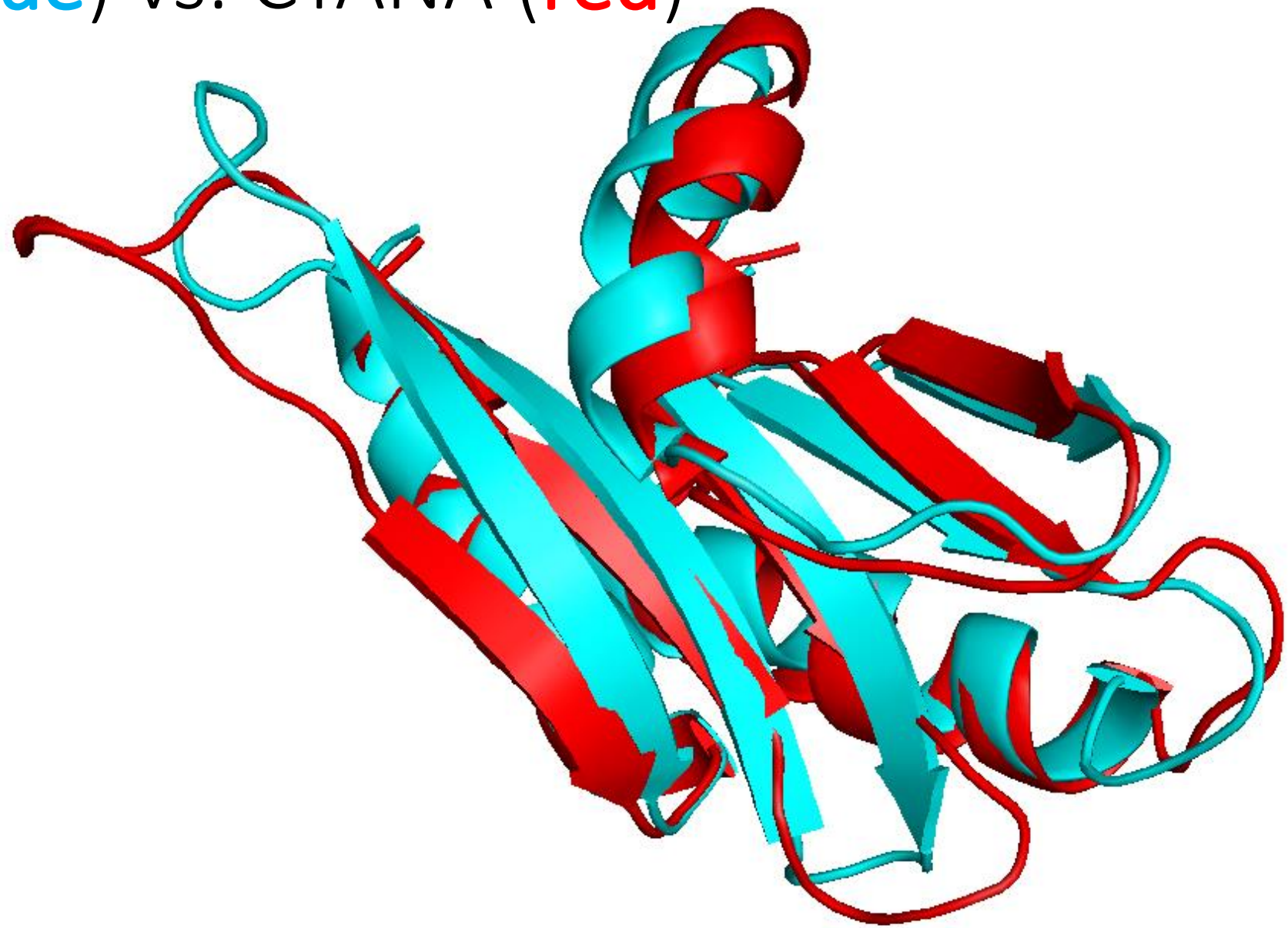
DALI server comparison c-term (blue) vs. 2g4c (red)



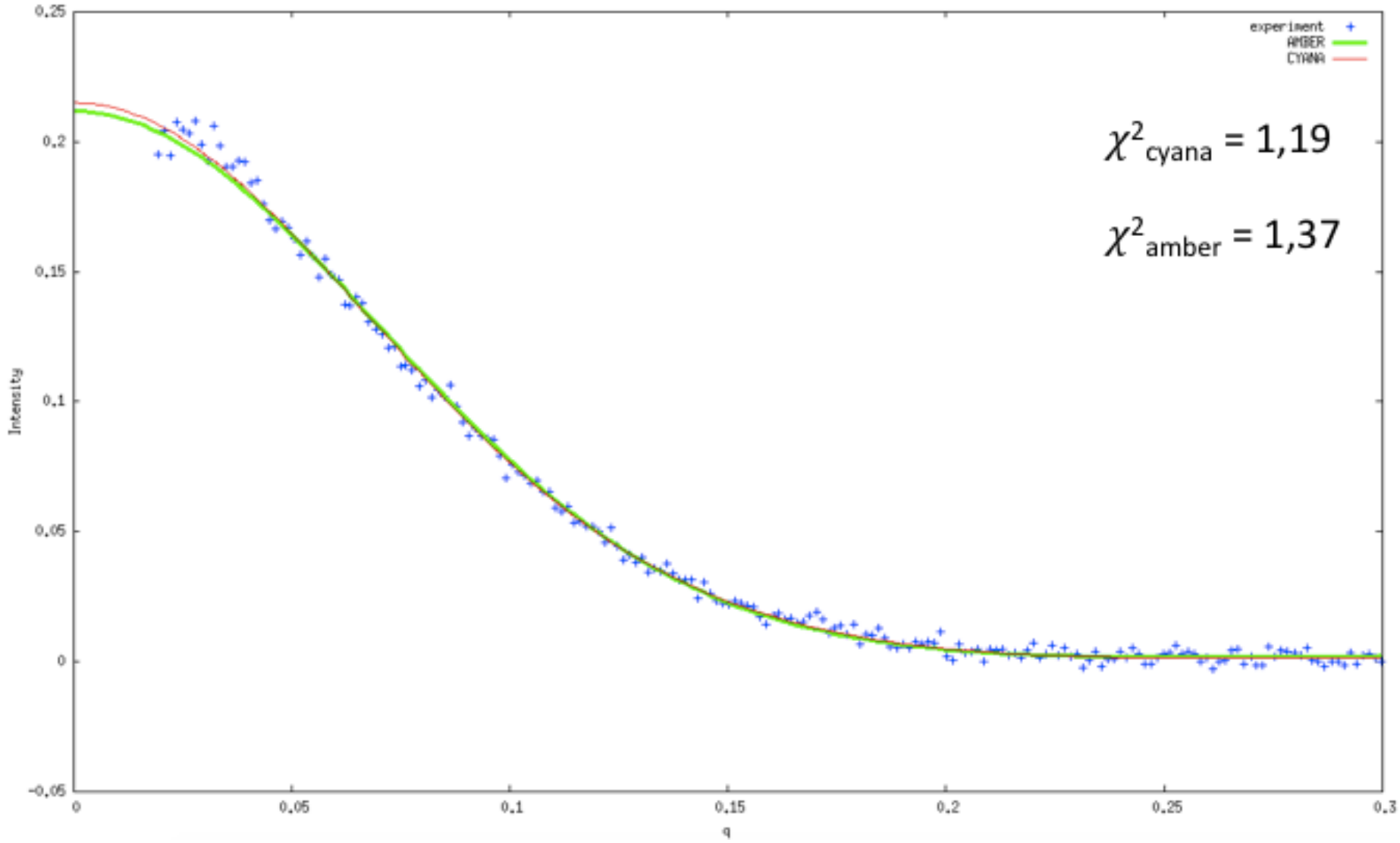
AMBER structure (best 10)



AMBER (blue) vs. CYANA (red)



SAXS comparison of expt. vs. calculated data



This story goes on but seems promising! Veeeery promising!

Take home message

Q: **Why didn’t you succeed?**

A: **‘cause we didn’t try hard enough!** Kvido Stříšovský

From the “Shawshank” movie

Q: Why R U here?

A: ’cause the lawyer screwed it up!

Acknowledgment

Sklenář Lab
Hana Křížová
Erik Caha
Michaela Wimmerová

Berini Lab
Cristina Del Bianco
Simone Ciofi-Baffoni
Isabella Felli

Griesinger/Carlomagno Lab
Marcel Reese
Victor Sanchez-Pedregal

Acknowledgment

The Stefl Lab:
Karel Kubicek
Olga Jasnovidova
Josef Pasulka
Roberto Pergoli
Maria Sanudo
Hana Cerna
Fruzsina Hóbor
Veronika Bacikova

Stepanka Vanacova:
Dominika Hrossova
Peter Holub

EMBO & HHMI

HFSP Human Frontier Science Program

Goethe Universität Frankfurt am Main:
Frank Löhr

The Fajkus lab:
Michal Zimmermann
Ctirad Hofr

The Sklenar lab

MŠMT

SEVENTH FRAMEWORK PROGRAMME

GAČR Grantová agentura České republiky
Czech Science Foundation

GAV The Grant Agency of AS CR