



Biophysical methods to guide protein crystallization

Protease Platform

Paul Erbel

RAMC: September 2013



Wiki:

There is no universally accepted definition of a mountain...

In the United Kingdom and the Irish Republic, a mountain is usually defined as any summit at least 2,000 feet (or 610 metres) high

The UN Environmental Programme's definition of "mountainous environment" includes any of the following:

Elevation of at least 2,500 m (8,200 ft);

Elevation of at least 1,500 m (4,900 ft), with a slope greater than 2 degrees;

Elevation of at least 1,000 m (3,300 ft), with a slope greater than 5 degrees;

Elevation of at least 300 m (980 ft), with a 300 m (980 ft) elevation range within 7 km (4.3 mi).

Cliche:

Mountain as a metaphor ... for a crystallization challenge

Biophysics and crystallization:

when is it useful,

when does it make a difference?

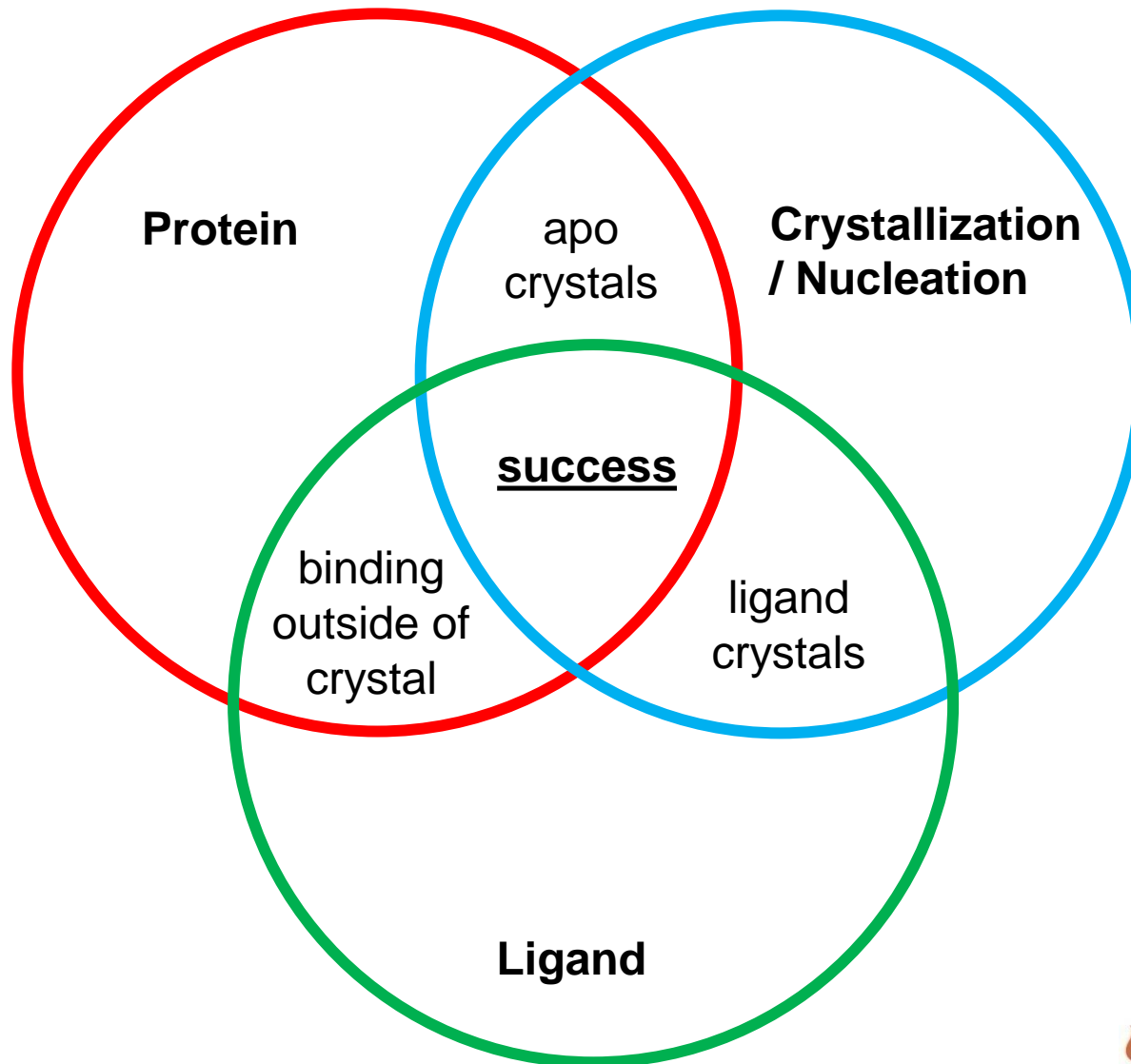
Parameters in protein crystallization 'in industry' *playing field*

Protein

**Crystallization /
Nucleation**

Ligand

Parameters in protein crystallization 'in industry' *playing field*



Parameters in protein crystallization 'in industry'

playing field

Protein

Protein quality

- purity, aggregation state, stability

Structural order of protein

- construct design, floppy termini or loops
- 'functional' disorder, changes upon ligand binding

Crystallization / Nucleation

Find crystallization condition

- Screening philosophy

Crystal contacts

- Natural variants, surface mutations
- Packing prevent compound binding?

Ligand

Characterization of ligand binding

- affinity, stoichiometry, kinetics

Ligand solubility versus IC₅₀

A real example ...

Introduction

- Metallo protease: ~35 kDa
 - No structure or structure of close analogue available
 - Part of protein complex, in isolation biochemical activity is very low.
- Summary of crystallization challenge from 2010-2012

A real example ...

Scouting

- Standard version: Protein production
 - About 10 constructs designed based on 'protease domain knowledge'
 - variations in N, C-termini
 - BEV used as main expression systems
 - Low expression yields, both soluble and insoluble!
 - Basic purification (Ni-chelating, reversed Ni-chelating, size exclusion)
 - protein overall not very stable based on precipitation during purification
 - crystallization trials up to a protein concentration of 10mg/ml

⇒ No sign of crystallization in limited screens

A real example ...

More scouting

- DeLuxe version: Protein production
 - Attempts to increase expression yield
 - Different fusion proteins, tags (his-tag, strep-tag, ZZ-fusion protein,...)
 - Different expression vectors (pTriEx, pET28, pFB)
 - Attempts to reduce dynamics by introducing loop mutations and truncation, disulphide bonds
 - Attempts to introduce crystal contacts by surface mutations

- ⇒ Over time >50 constructs explored
- ⇒ No sign of crystallization even in extensive screens
 - Mr. D'Arcy in collaboration with Hauptmann Woodward Institute (1536 conditions)

A real example ...
intermezzo

Are you convinced this is a rather difficult target?

NO?

A real example ...

Adding a binding partner

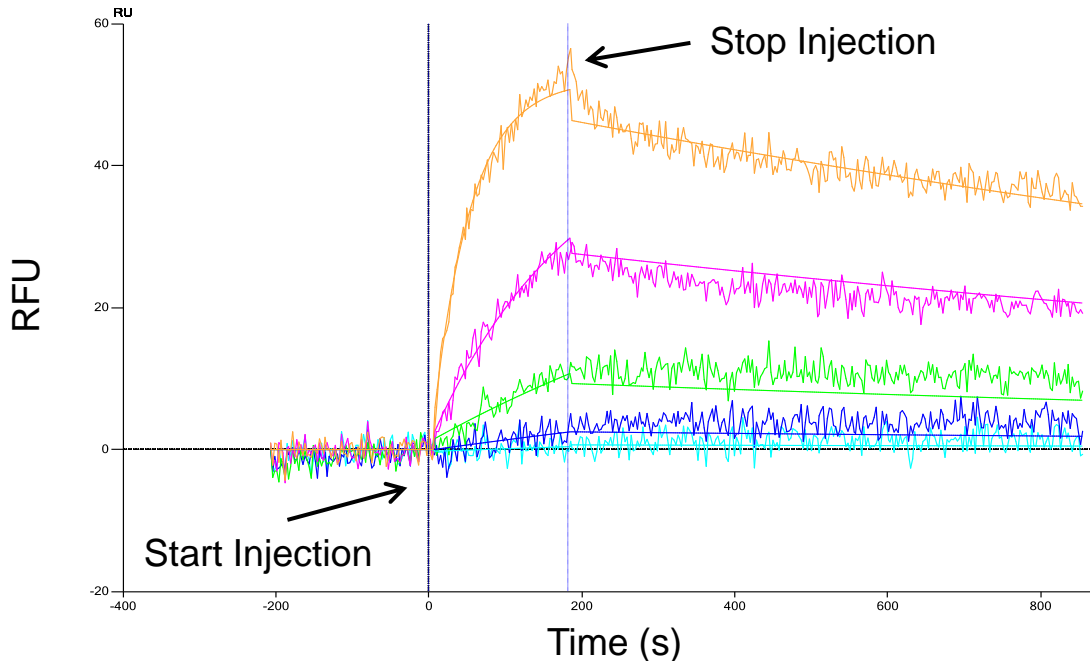
- Phase 1: no crystals obtained after extensive scouting of constructs and crystallization conditions
 - Hypothesis: protease is rather flexible outside protein complex
- Strategy: Look for binding partner
 - Compounds ongoing (weak HTS hit became available during phase 1)
 - Natural binding partner(s)
 - Generate FAB fragments (in-house possibilities with biologics group)

A real example ...

Adding a binding partner: FAB

- Panning and validation
 - 6 Fabs confirmed and expressed
- Characterization of FABS by SPR
 - Binding affinity highest for 1P8
 - Epitope mapping: all FABS (by competition) except 1P8 bind to a similar epitope

Candidate	EC ₅₀ [nM]	k _D [nM]	k _D [nM] c-truncation
	Elisa	SPR	SPR
1024.2J23	1250	Weak	187
1024.1P8	4	4	No binding
1024.1P1	5	25	33
1024.1J23	143	39	46
1024.1M3	30	42	65
1024.1L17	21	10	10



A real example ...

Take home message

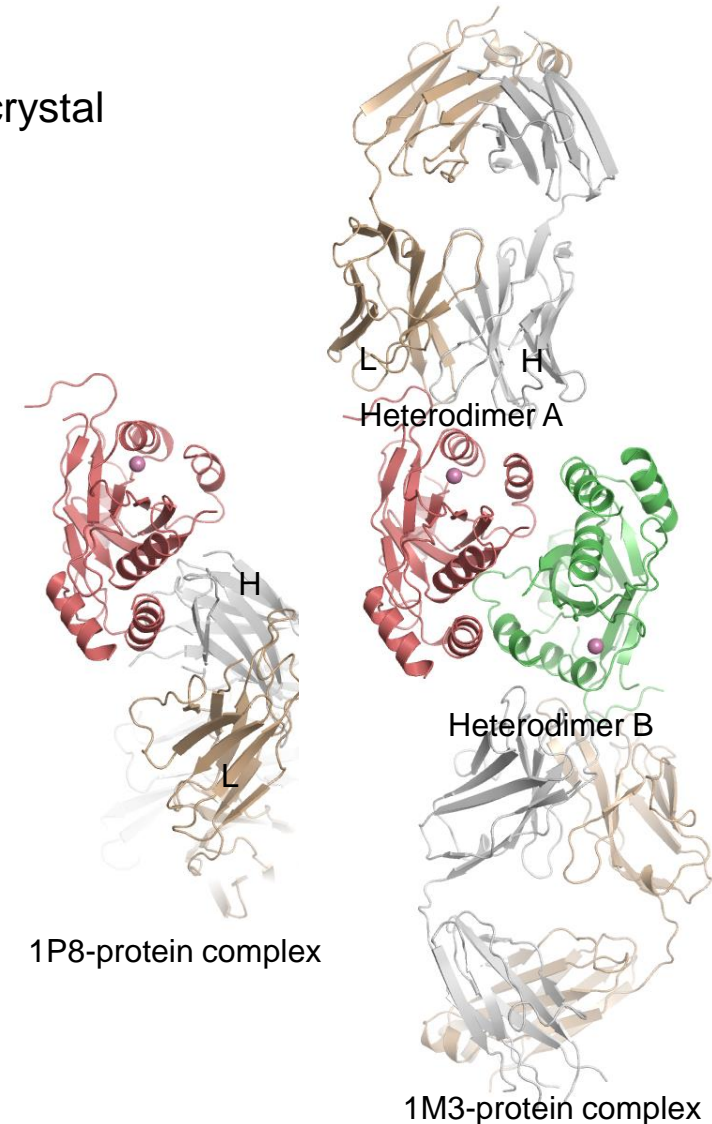
Biophysics (SPR) used to select two out of 6 FABs that recognize different epitopes

Biophysics used to prioritize

A real example ...

Adding a binding partner: FAB – protease complex obtained

- FAB-Protease structures obtained
 - 1M3 binds close to the active site, complex is a dimer in crystal
 - 1P8 binds far from the active site, recognizes C-terminus
 - How does this protease look like in solution?
 - Current assumption: It looks like in complex with FAB as protein is similar in both complexes
 - Why does protease not crystallize without FAB?
 - because it is too flexible?
 - FAB complexes suitable for drug discovery?
 - Soak into existing FAB-Protease complex crystals
 - Attempt to co-crystallize ternary complexes: compound-FAB-Protease
- ⇒ No cocrystal structures with compound obtained!



A real example ...

Take home message

Adding binding partners a good approach to obtain crystals

Now convinced that this is a rather difficult target?

A real example ...

Using biophysics...

- Phase 1 - Scouting: no crystals obtained
 - Hypothesis: protease is rather flexible outside protein complex

- Strategy: Look for binding partner
 - Compounds ongoing (weak HTS hit became available during phase 1)
 - Natural binding partner(s)
 - Generate FAB fragments (in-house possibilities with biologics group)

- 'Biophysical' approach
 - Limited proteolysis
 - Protein purification
 - Compound binding
 - DSF
 - NMR

A real example ...

Using biophysics: prioritization of protein construct

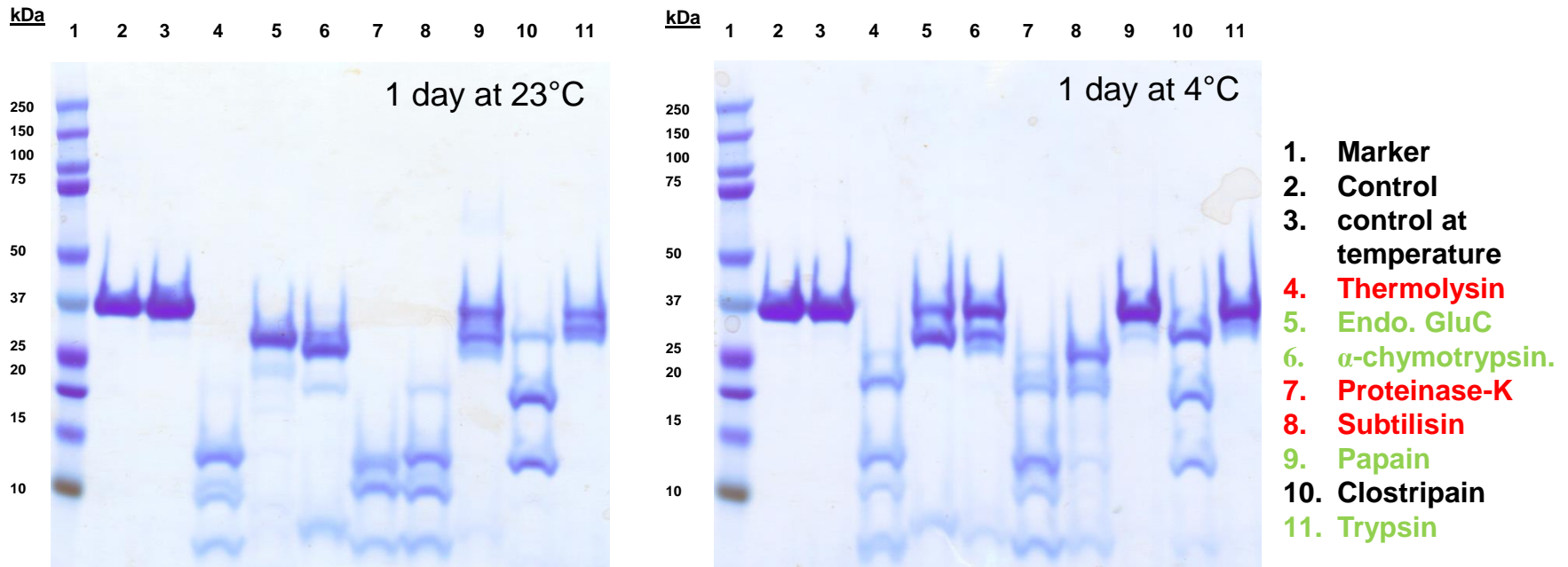
- Which protein construct to select for detailed analysis?
 - Protein behavior during purification
 - Protein folding (limited proteolysis, NMR)
 - Melting temperature (DSF)
 - Aggregation state (DLS)

 - Bonus info: FAB-protease complex structures obtained with construct-1 !

- Protein expression in BEV vs *E.COLI*?
 - If no difference, prefer *E.Coli*
 - Faster turnaround
 - Easier up-scaling
 - Isotope labeling for NMR

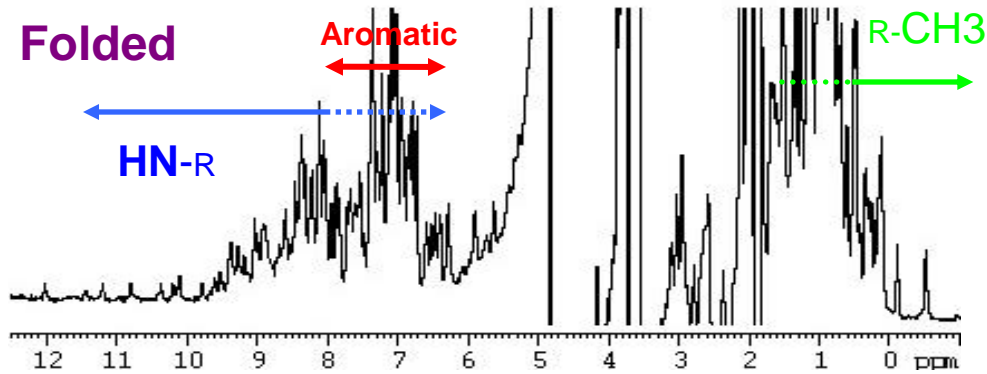
A real example ...

Using biophysics: prioritization of protein construct (limited proteolysis)



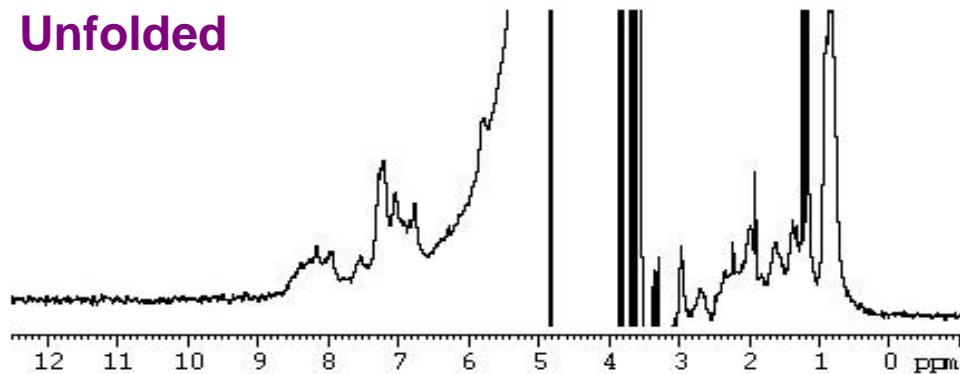
- Thermolysin, proteinase K and subtilisin – proteases with a low specificity – degrade the protein under the conditions used
- 30 kDa, stable fragment obtained with Endo GluC, α -chymotrypsin, papain and trypsin
 - Specific proteases cleave consistently at the C-terminus
 - No cleavage event observed in the N-terminal region

Some background: Protein folding as seen by NMR



MMP12; Mw: 18 kDa

- + chemical dispersion in methyl region
- + chemical dispersion in amide region
- + sharp resonances (=unique conformation)



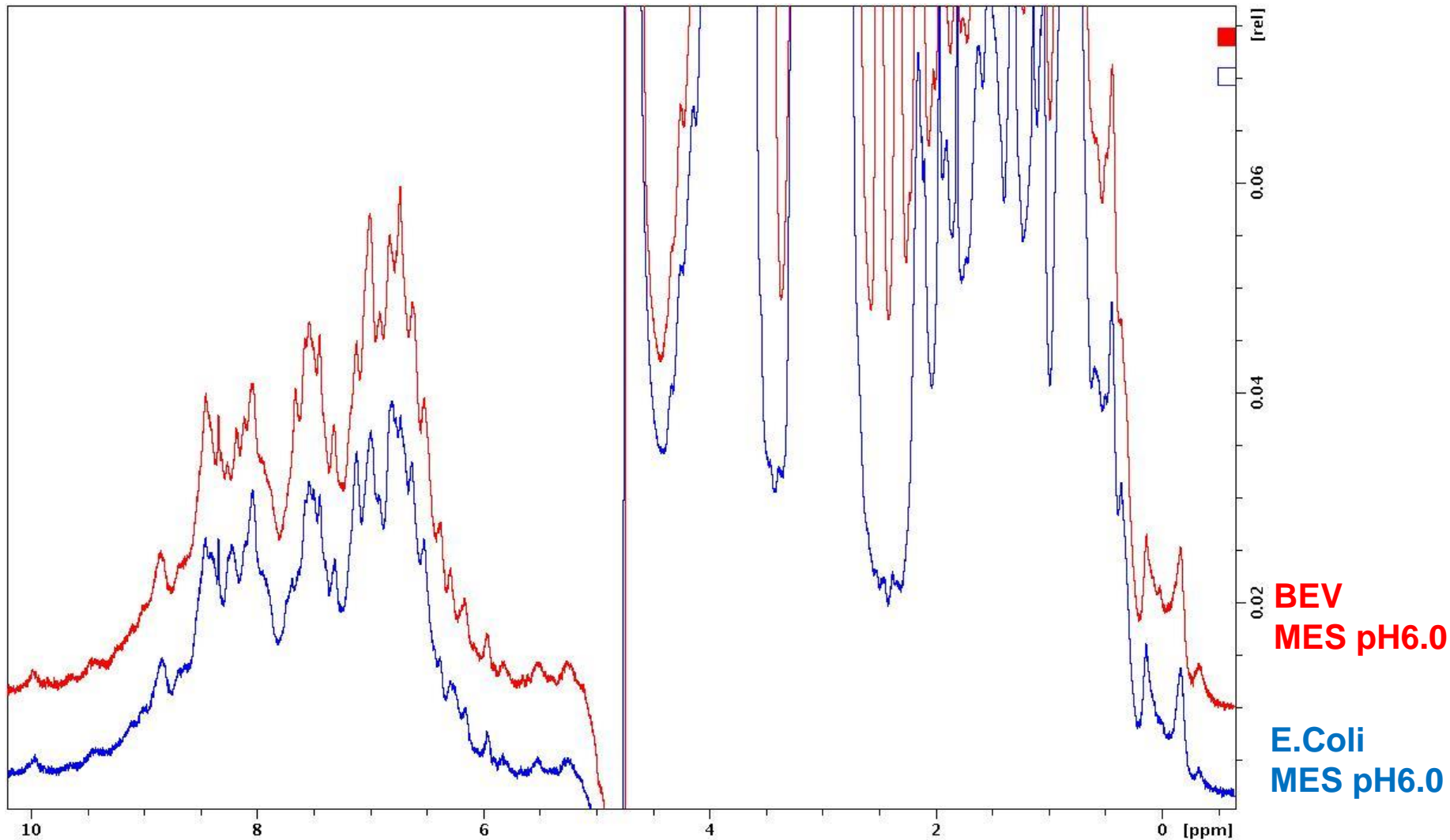
MMP9; Mw: 18 kDa (no fibronectin)

- poor chemical dispersion in methyl region
- poor chemical dispersion in amide region
- broad resonance (=multiply conformations)

- Simple experiment: requirements ~0.2mg of protein, Mw < 50kDa
 - Protein can be reused after NMR (concentrate for crystallization?)
- Folding assessment by NMR: situation often not black or white

A real example ...

Using biophysics: prioritization of protein construct (NMR)



A real example ...

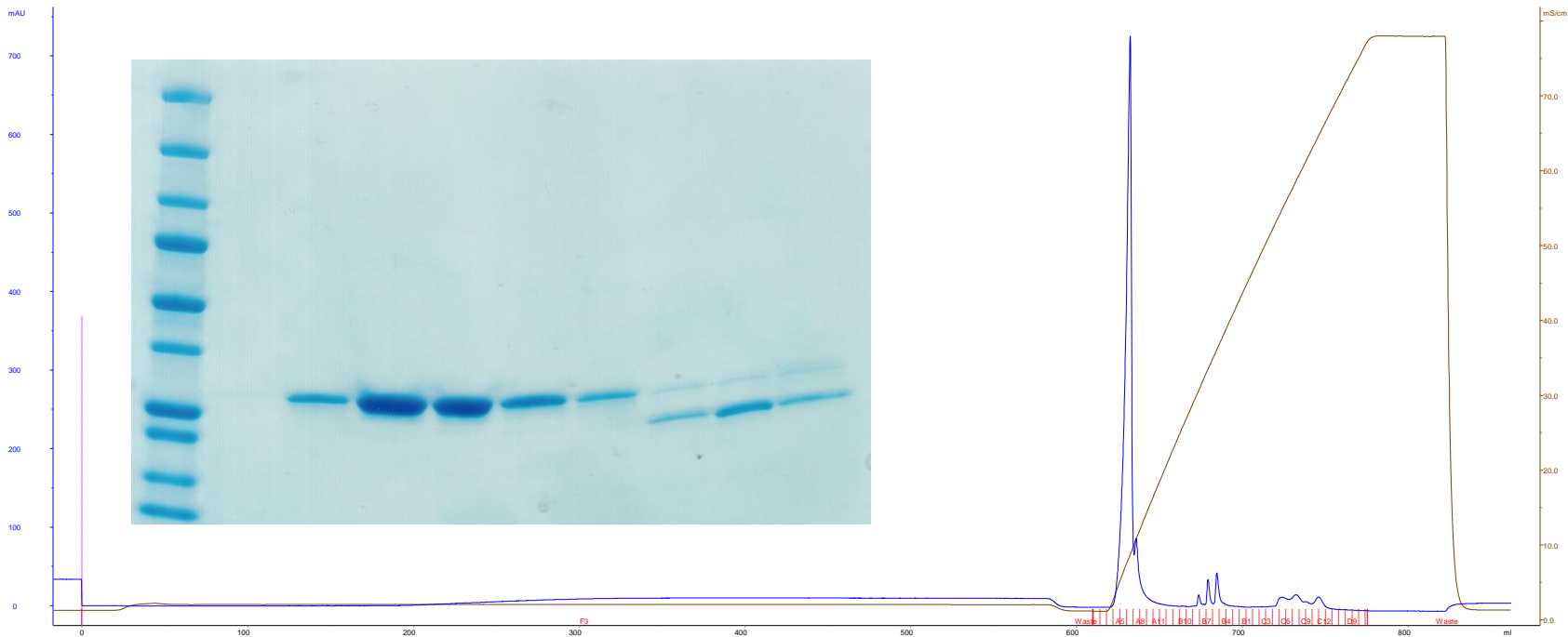
Using biophysics: prioritization of protein construct

Construct-1 expressed in *E.coli* is fine!

A real example ...

Using biophysics: SOP protein purification

- Switch expression system to *E.Coli* and scale up (24L batches)
- SOP protein purification: 3 purification steps
 - Affinity purification (typically Ni-NTA)
 - Ion exchange (or HIC)
 - Size exclusion: only polishing step (=analytics, buffer exchange)



A real example ...

Using biophysics: SOP protein purification

Improved purity: often not obvious on a gel / MS

- Removal DNA/RNA, others?

Improved homogeneity: often not obvious from DLS or size exclusion profile

In-house examples where SOP made the difference:

D'Arcy, Acta Cryst. F62 (2006); Erbel, Nat. Struct. Biol. 13 (2006)

Wiesmann, JMB 419 (2012)

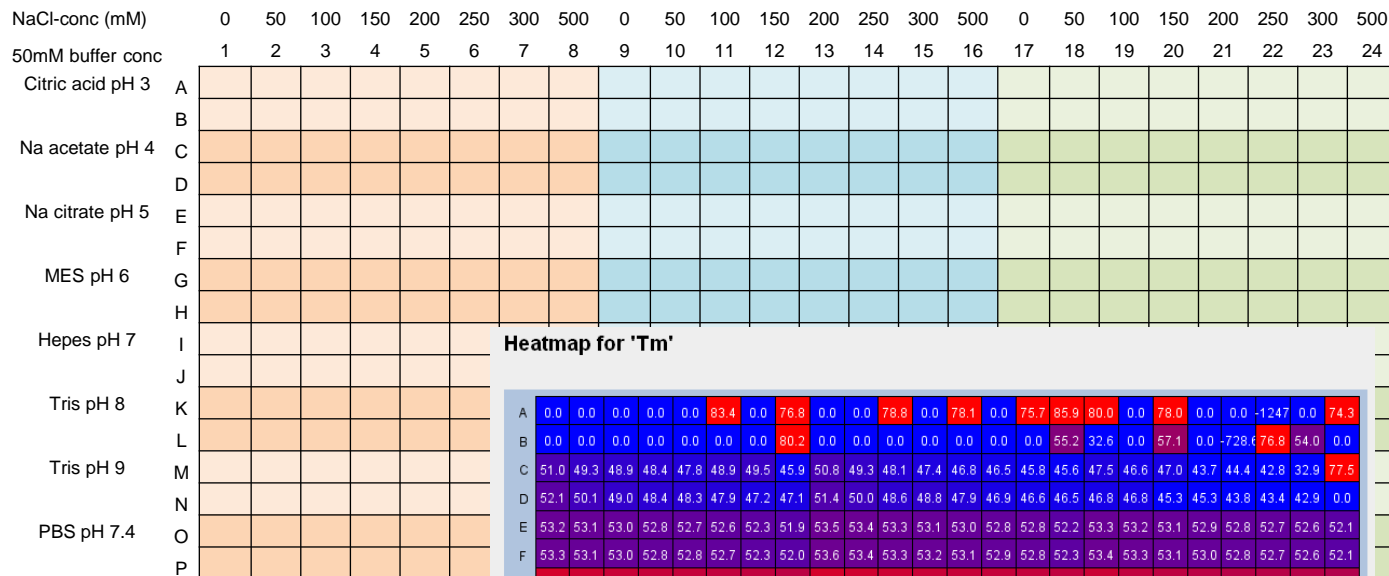
A real example ...

Using biophysics: Buffer screen using Differential Scanning Fluorimetry

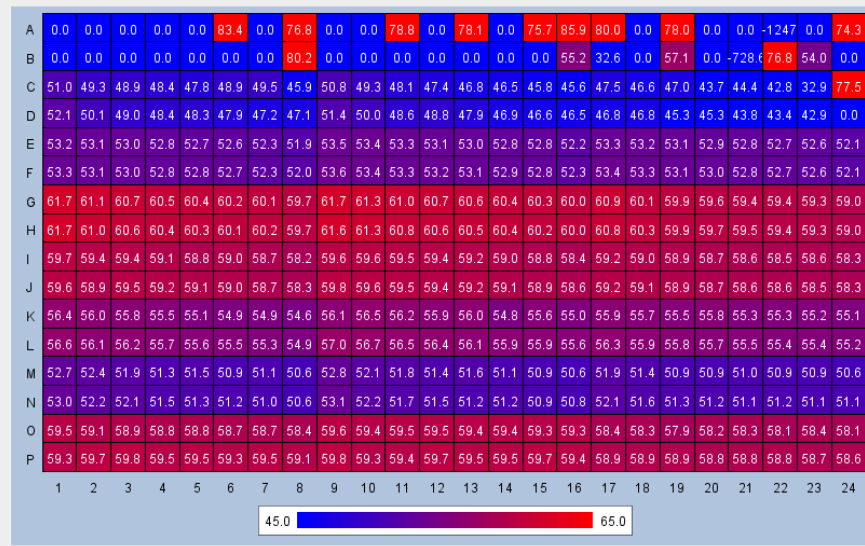
Buffers with
pH 3-9

NaCl-conc.
0-500mM

duplicates



Heatmap for 'Tm'



Highest Tm ~62°C obtained in
MES pH 6.0 with 0-50mM NaCl

(theoretical IEP: 7.0 – 6.8)

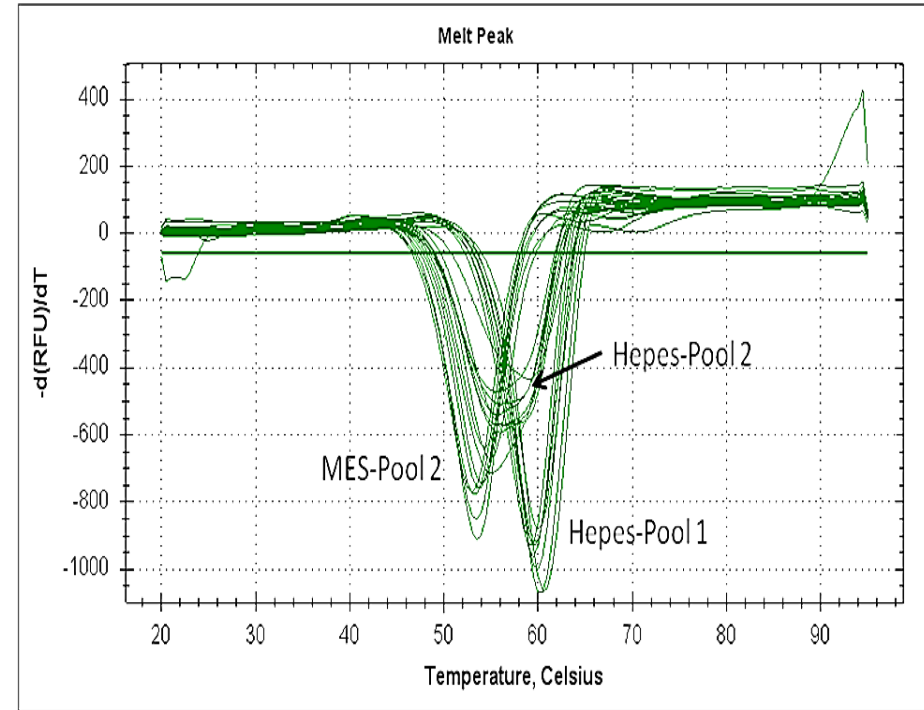
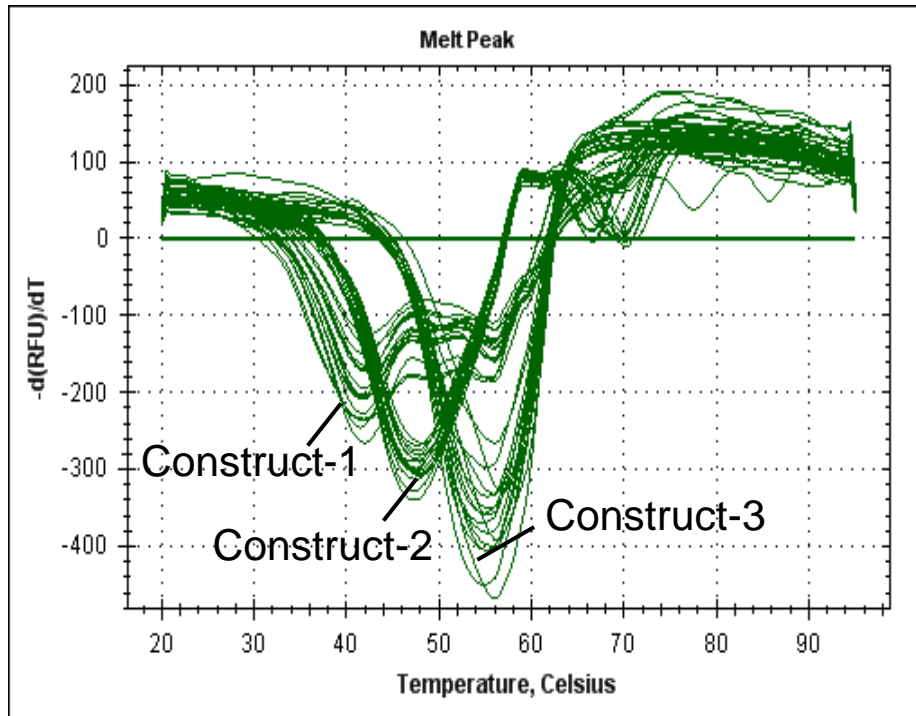
A real example ...

Using biophysics: Buffer screen using Differential Scanning Fluorimetry

Black and white results?

A real example ...

Using biophysics: Differential Scanning Fluorimetry comparing protein batches



- Different constructs, same conditions:
 - 50mM MES pH 6.0, 0-500mM NaCl

- Same construct(-1) different preparation:
 - Comparing fractions from IEX
 - Pool 1: sticks to IEX
 - Pool 2: does not stick to IEX

A real example ...

Using biophysics: Differential Scanning Fluorimetry comparing protein batches

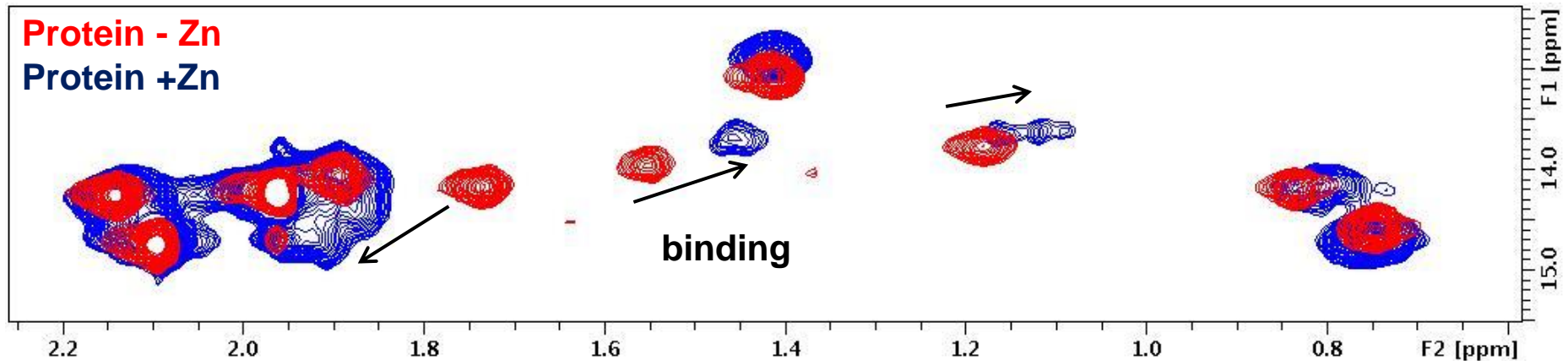
Black and white results?

No – what is going on?

A real example ...

Using biophysics: Measuring Zn^{++} binding using NMR

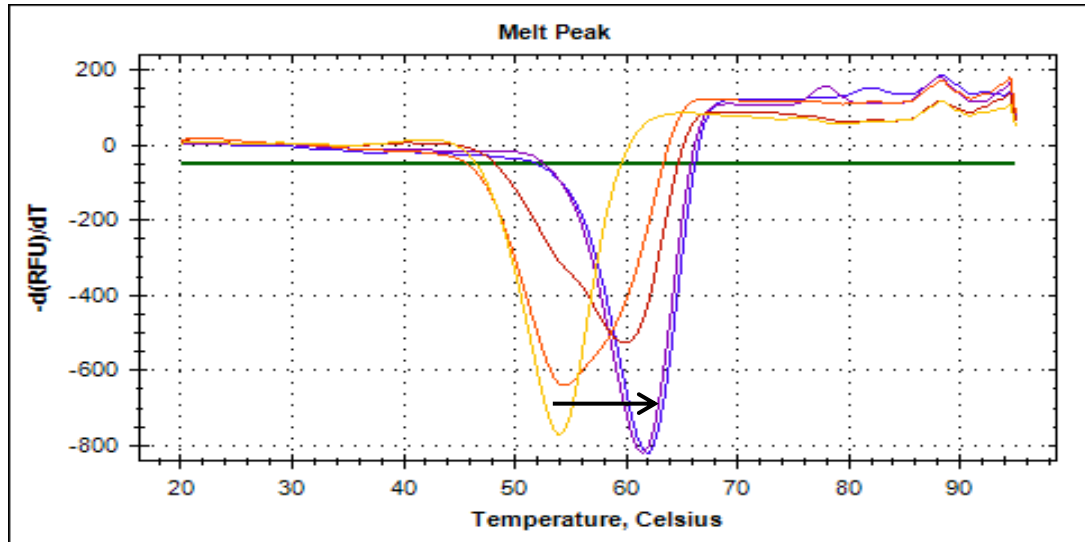
- ^{13}C -Methionine labeling for NMR
 - Fingerprint of protein
 - Similar way as selenomethionine labeling for crystallization
 - ^{13}C -methionine is not toxic



- Conclusion: pronounced effects on the Methionine signals upon addition of Zn^{++}
 - Active-site Zn^{++} ?
 - Second Zn^{++} binding site?

A real example ...

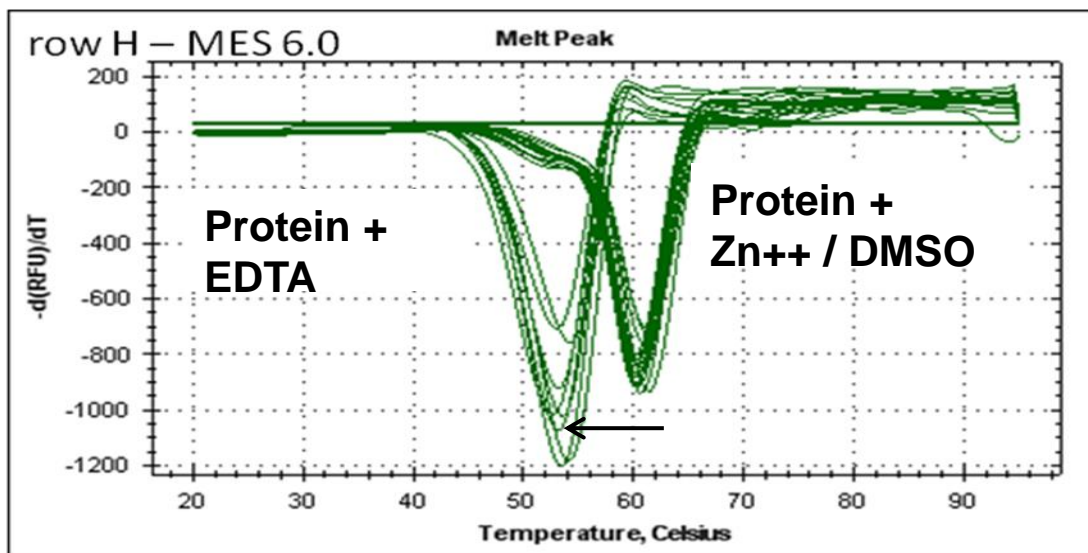
Using biophysics: Measuring Zn⁺⁺ binding by DSF



Zinc (μM)	Tm (°C)
0	54
1	54.5
2	59.6
4	61.3
8	62

+8°C

T_m stays at 62°C for zinc-conc. up to 128μM



A real example ...

Using biophysics: Summary Zn⁺⁺ binding

- Zn⁺⁺ atom lost during purification
 - Zn⁺⁺ loading depend on expression yield, expression system and degree of purification
 - iPO4 buffer had a negative effect (mixing with FABs delivered in PBS)
- Mind set: Initially not aware of this
 - Biased based on experiences with other metal proteases (like TLL, MMPs)
 - Protein reasonable behavior, but rather heterogeneous
- Zn⁺⁺ problem identified using DSF and NMR
 - Zn⁺⁺ loading step added during concentration (excess of Zn precipitates protein)
 - Buffer condition changed:

50mM HEPES pH 7.5
150mM NaCl
1mM TCEP



50mM MES pH 6.0
50mM NaCl
1mM TCEP
10μM ZnCl₂

A real example ...

Using biophysics

All set not?

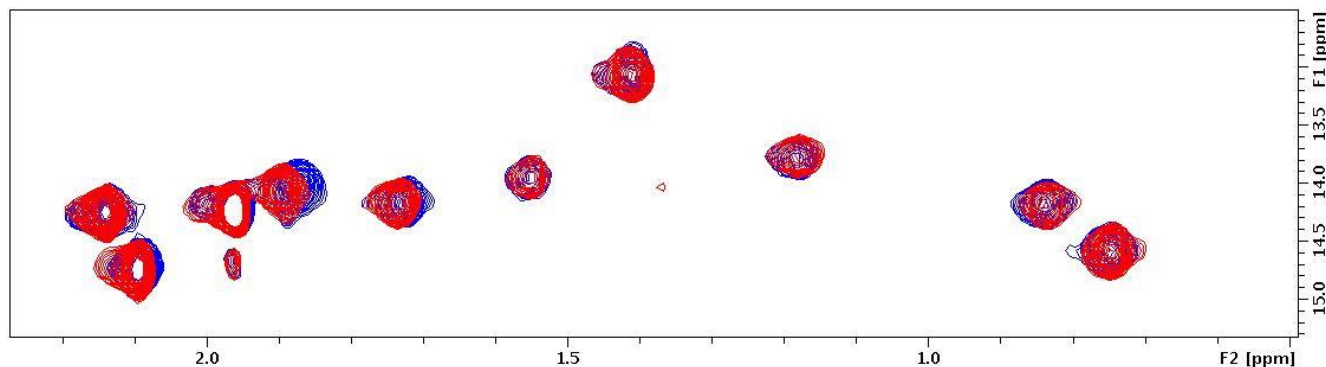
No, still no crystals

✓ Added Zn^{++}

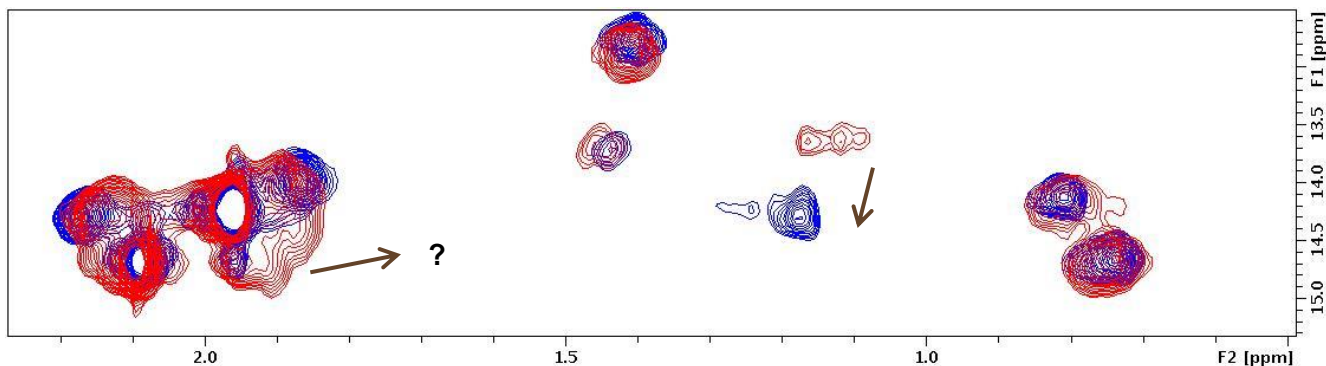
=> Add: compound!

A real example ...

Using biophysics: compound binding as seen by NMR



Protein w/o zinc
Red: apo
Blue: + compound



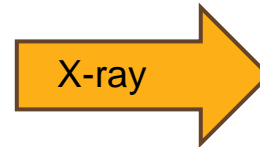
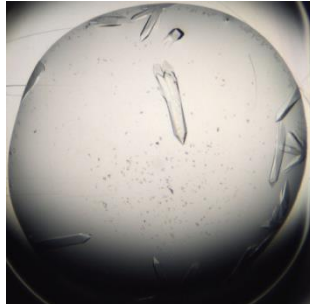
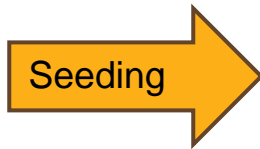
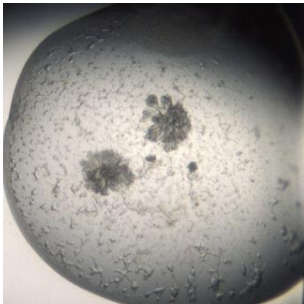
Protein with zinc
Red: apo
Blue: + compound

Compound does not bind to **zinc free** protein

Compound does bind to **zinc loaded** protein

A real example ...

Using biophysics: crystals obtained!

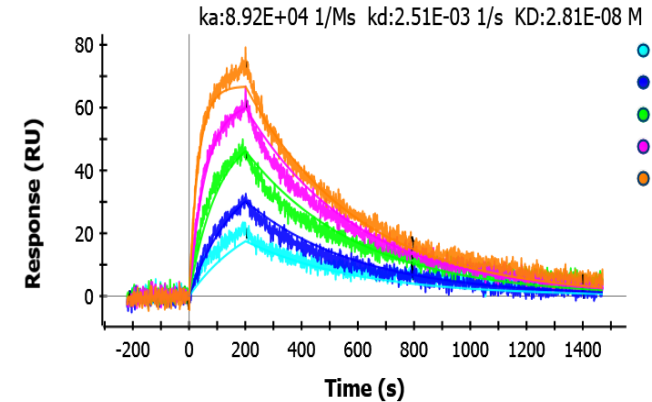


Structure
obtained

A real example ...

Using biophysics: crystal obtained

- Expression system: E.coli
 - Greater flexibility
- Select construct and focus
 - Enabled us to crystallize the first construct
 - Protein purification (no crystals obtained without IEX step)
- Optimize buffer conditions (DSF)
 - MES pH6.0 buffer
 - Zn⁺⁺ loading
- Add binding partners
 - Crystallization in presence of compounds, selection based on NMR/SPR data
- Crystallization
 - crystallization setup and plates storing done at 4°C (room temperature did not work)
 - Additional screens were tested: MPD condition
 - Seeding



Different Topic: selection of ligands for crystallization

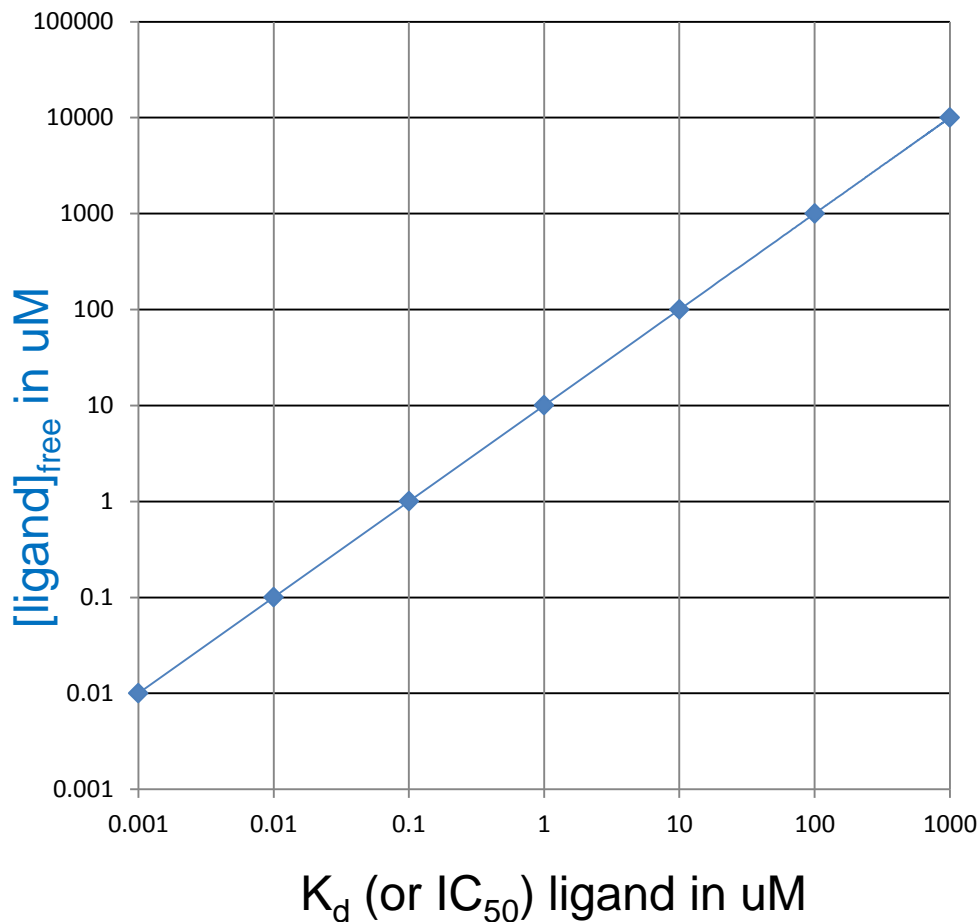
- Ligand binding often important for crystallization
 - Structural rearrangements upon binding
 - Chemist like them too

- How can we select the best ligands for cocrystallization?
 - Critical parameters
 - Binding to target
 - Potency, stoichiometry
 - Solubility
 - Also ligand purity, stability

Objective: >90% protein is occupied by ligand

Ligand concentration?

Correlation between K_d and ligand concentration to achieve 90% protein occupancy



$$\text{fraction occupied protein} = \frac{[Ligand]_{free}}{K_d + [Ligand]_{free}}$$

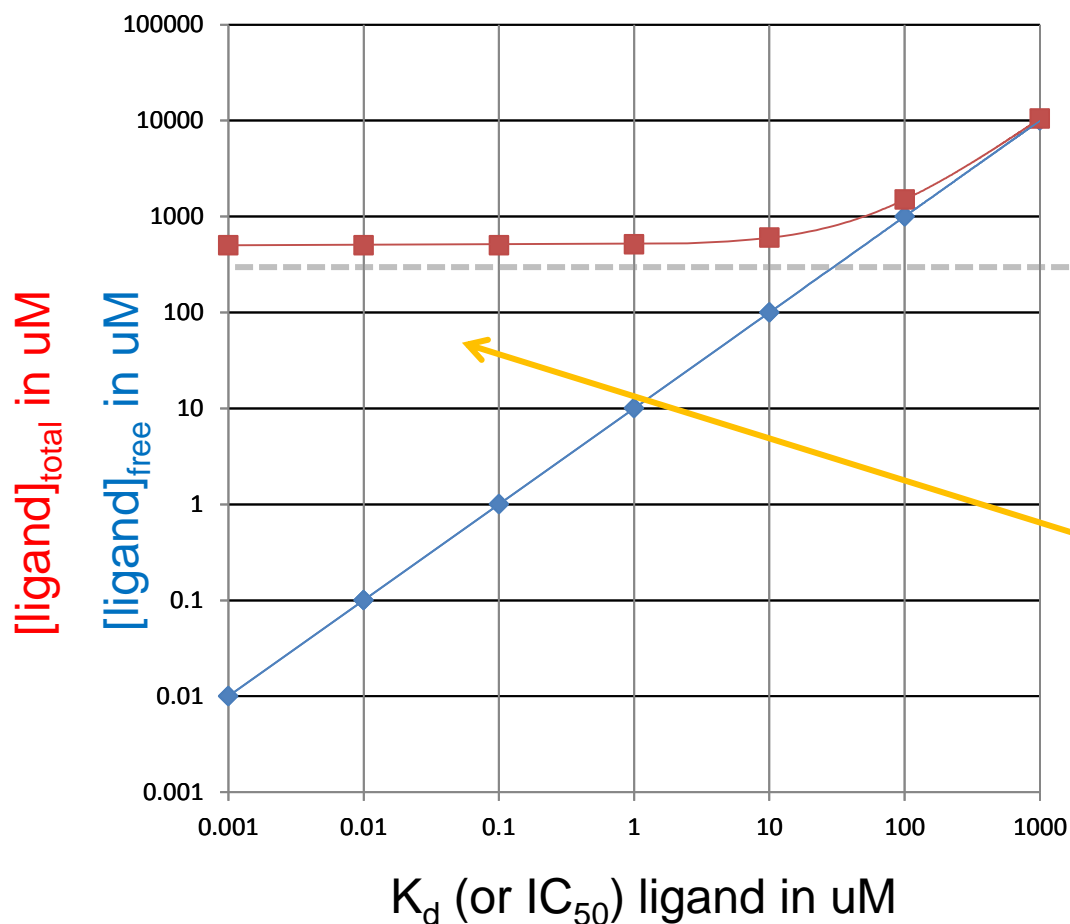
If fraction bound > 0.9
 $K_d < 0.11 [Ligand]_{free}$

90% of protein occupied if
 $[Ligand]_{free} \sim 10 \times K_d$

Objective: >90% protein is occupied by ligand

Ligand concentration?

Correlation between K_d and ligand concentration to achieve 90% protein occupancy



Crystallization ~ 10mg/ml
~ 200uM (50kDa protein)

Experimental details are important

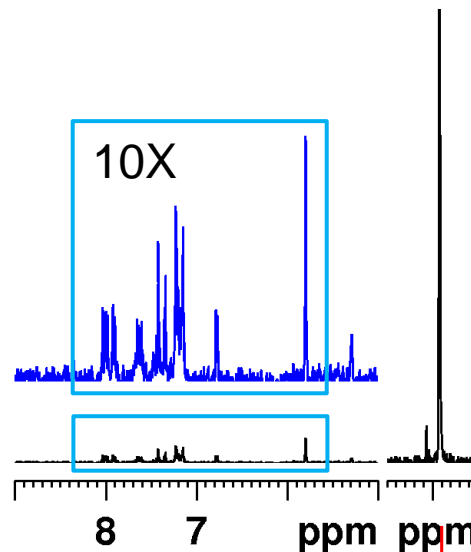
- Proper mixing of DMSO stock with protein solution
- Time dependence
- Work dilute!

Objective: >90% protein is occupied by ligand

solubility determination in buffer

■ Impact of solubility determination

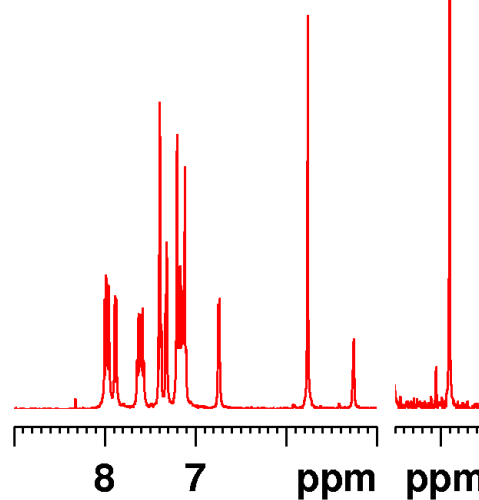
- ratio IC_{50} vs solubility
 - problematic if solubility < IC_{50}
 - IC_{50} valid? What is the assay measuring?
- 'absolute' solubility
 - Prioritization of compounds
 - Design co-crystallization experiments (solubility > $10 \times K_d$)



Cpd in PBS pH7.4
- Solubility ~65uM

■ Compound solubility by NMR

- reliable method in our experience
- acceptable throughput (2-30min / sample)
- solubility can be determined >5uM
- good range of buffers (including detergents)



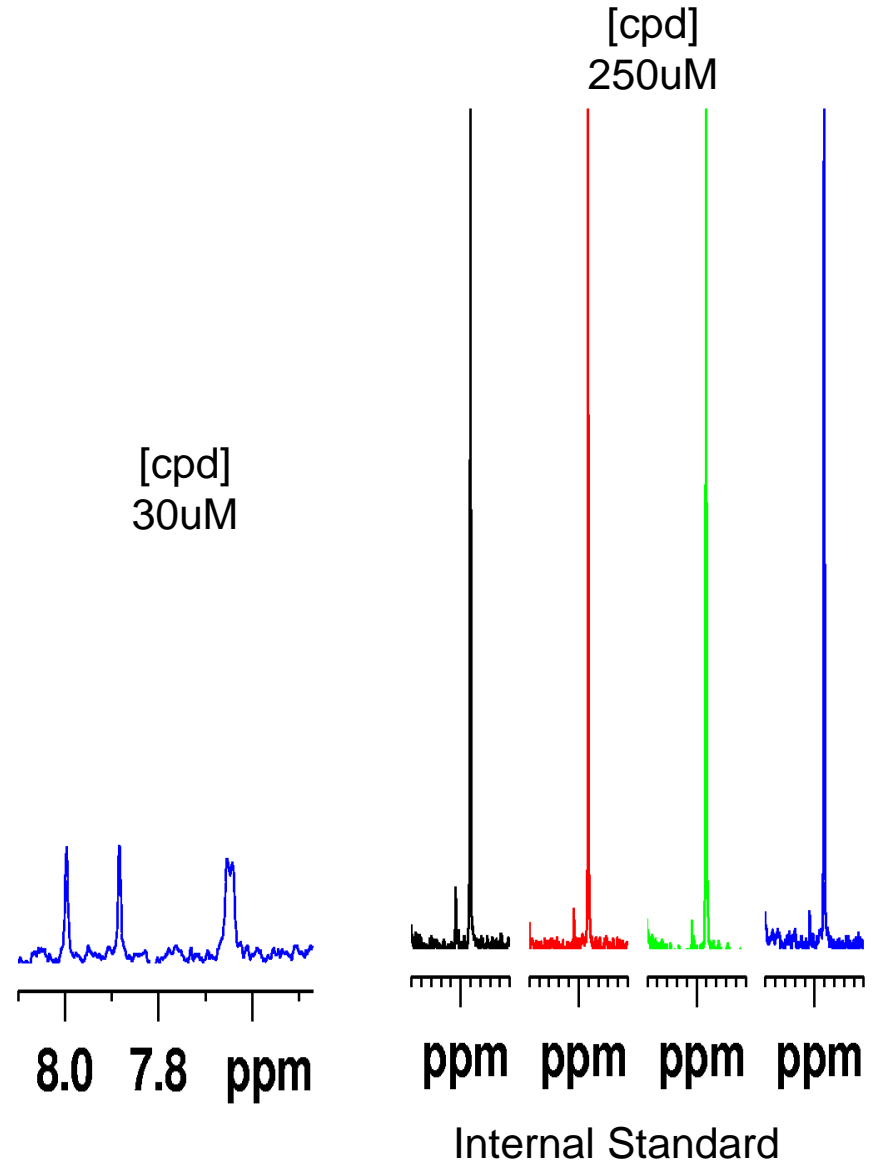
Cpd in Acetate pH4.5
- Solubility >1000uM

⇒ Compound solubility:
low tech, high impact

Objective: >90% protein is occupied by ligand

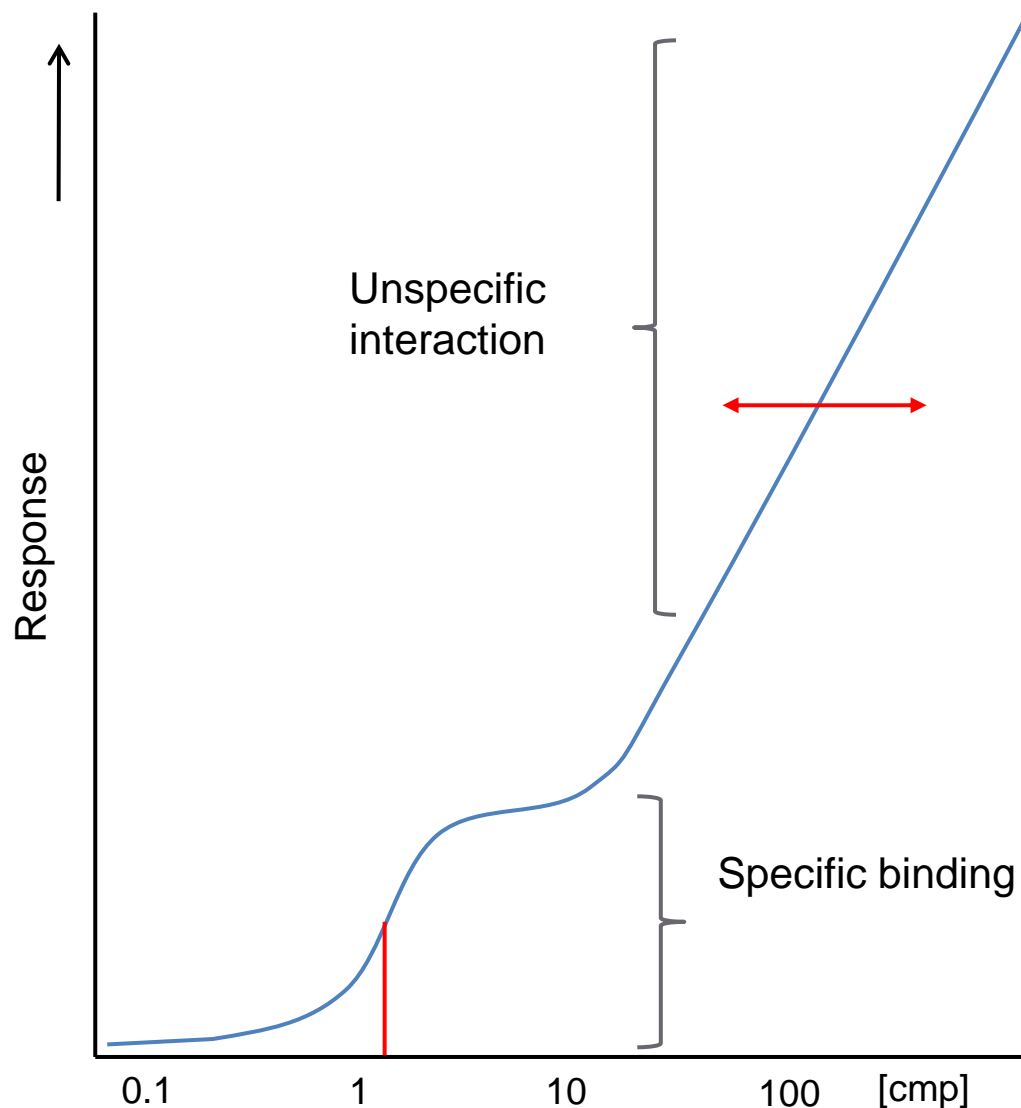
solubility determination in buffer

- Compound solubility by NMR
 - Illustration of micel formation



Solubility and potency: specific vs unspecific binding

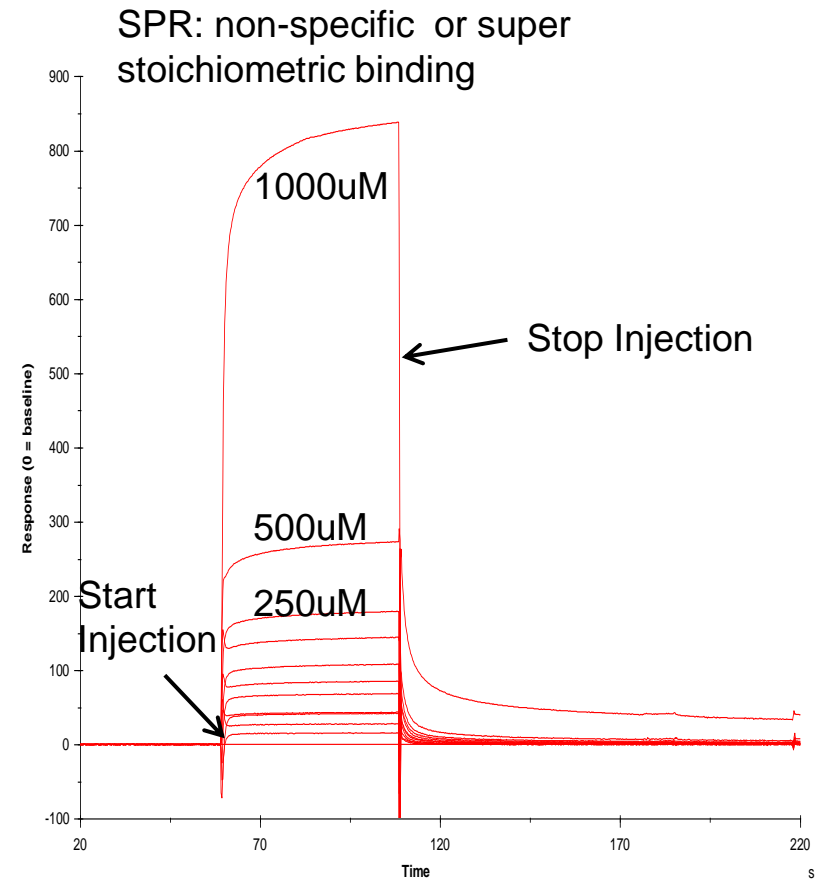
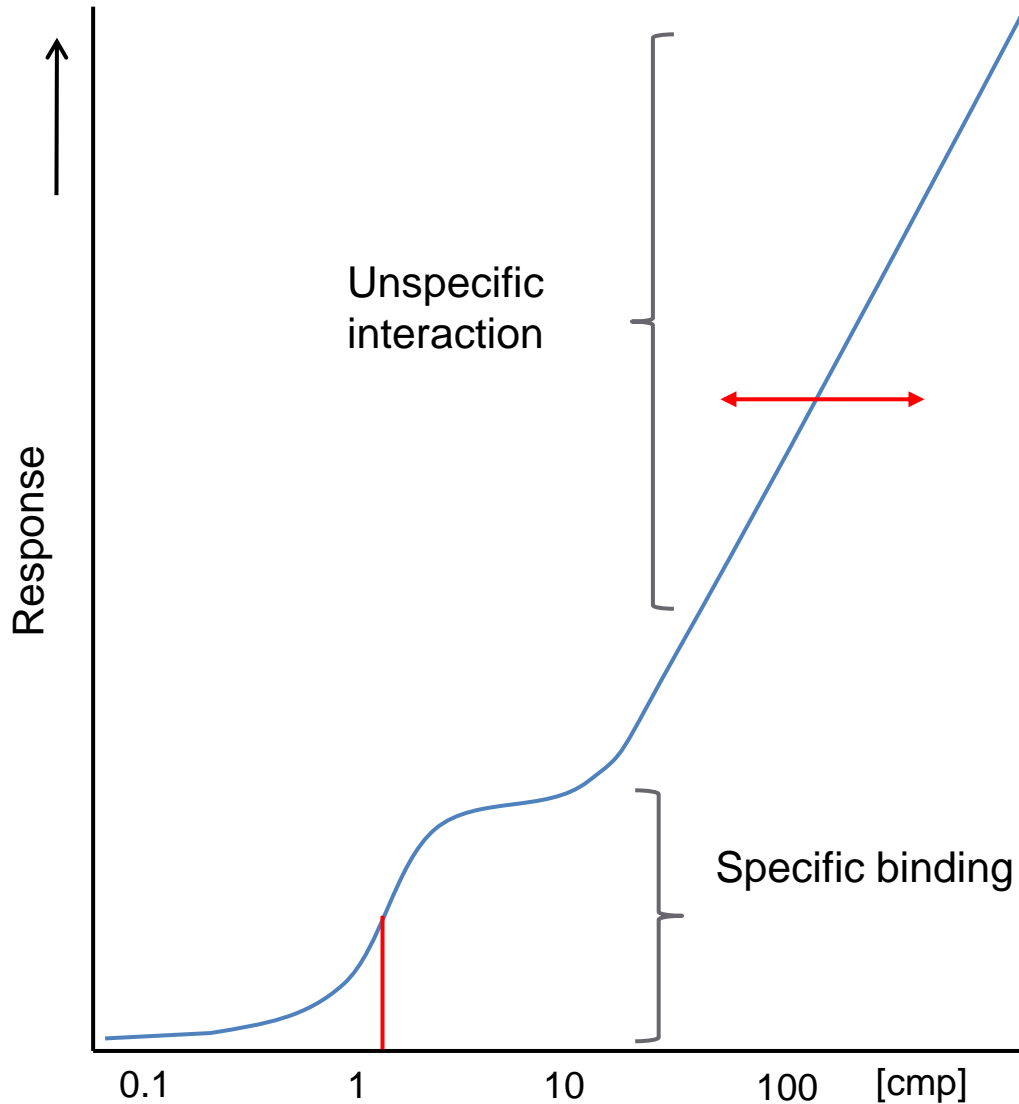
different technologies



- Separation between specific and unspecific binding is cmpd dependent
- Assays often measure at different concentration range
 - >90% protein occupancy at 10 x IC₅₀
 - Protein concentration
Crystallization / NMR > DSF > SPR > Assay
- Response unspecific binding often larger than specific binding
 - NMR: line broadening, signal reduction
 - DSF: no transitions
 - SPR: response 10x bigger

Solubility and potency: specific vs unspecific binding

different technologies



Biophysical characterization for compound binding

Affinity (K_D), Specific or Unspecific binding

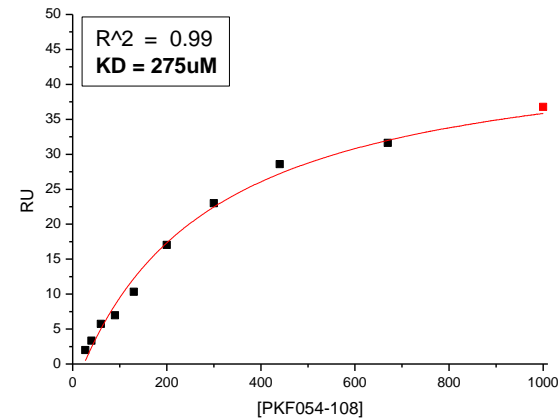
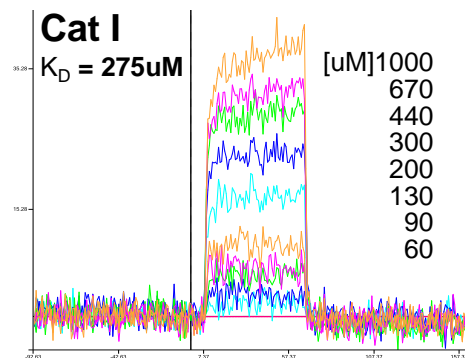
Category 1:

- NMR binding (yes)
- SPR: 1-1 binding with K_D

>60%

⇒ NMR and SPR aligned

⇒ First priority for crystallization

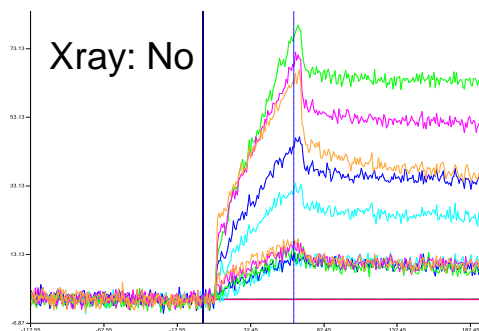
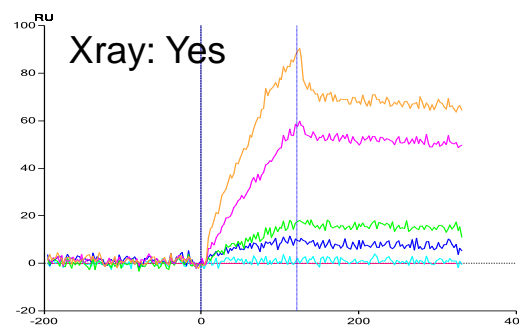


Category 2:

- NMR binding (yes)
- SPR: response
 - slow kinetics, super stoichiometric
 - Binding to surface?

~20%

⇒ Low priority for crystallization



Lessons learnt:

- SPR: strict filter for crystallization (solubility, cpd aggregation, non specific binding, affinity)?
- Combination of NMR and SPR links robustness, sensitivity and hit characterization (reversed strategy?)



Thank you:
Daniela Ostermeier
Frederic Villard
Simon Ruedisser
Martin Renatus
Allan d'Arcy
Christian Wiesmann