# Applications of NMR spectroscopy in Drug Discovery

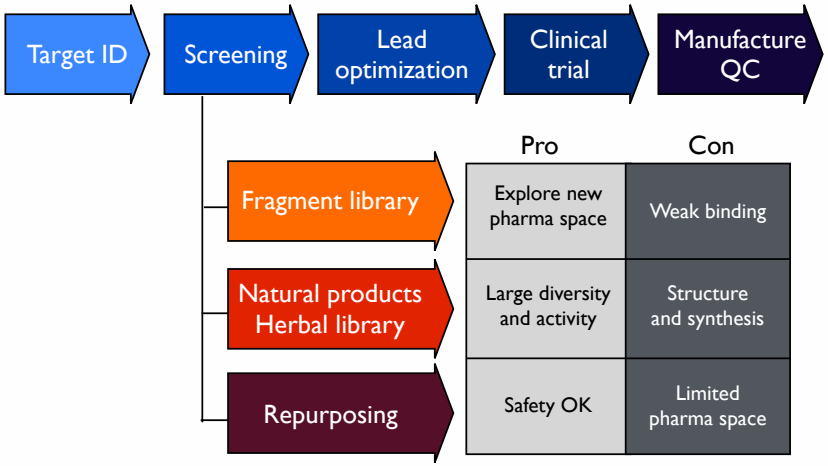
Shang-Te Danny Hsu
Institute of Biological Chemistry
Academia Sinica
Taiwan

2016.6.24 HFNMRC workshop on NMR screening approach



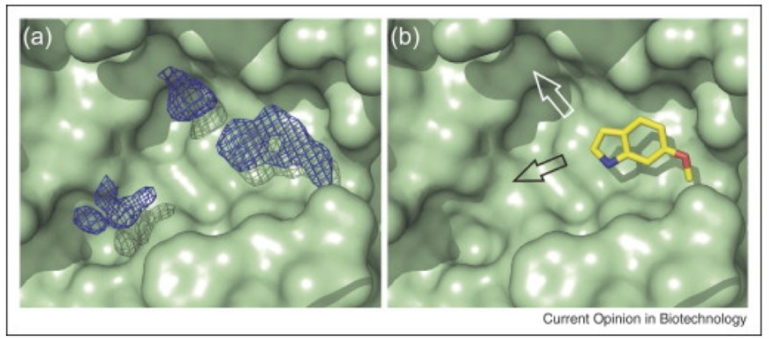


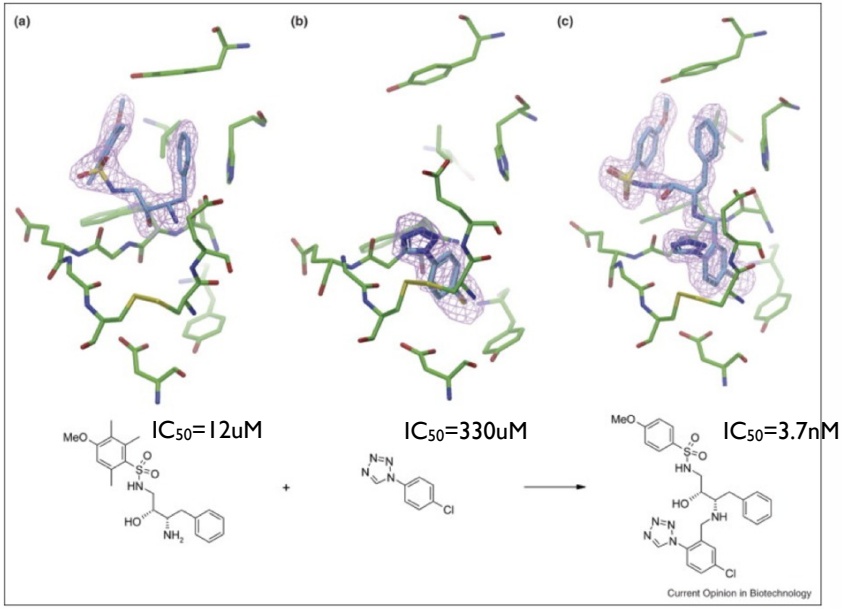

# Drug development pipeline





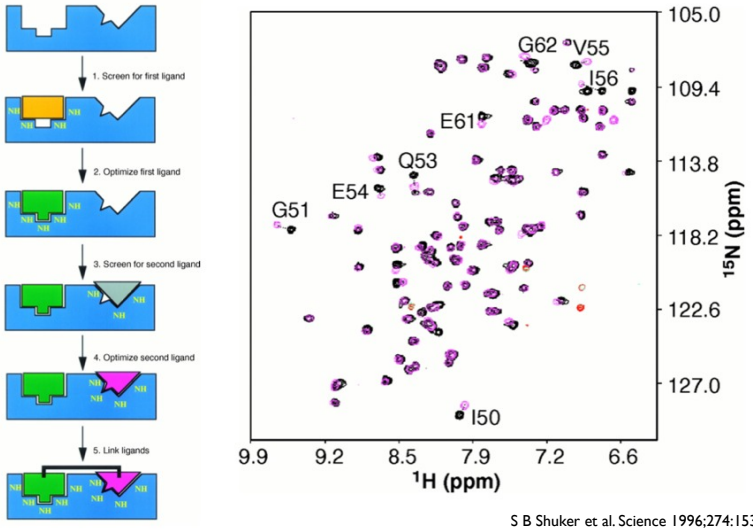

# Maximize ligand-protein complementarity by fragment-based drug development (FBDD)








# Structure activity relationship by NMR (SAR-NMR)




S B Shuker et al. Science 1996;274:1531-1534



# SAR in FBDD

Free energy of binding is associated with binding constant (or dissociation constant, $K_d$)

$$\Delta G = -RTlnK_d$$

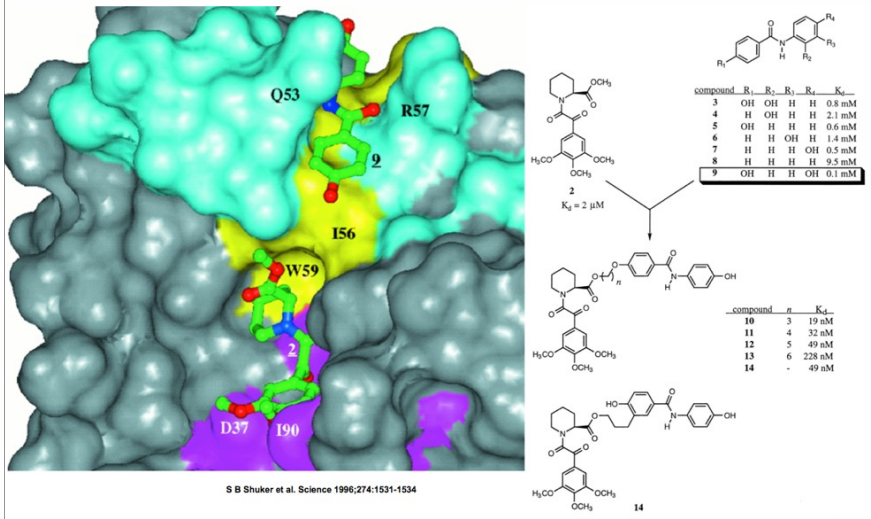
Free energy is additive. So if two fragments, A and B, bind to the target protein independently. Assuming that the linkage of the two fragments will not affect their binding to the target protein, then

$$\begin{aligned}\Delta G_{A+B} &= \Delta G_A + \Delta G_B \\ &= -RTlnK_{d,A} + -RTlnK_{d,B} \\ &= -RTln(K_{d,A} \times K_{d,B})\end{aligned}$$

The sum of the two contributions will enhance the binding affinity
Two sub-mM binders can become a nM binder

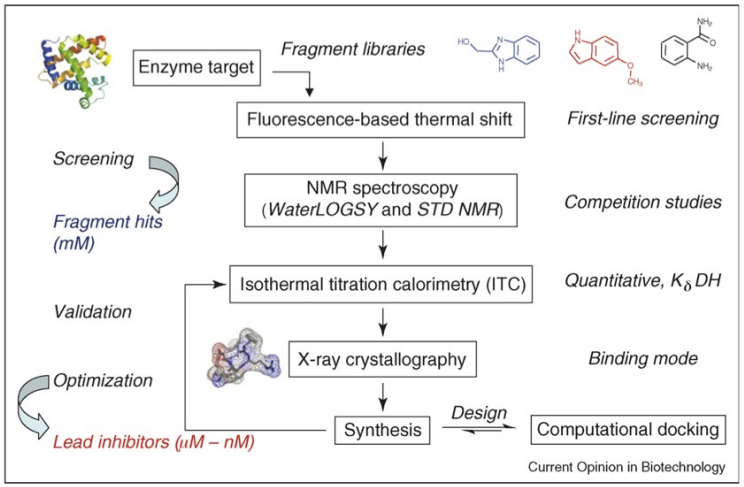


# A surface representation of FKBP showing the locations of 2 and 9, as determined from $^{15}N$-$^{13}C$ filtered NOE data




S B Shuker et al. Science 1996;274:1531-1534

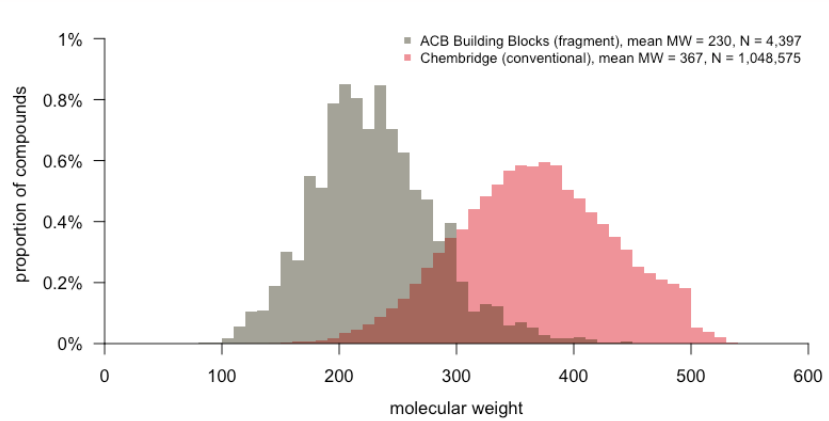
# Fragment-based drug development (FBDD)




Ciulli & Abell (2007) Curr. Opin. Biotech. 18:489



# Fragment characteristics




http://www.cureffi.org/2016/01/13/nmr-fragment-screening/



# Fragment screening tools

- X-ray crystallography
- Nuclear magnetic resonance (NMR) spectroscopy
- Surface plasmon resonance (SPR)
- Isothermal titration calorimetry (ITC)
- Biolayer interferometry (BLItz)
- Microscal thermophoresis (MST)
- Thermal shift assay
- Automated far-UV CD spectroscopy
- Alpha screen

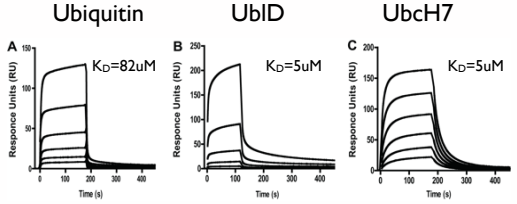


# Surface plasmon resonance (SPR)

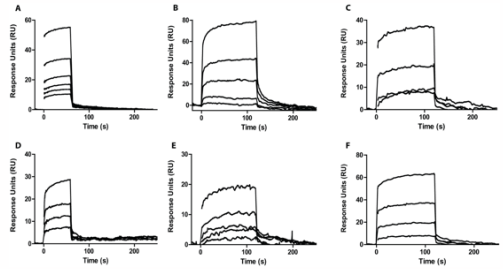
Biacore T200 w/ CM5 chip

Low MW fragments generally binds to protein targets weakly resulting in small response units and fast on- and off-rates

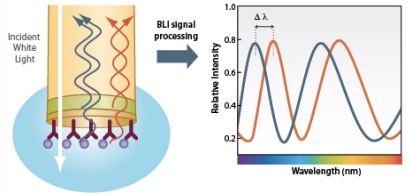
Regnström *et al. PLoS ONE* (2013)

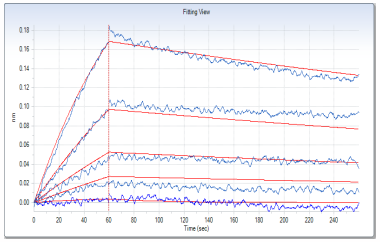


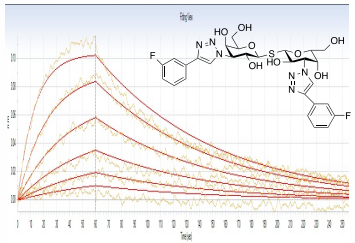

Fragment screening against Parkin



# Biolayer interferometry (BLItz)



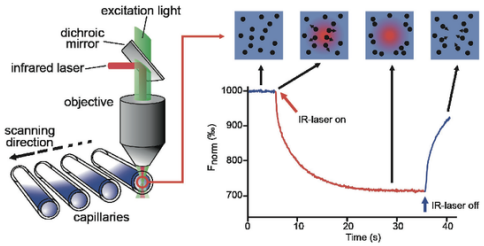


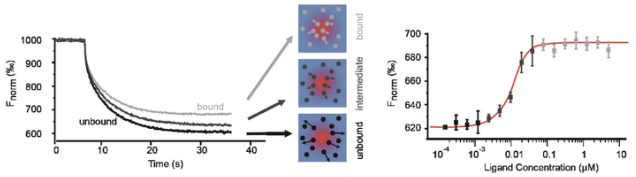
BLItz is similar to SPR but is less sensitive with the advantage of being able to recover samples and less expensive functionalized tips



# MicroScale Thermophoresis (MST)

www.nanotemper-technologies.com

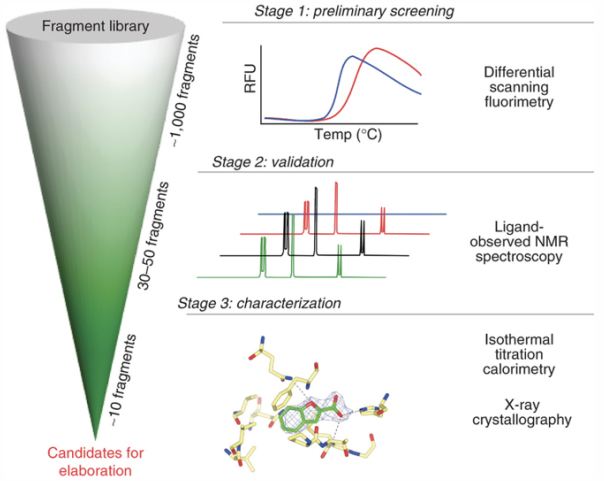



MST can probe biomolecular interactions across a wide-range of binding affinities (pM-mM) but it requires dye-labeling @ BCF-AS



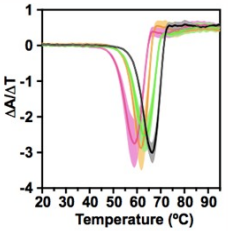
# A three-stage biophysical screening cascade for fragment-based drug discovery

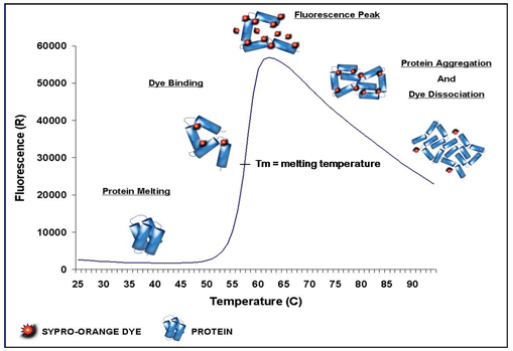



Mashalidis *et al. Nature Protocols* **8**, 2309–2324 (2013)



# Thermal shift assay

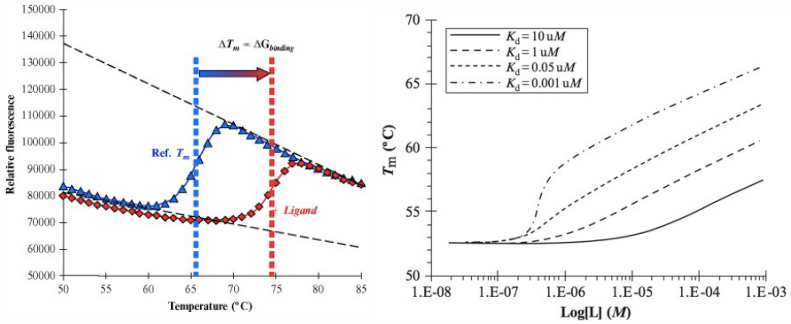



https://en.wikipedia.org/wiki/Thermal_shift_assay

Thermal shift assays are routinely used to optimize buffer conditions with different additives for crystalization and it can also be used in a high-throughput manner to scout for stabilizing fragments

# Thermal shift assay



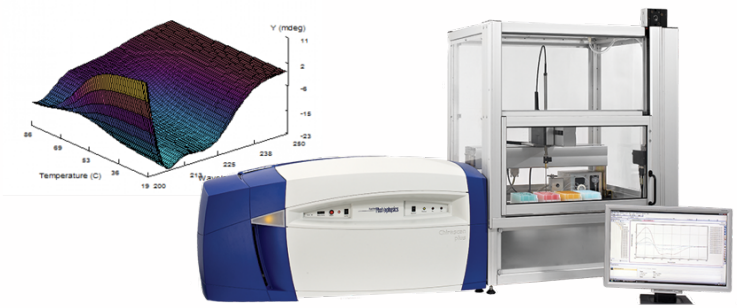

Weak binding (mM) requires excess fragments to result in relatively small increases in Tm (typically less than 2 °C)

Kranz & Schalk-Hihi (2011) Methods in Enzymology 493:277



# Automated CD spectrometer

- Samples presented in 96-well plate format, up to four plates
- Robotic transfer of sample from plate to spectrometer
- Wash-dry protocols eliminate dilution and cross contamination
- Low sample usage (<40$\mu$g protein per far-UV denaturation)
- Conformation and aggregation monitored by CD and absorption

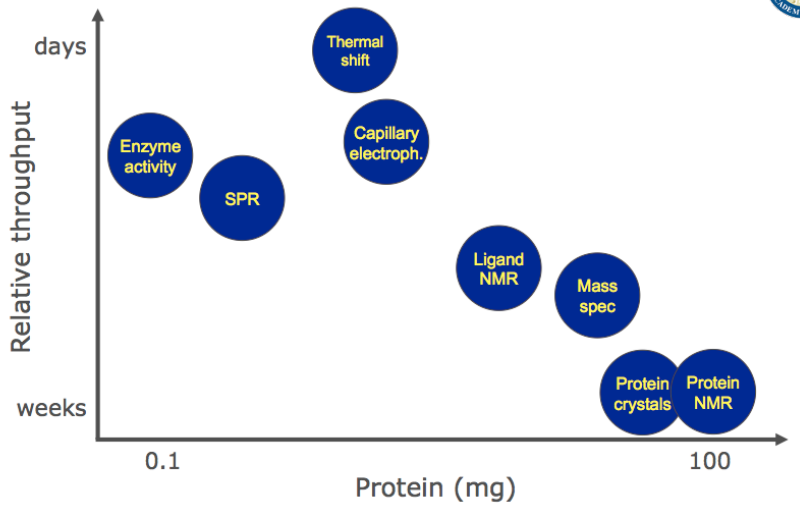




# Companies that offer screening services

Astex X-ray, NMR, ITC 1,600 compound library
Alveus Pharma "Details coming soon"
Beactica SPR 1946 compound library
BioFocus SPR 1,500 compound library, also offer X-ray and Mass Spec
Biosensor Tools SPR 1500 compound library
Carmot Therapeutics Chemotype evolution (tethering) Crelux Microscale thermophoresis X-ray 1,000 compound library
Crown Biosciences X-ray, SPR X-ray 3,400 compound library
Emerald BioStructures X-ray, NMR 1,500 compound library
Evotec Fluorescence correlation spectroscopy SPR, NMR, X-ray, biochemical **30,000** compound library
IOTA Pharmaceuticals SPR, biochemical, and NMR X-ray 5,500 compound library
Kinetic Discovery SPR 700 compound library
Kinomed X-ray Novalix SPR array MS, NMR **24,000** compound library
Pharma Diagnostics Bead-based SPR compound library
Polyphor SPR 1,000 compound library
Proteros TR-FRET X-ray 8,000 compound library
Selcia Capillary electrophoresis 1,300 compound library
Sprint Bioscience X-ray,
Structure Based Design X-ray, 1000 compound library
Vernalis NMR, biochemical, SPR, ITC X-ray 1,400 compound library
Viva Biotech X-Ray, NMR, SPR 2000 compound library
Zenobia X-ray, SPR X-ray 1,000 compound library
ZoBio TINSa NMR, SPR 1,500 compound library

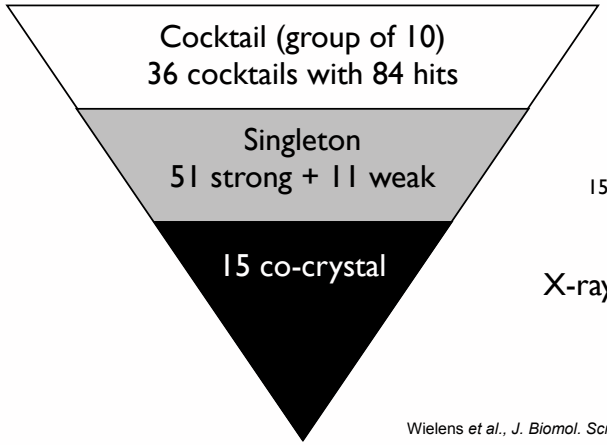





http://www.cambridgemedchemconsulting.com/resources/hit_identification/fragment_based_screening.html

# Protein target: HIV-1 integrase
# Approach A

A library of 500 compounds <MW>=198Da




STD-NMR

STD-NMR
$^{15}$N-$^{1}$H HSQC

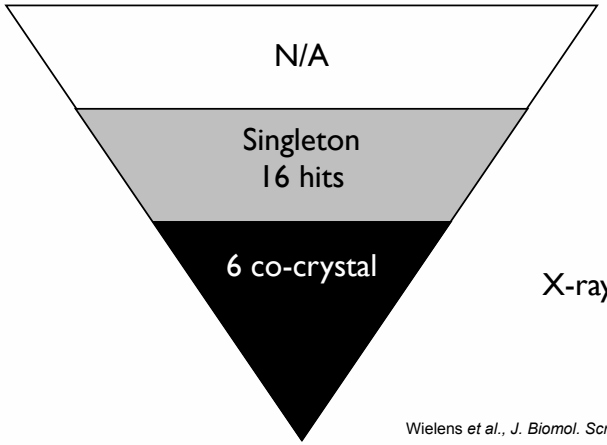
X-ray crystallography
Soaking

Wielens *et al., J. Biomol. Screening* 18(2) 147–159 (2013)



# Protein target: HIV-1 integrase
# Approach B

A library of 500 compounds <MW>=198Da



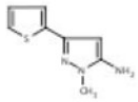

SPR

X-ray crystallography
Soaking

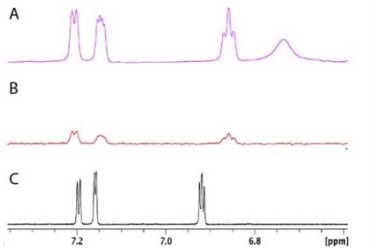
Wielens *et al., J. Biomol. Screening* 18(2) 147–159 (2013)



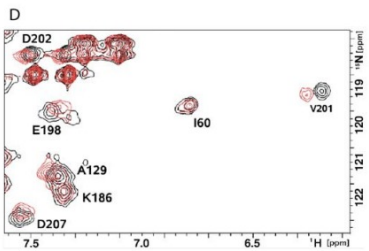
STD-NMR



10



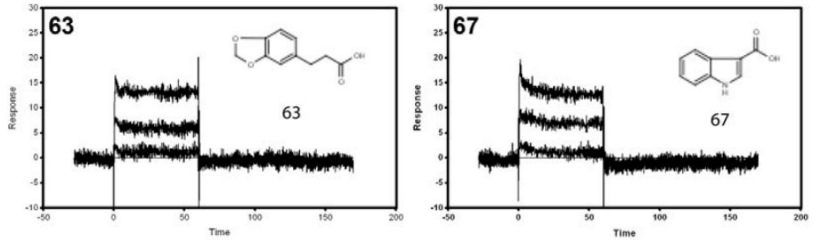

$^{15}$N-$^{1}$H HSQC




Wielens *et al., J. Biomol. Screening* 18(2) 147–159 (2013)



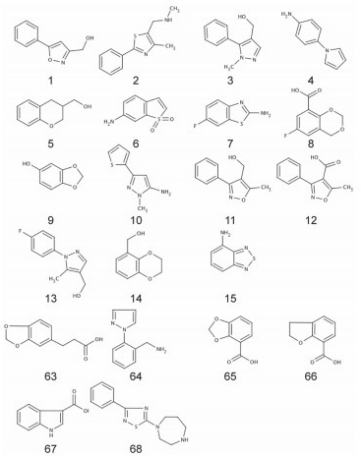
# SPR has poor S/N ratios for fragments and the binding kinetics is too fast to be accurately determined

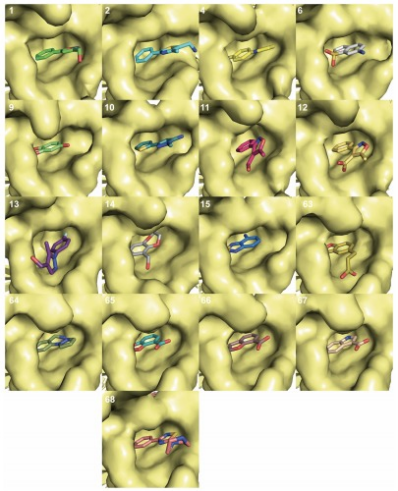



Generally a limited number of titration points are used for initial SPR-based screens and the steady state signal changes are used for rough estimation of dissociation constants

Wielens *et al., J. Biomol. Screening* 18(2) 147–159 (2013)

## All hits target the same pocket

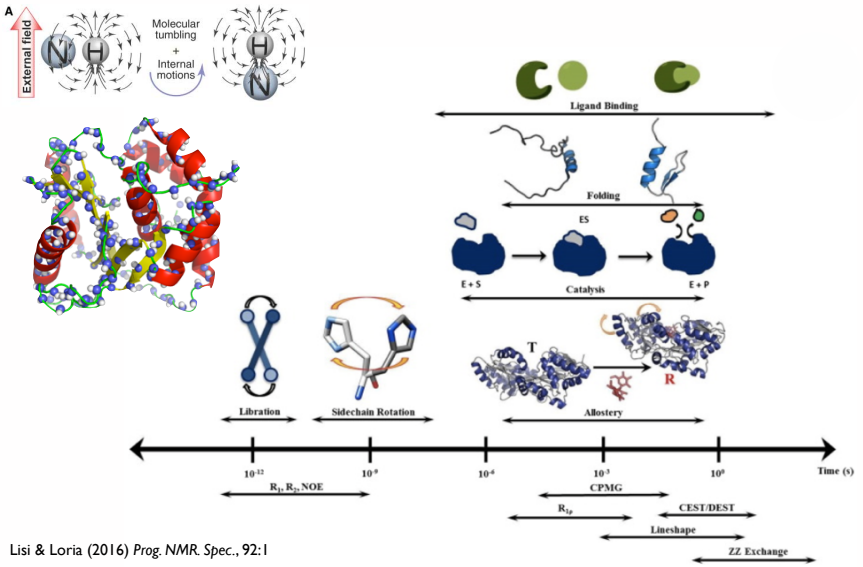




Wielens *et al., J. Biomol. Screening* 18(2) 147–159 (2013)

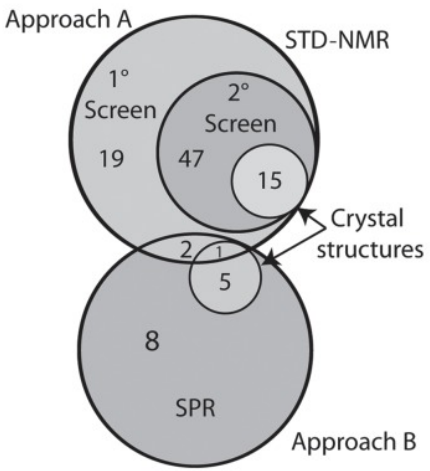


## Biomolecular dynamics probed by NMR




Lisi & Loria (2016) *Prog. NMR Spec.*, 92:1






A combined use of STD-NMR and $^{15}N$-$^{1}H$ HSQC in two stages (initial cocktails followed by singletons as in Approach A) increased the number of positive hits that eventually lead to a three-fold increase of co-crystal structures compared to individual SPR screening approach (Approach B)

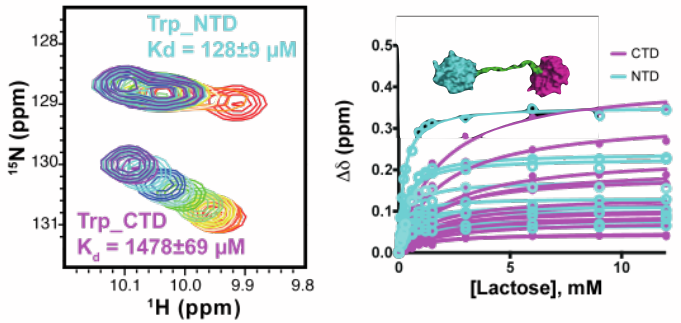
Wielens *et al., J. Biomol. Screening* 18(2) 147–159 (2013)



## Advantages of solution state NMR spectroscopy

- Provide structural and dynamic information at atomic resolution
- Broad range of experimental conditions can be accommodated: temperature, salt, buffer, additives (to be as close to physiological conditions as possible)
- Very sensitive to conformational changes, e.g., folding
- Ideal for weak and transient biomolecular interactions - particularly useful for FBDD

# Chemical shift perturbation as a probe for protein-ligand interactions

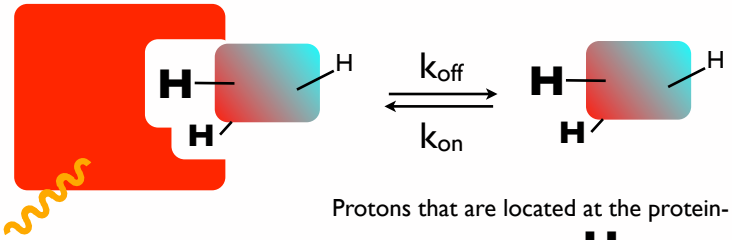





# Group epitope mapping (GEM) by saturation transfer difference (STD)

Protein-ligand complex

Free ligand



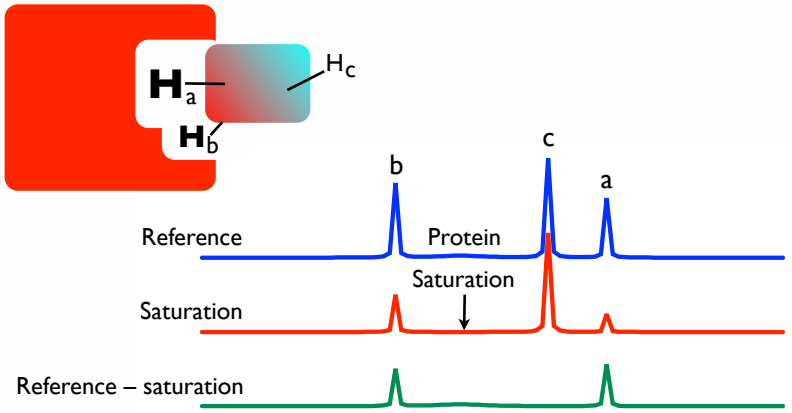

Selective saturation pulse

Protons that are located at the protein-ligand binding interface (**H** & **H**) will experience a higher degree of saturation transfer than those that are not involved in intermolecular interactions (H)

Mayer & Meyer *J.Am. Chem. Soc.*, **2001**, *123* (25), pp 6108–6117

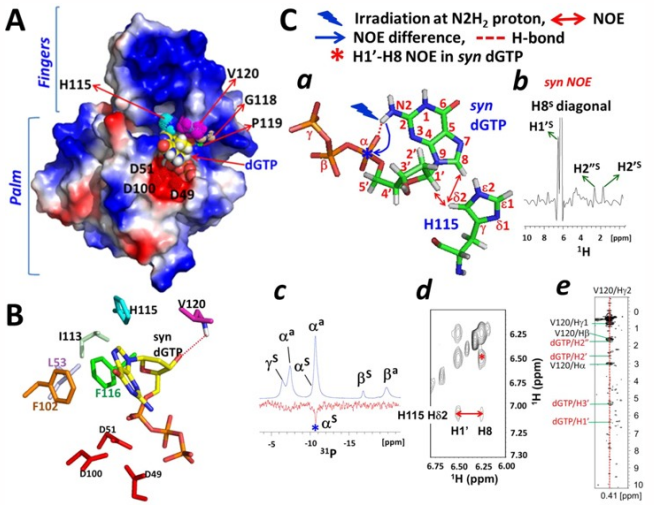


# Group epitope mapping (GEM) by saturation transfer difference (STD)



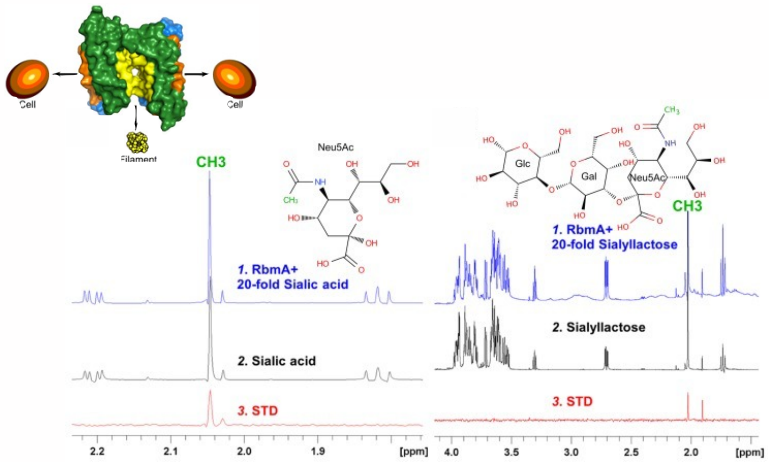



# NMR evidence for the Pol X:Mg dGTP binary complex.




Wu *et al. J.Am. Chem. Soc.*, **2014**, *136* (13), pp 4927–4937

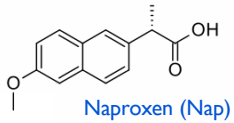
# STD of sialic acid and sialyllactose with RbmA



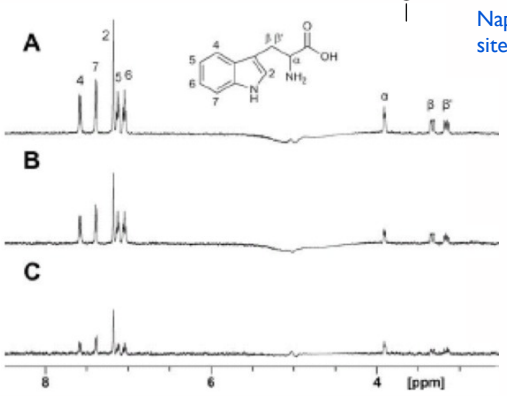

Maestre-Reyna *et al.* PLoS One. 2013; 8(12): e82458.



# Specific vs. non-specific binding ligands




Naproxen (Nap) binds to multiple sites of HSA with sub-uM Kd




0.1mM Human serum albumin + 6mM L-trptophan
(all binding events)

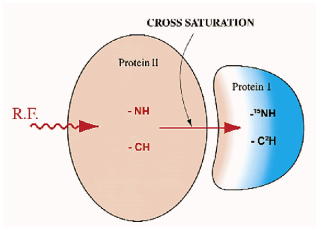
0.1mM human serum albumin + 6mM L-trptophan & 0.1 mM Nap
(only non-specific binding)

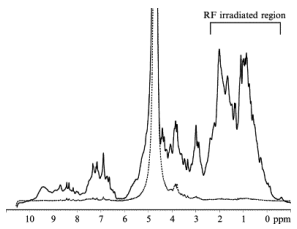
Spectrum A - spectrum B
(only specific binding)

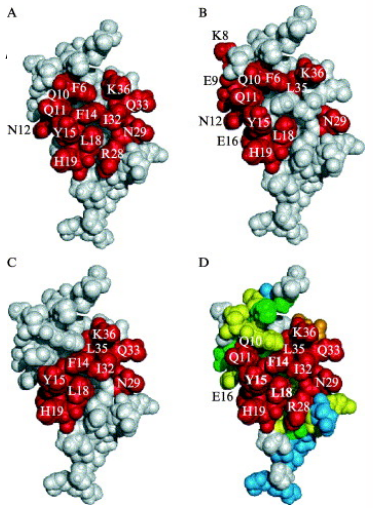
Mayer & Meyer *J.Am. Chem. Soc.*, **2001**, *123* (25), pp 6108–6117



# Transferred cross saturation (TCS) for large biomolecular complexes



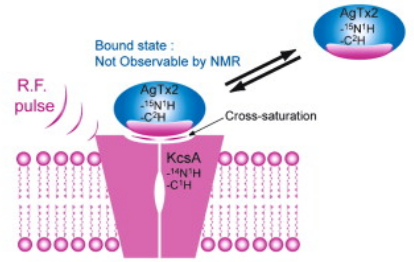



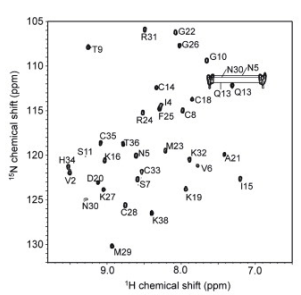



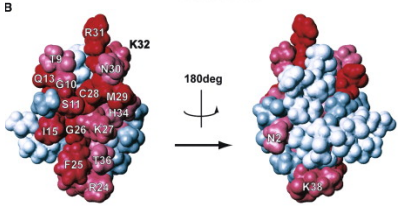

Shimada, Methods in Enzymology (2005) 394, 483–506

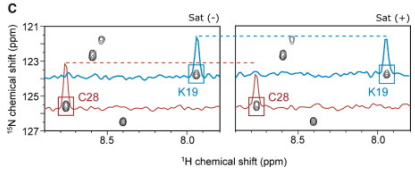


# Free state : Observable by NMR




Takeuchi et al., (2003) Structure, 11:1381

*Journal of Biomolecular NMR*, **21:** 349–359, 2001.
*KLUWER/ESCOM*


**WaterLOGSY as a method for primary NMR screening: Practical aspects and range of applicability**

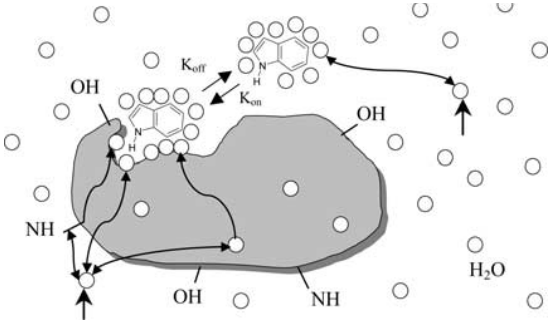
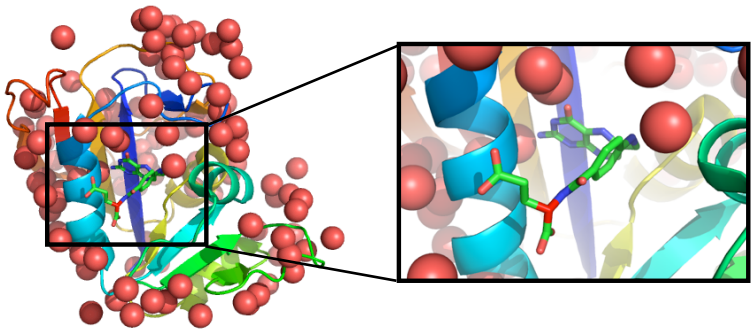



*Figure 1.* WaterLOGSY principle. The protein is shown with the buried cavities and the active binding site. The ligand is shown in the bound and free states. Excitation of bulk water (circles) is shown with a solid arrow and some of the different magnetization transfer pathways are shown with dashed lines.

Dalvit et al., *J. Biomol. NMR*, **2001**, *21*, pp 349-459



# Structured water



DFHR in complex with NADPH



# Cross relaxation rate between water and protein-ligand complex

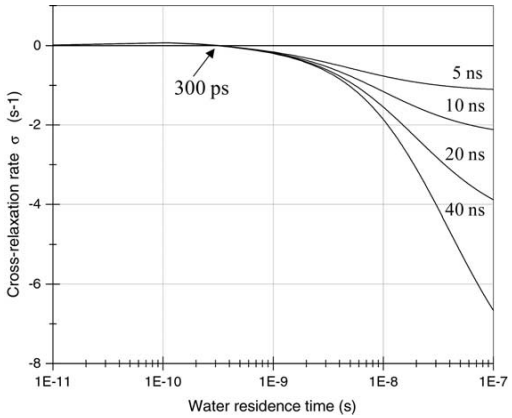



*Figure 4.* Intermolecular cross relaxation between water and protein-ligand complex according to Equation 1 as a function of water residence time (x axis). The simulation was performed for a Larmor frequency of 600 MHz, using a $r_{wp}$ of 2.5 Å. Simulations were performed for different rotational correlation times of the protein (values indicated with the curves).

Dalvit et al., *J. Biomol. NMR*, **2001**, *21*, pp 349-459



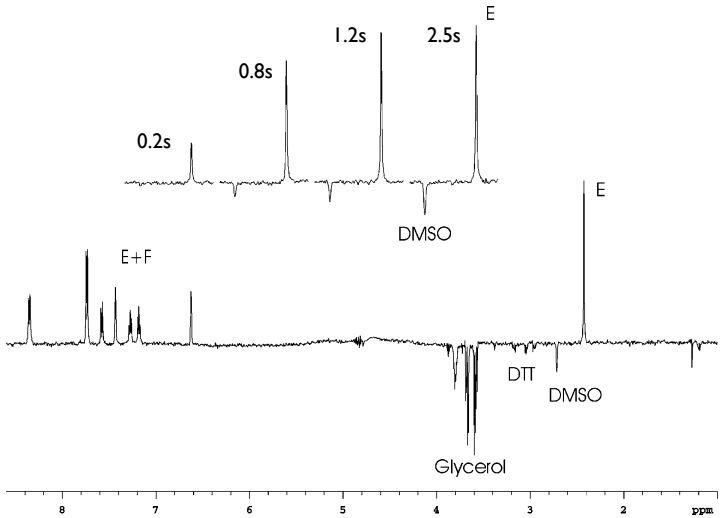



*Figure 6.* One-Dimensional WaterLOGSY for ligands E and F (200 μM) in the presence of 10 μM cdk2. The spectrum was recorded with 512 scans. The relaxation and mixing times were 2.6 and 2.0 s, respectively. Positive and negative signals identify cdk2 binding and non-interacting molecules, respectively. Glycerol was present at mM concentration. The insert shows expansions of the spectral region between 2.3 and 2.8 ppm which were measured by WaterLOGSY experiments with mixing times of, respectively, 0.2, 0.8, 1.2 and 2.5 s (from left to right).

Dalvit et al., *J. Biomol. NMR*, **2001**, *21*, pp 349-459

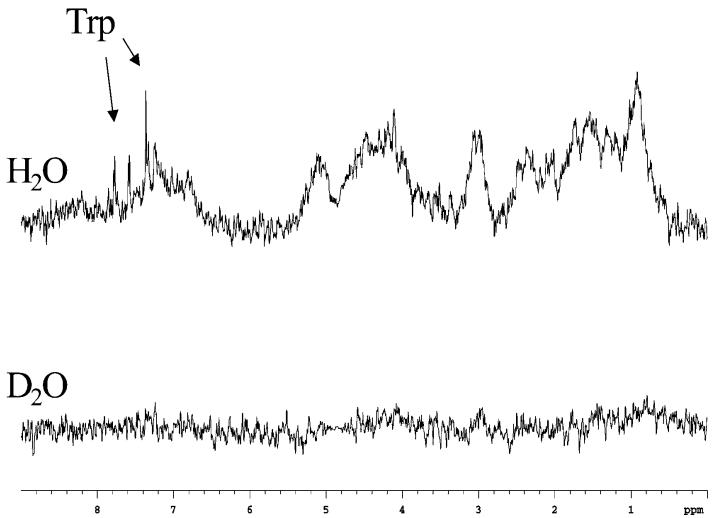
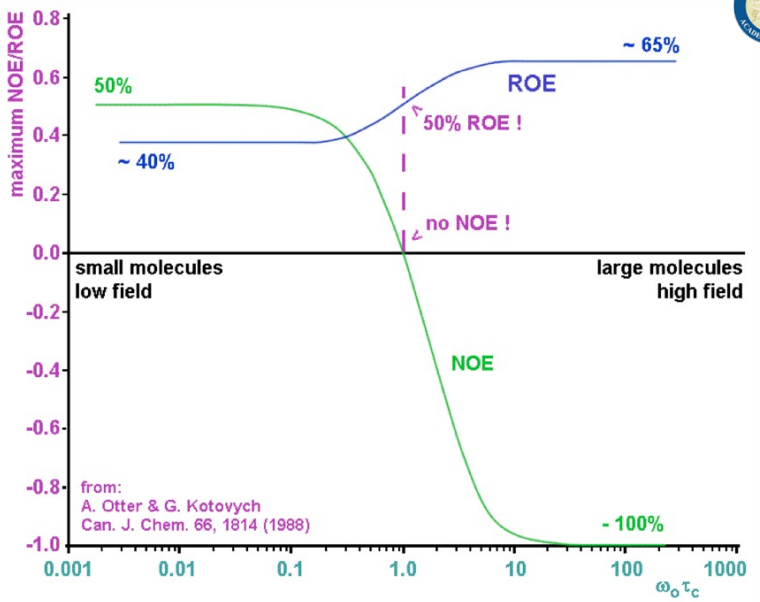



*Figure 7.* One-Dimensional WaterLOGSY spectra recorded for a 10 μM HSA solution in the presence of 100 μM L-Trp in $H_2O$ (upper trace) and $D_2O$ (lower trace). The spectra were recorded with 1024 scans and a 2.65 s repetition time. Other parameters are the same as described in Figure 6.

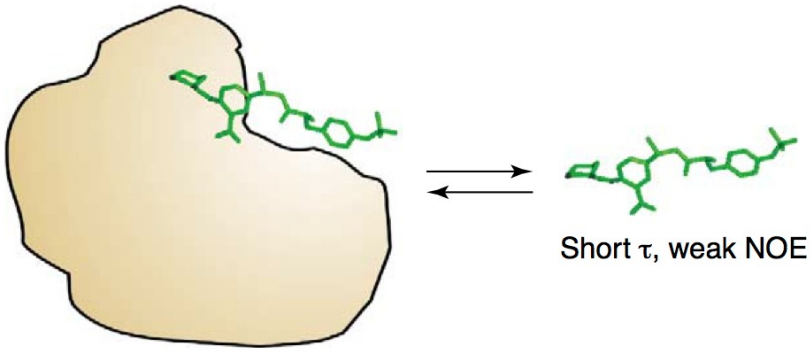
Dalvit et al., *J. Biomol. NMR*, **2001**, *21*, pp 349-459







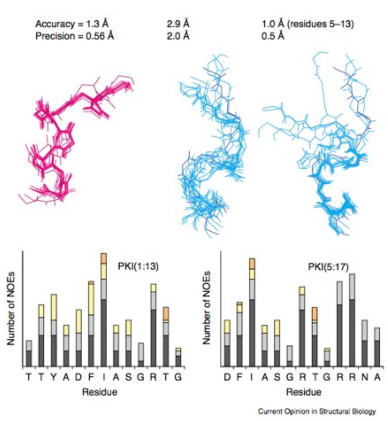

# Exchange transfer-NOE for determining ligand structure in bound state

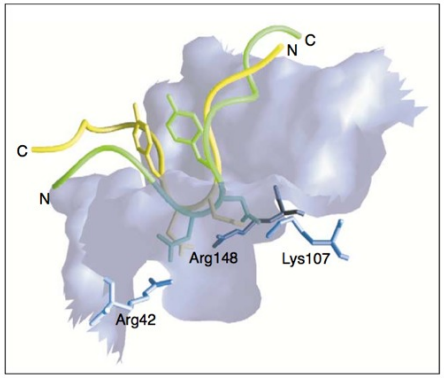



Current Opinion in Structural Biology 2003, 13:581–588



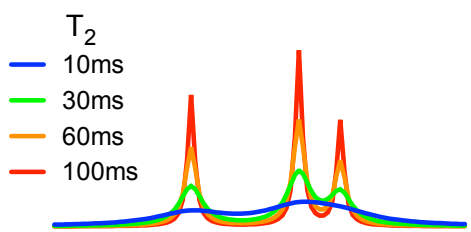
# trNOE provide bound-form structural information




Peptide docked by et-NOE restrained molecular dynamics. Two orientations that differ by an ~180° rotation are favorable (yellow and green). Reproduced with permission from [22].

# Effect of transverse relaxation CPMG




Transverse relaxation time ($T_2$) proportional to the correlation time of the sample, which is proportional to the rotational diffusion rate. According to Stoke-Einstein equation:

$$T_2 \propto \tau_c \propto D_R = \frac{k_B T}{8\pi\eta r_h^3}$$

Boltzmann constant; Temperature; Viscosity; Hydrodynamic radius



# Summary

- Advantages of NMR spectroscopy in drug development
  - Data collection under physiological conditions, i.e., in solution or even in cell
  - Atomic structural and dynamic information without crystallization
  - High information content which can be quantitated for SAR
  - Broad range of binding affinities can be explored (nM–mM)
- Requirements for NMR spectroscopy
  - High sample quantity and concentration (mM range) can cause solubility problem
  - Stable isotope ($^{13}C/^{15}N$) labelling can be expensive
  - High maintenance costs for NMR spectrometers

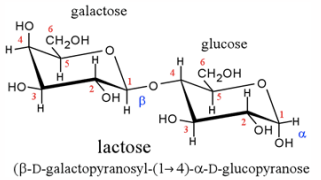


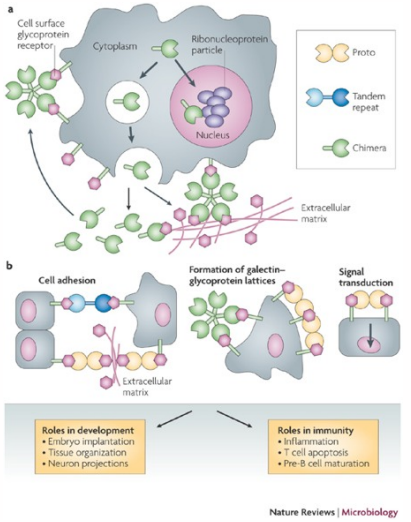
| | SAR-NMR | STD NMR | Spin labelling | Diffusion editing | Inverse NOE | Water-LOGSY |
|---|---|---|---|---|---|---|
| MW > 30kDa | limited | O | O | X | O | O |
| MW < 30kDa | O | X | O | O | X | X |
| Isotope labelling | O | X | X | X | X | X |
| Binding epitope on protein | O | X | X | X | X | X |
| Binding epitope on ligand | X | O | X | X | O | O |
| Protein quantity [nmol] | 25 | 0.1 | ~1 | ~100 | ~25 | ~25 |
| $K_D$ tight binding | no limit | 100pM | 100pM | ~100nM | 1nM | ~100pM |
| $K_D$ weak binding | ~1mM | ~10mM | ~10mM | ~1mM | ~1mM | ~10mM |
| ID of ligand | X | O | O | X | O | O |



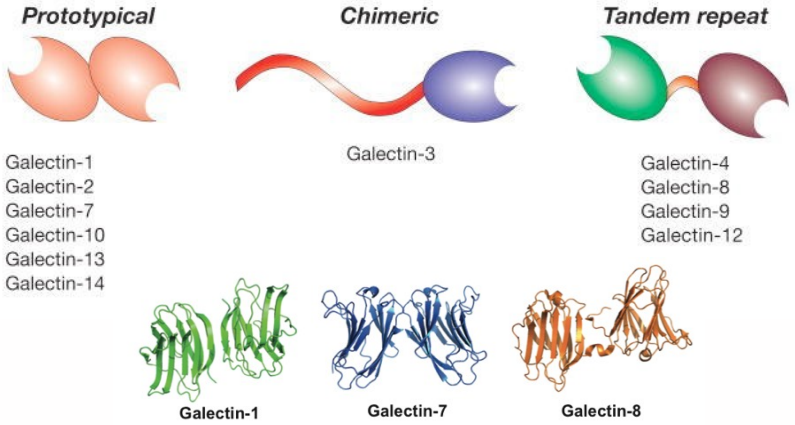
# Case study

- **Galectins** are a family of proteins defined by their binding specificity for β-galactoside sugars
- They have a broad variety of functions including mediation of cell–cell interactions, cell–matrix adhesion and transmembrane signalling.
- Their expression and secretion is well regulated, suggesting they may be expressed at different times during development.



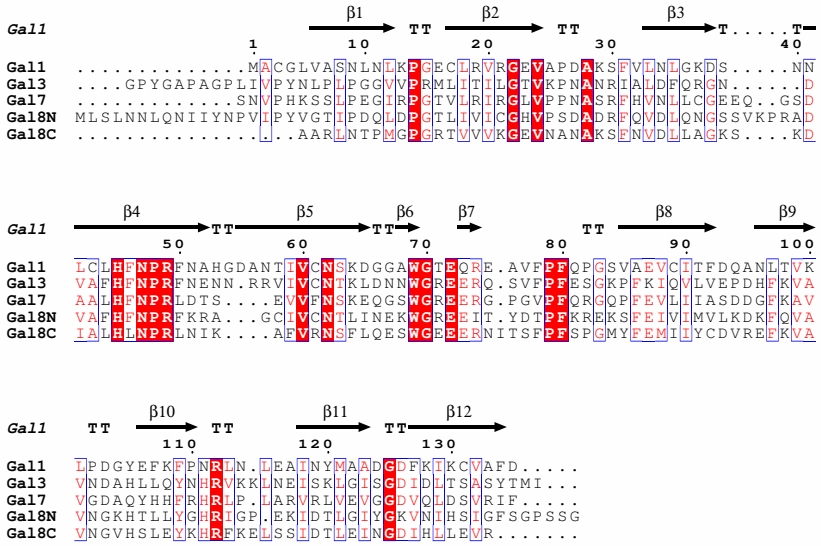
# Different types of galectins




Chapter 33, Galectins, Cummings & Liu "Essentials of Glycobiology, 2nd Ed." Varki A, Cummings RD, Esko JD, et al., editors. Cold Spring Harbor Laboratory Press; 2009

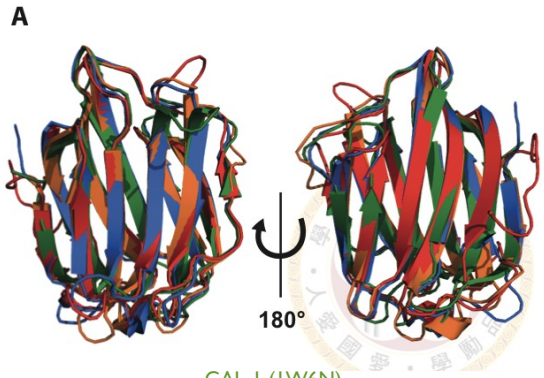


# Galectins have limited sequence homology

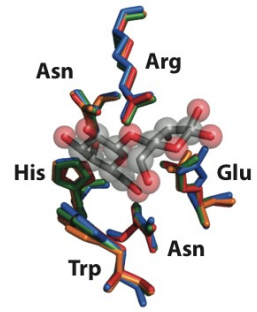





# Galectins use nearly identical sets of side-chains to recognise carbohydrate substrates



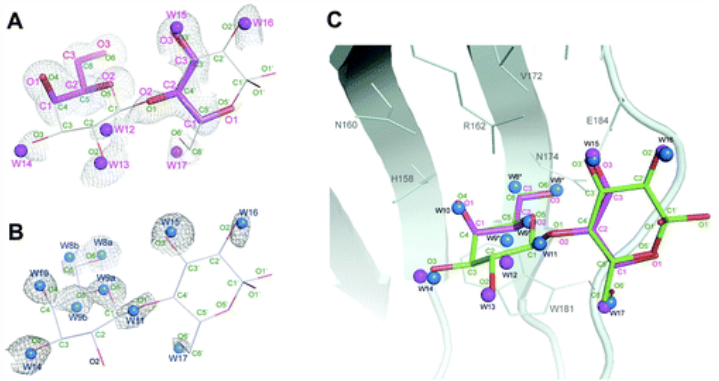

GAL-1 (1W6N)
GAL-7 (1BKZ)
GAL-8N (3KVN)
GAL-8C (3OBJ)




Tryptophan mediated CH-π interaction is universal in all galectin-glycan interactions

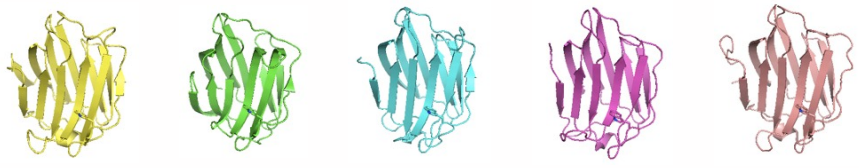


# Carbohydrate-binding site in galectin-3 is pre-organised to recognise a sugar-like framework of oxygens




Saraboji *et al., Biochemistry* **2012** 51:296

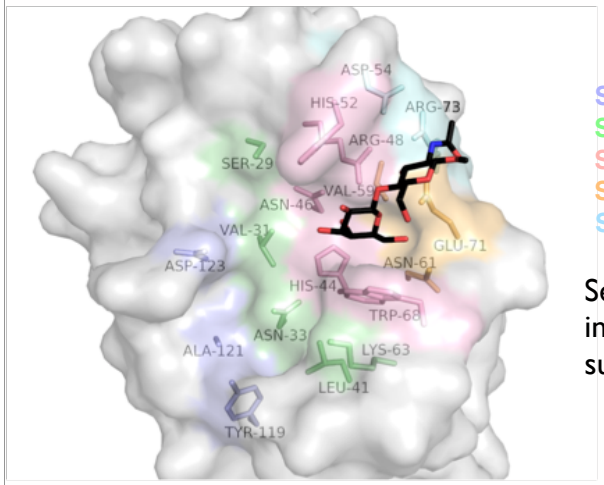
## Galectin substrate selectivity can be tuned by exploring the structural differences outside the direct glycan contact sites



Gal1 Gal3 Gal7 Gal8N Gal8C



## Galectins have relatively shallow binding pocket for carbohydrates

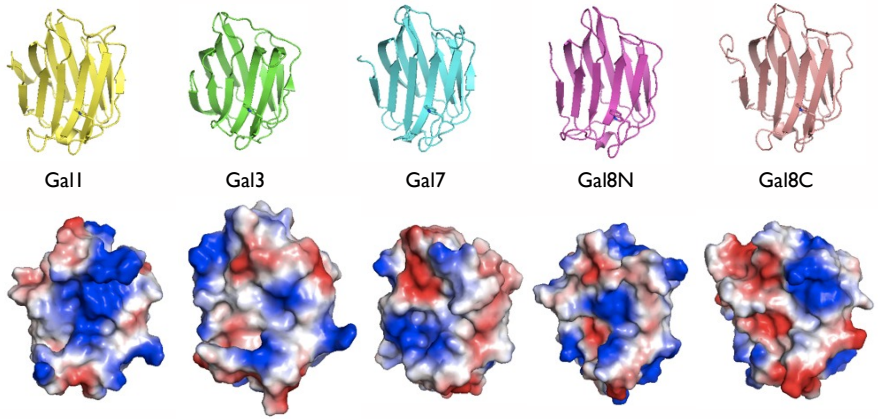



Subsite A
Subsite B
Subsite C
Subsite D
Subsite E

Second-shell interactions provides substrate specificity



## Galectin substrate selectivity can be tuned by exploring the structural differences outside the direct glycan contact sites

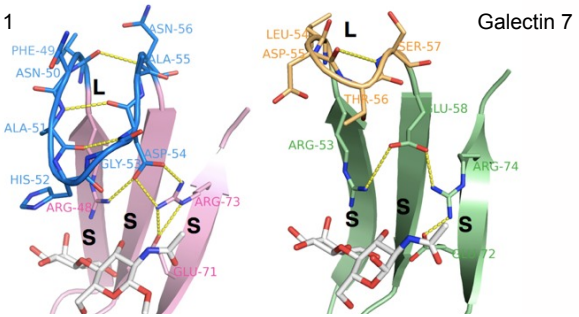


Gal1 Gal3 Gal7 Gal8N Gal8C



## Substrate specificity may be associated with loop structures containing an R-D/E-E-R salt-bridge network

Galectin 1 Galectin 7




```
hGal1                          44 HFNPRFNAHGDANTIVCNSKDGGAWGTEQR
hGal3_CRD                     158 HFNPRFNENNR-RVIVCNTKLDNNWGREER
hGal7 group consensus          49 HFNPR--LDtS--EVVGNskQGGsWGREER
hGal7-like group consensus     XX HFNPR--LdES--tVVgNtLpgGXWGrEER
Mammalian Gal2 consensus       XX HFNPR--FXES--tIvCNsXdGXSWGqEQR
                                  * *       * *       * * *
```

Hsieh *et al.* PLoS One. 2015; 10(5): e0125946.

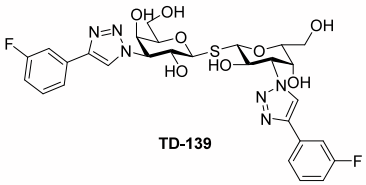
Galecto Biotech

Home Products About News


**FDA approves an IND from Galecto Biotech for TD139 treatment of IPF patients. - Phase Ib/IIa Trial to Start in the UK**

(COPENHAGEN – March 1, 2015) Galecto Biotech AB today announced that the U.S. Food and Drug Administration (FDA) has approved an Investigational New Drug (IND) application for the company's galectin-3 inhibitor, TD139, for the treatment of patients with idiopathic pulmonary fibrosis (IPF). Simultaneously, the company announced that a phase Ib/IIa trial in IPF patients is due to start shortly in several centers in the United Kingdom (UK).

Bristol-Myers Squibb

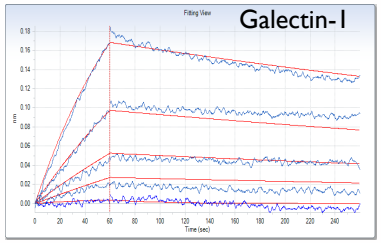


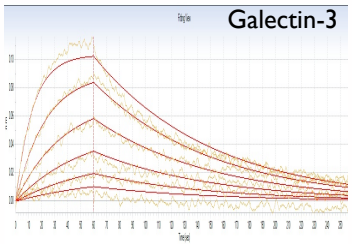
**TD-139**

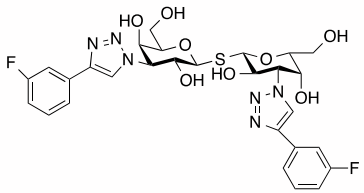
Leffler, Nilsson, Salameh, Sethi and Schambye



## TD139 binds to galectin-1 and -3 in nM range with very slow off rates



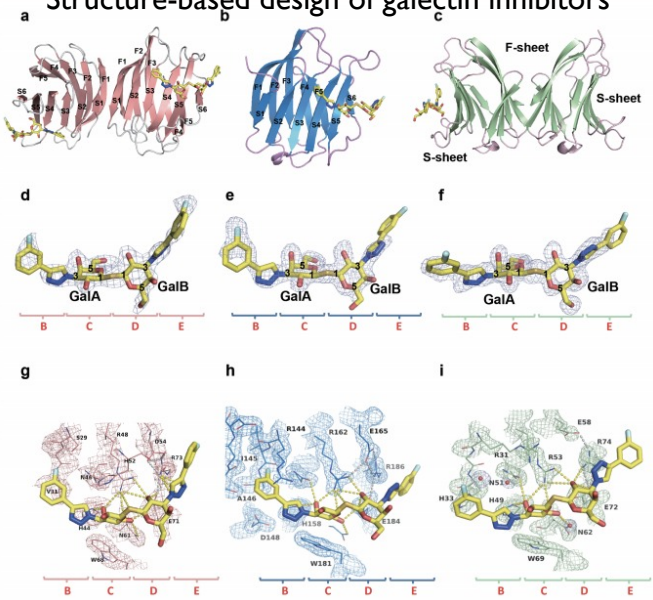



| | Galectin-1 | Galectin-3 |
|---|---|---|
| $K_D$ | 56 nM | 18 nM |
| $k_{on}$ ($M^{-1}$ $s^{-1}$) | 2E+05 | 6E+04 |
| $k_{off}$ ($s^{-1}$) | 1E-02 | 1E-03 |

Hsieh, Lin *et al., Sci. Rep.* (2016) *in press*

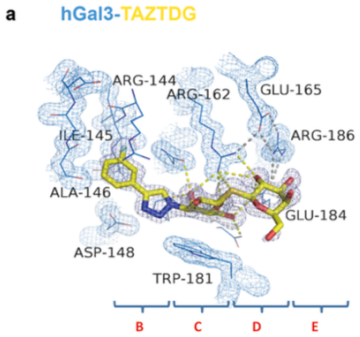


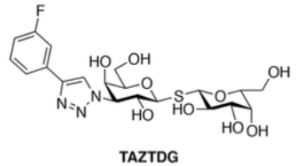
## Structure-based design of galectin inhibitors



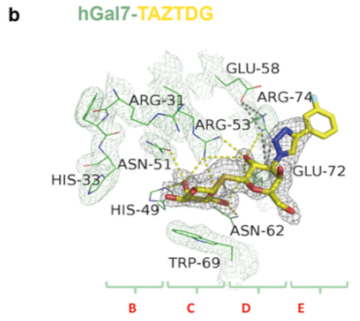

Hsieh, Lin *et al., Sci. Rep.* (2016) *in press*

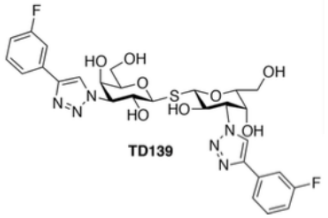




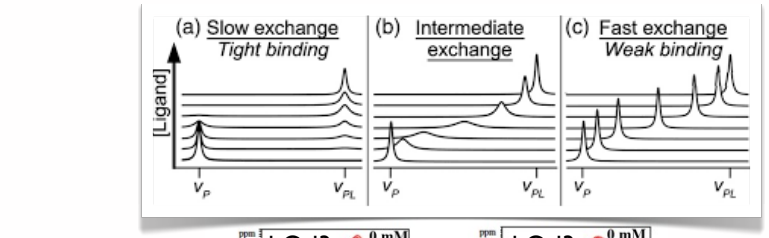
**TAZTDG**

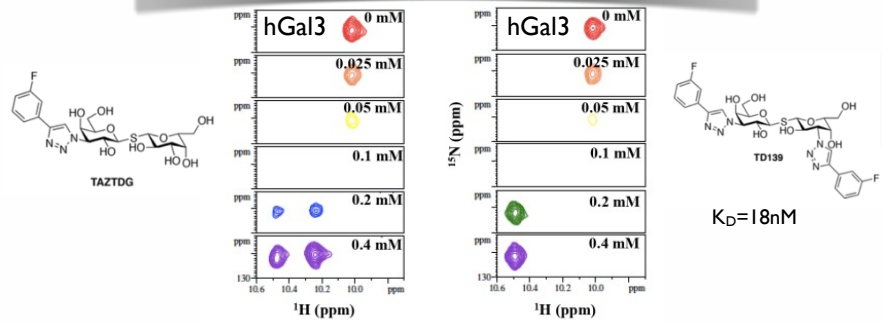


**TD139**

Hsieh, Lin *et al., Sci. Rep.* (2016) *in press*







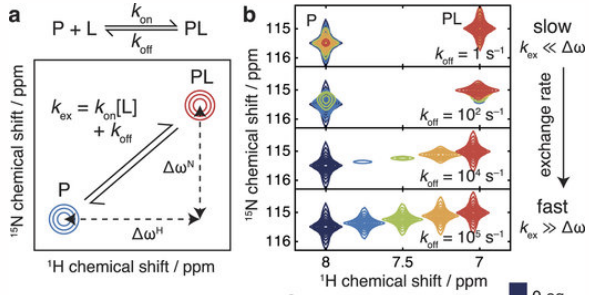


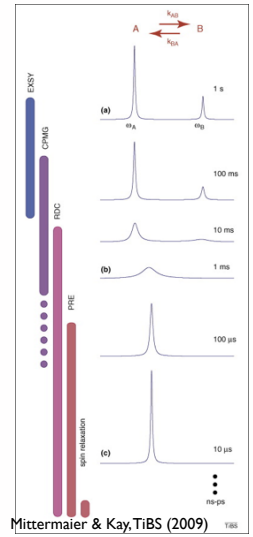

$K_D$=18nM

Hsieh, Lin *et al., Sci. Rep.* (2016) *in press*



# Two-dimensional NMR lineshape analysis

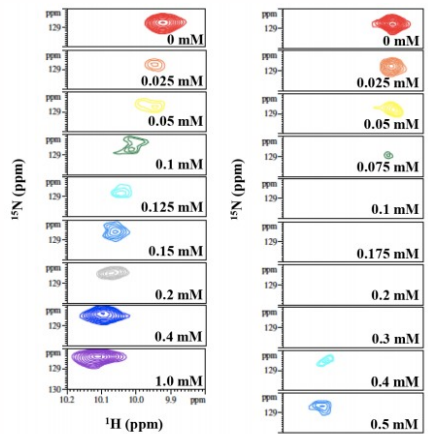



Mittermaier & Kay, TiBS (2009)

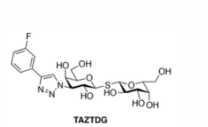
Waudby *et al., Scientific Reports* **6**, Article number: 24826 (2016)

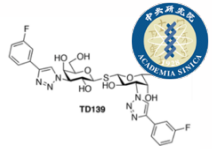


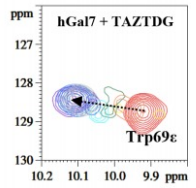


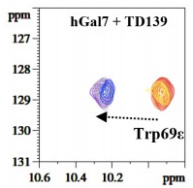


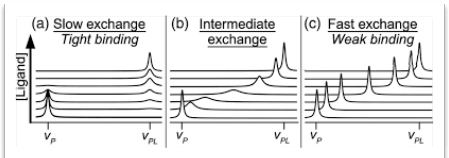




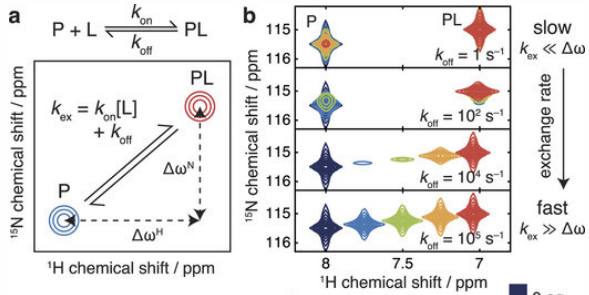




Hsieh, Lin *et al., Sci. Rep.* (2016) *in press*

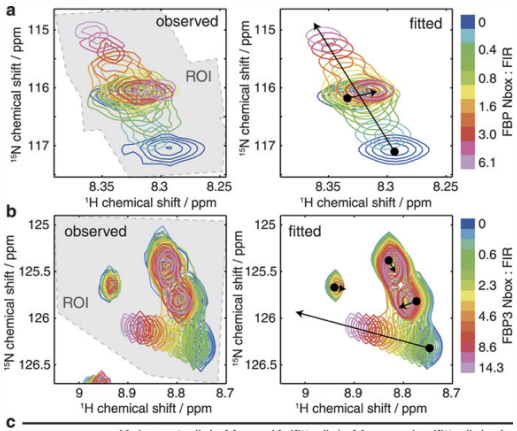


# Two-dimensional NMR lineshape analysis




Waudby *et al., Scientific Reports* **6**, Article number: 24826 (2016)

# Two-dimensional NMR lineshape analysis




c

| | $K_d$ (reported) / μM | $K_d$ (fitted) / μM | $k_{off}$ (fitted) / $s^{-1}$ |
|---|---|---|---|
| FBP Nbox | 14 ± 8 | 22.7 ± 0.5 | 2700 ± 100 |
| FBP3 Nbox | 280 ± 36 | 283 ± 2 | 15000 ± 300 |

Waudby *et al.*, *Scientific Reports* **6**, Article number: 24826 (2016)

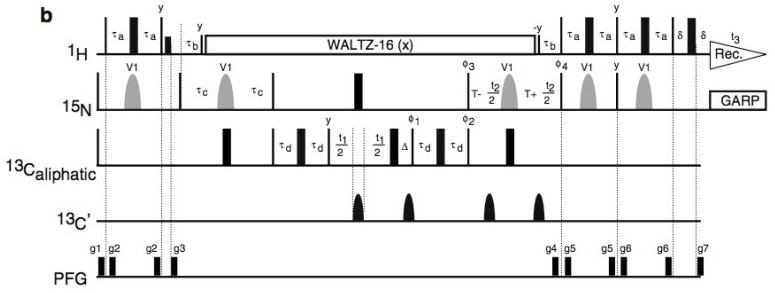


Journal of Biomolecular NMR (2006) 36:251–257
DOI 10.1007/s10858-006-9089-7

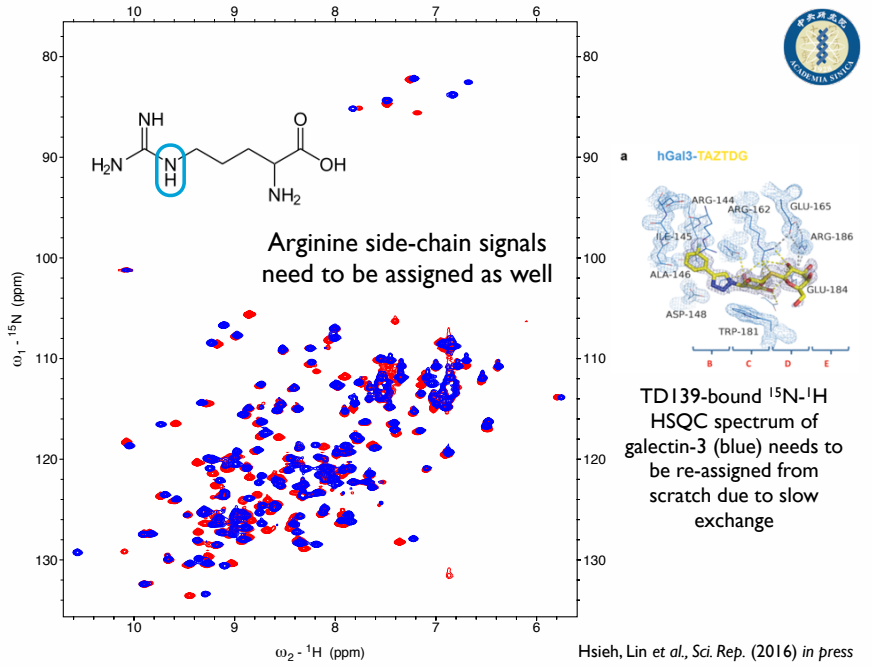

*Article*

## Sensitivity improvement for correlations involving arginine side-chain Nε/Hε resonances in multi-dimensional NMR experiments using broadband $^{15}$N 180° pulses

Junji Iwahara & G. Marius Clore*
*Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Disease, National Institutes of Health, Bethesda, MD 20892-0520, USA*





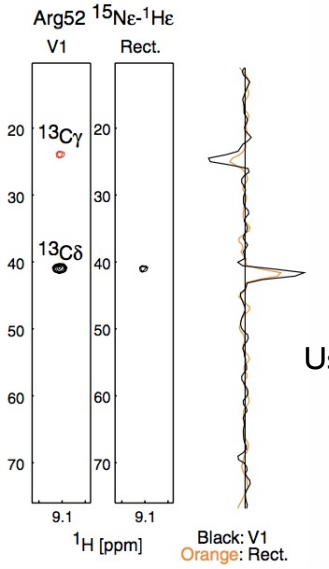


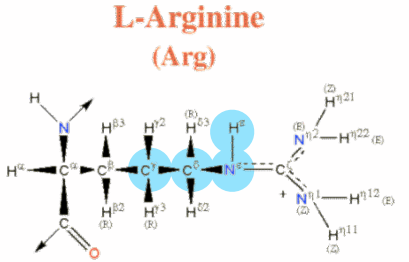


TD139-bound $^{15}$N-$^{1}$H HSQC spectrum of galectin-3 (blue) needs to be re-assigned from scratch due to slow exchange

Hsieh, Lin *et al.*, *Sci. Rep.* (2016) *in press*

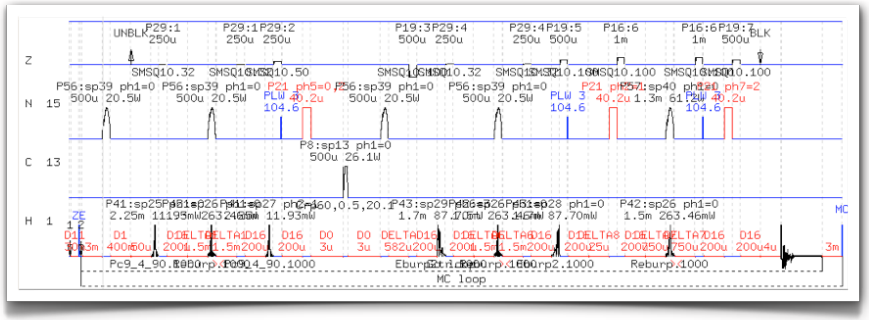





Use HNCACB to connect Nε-Cδ-Cγ

# BEST-TROSY pp:b_trosyetf3gpsi.3



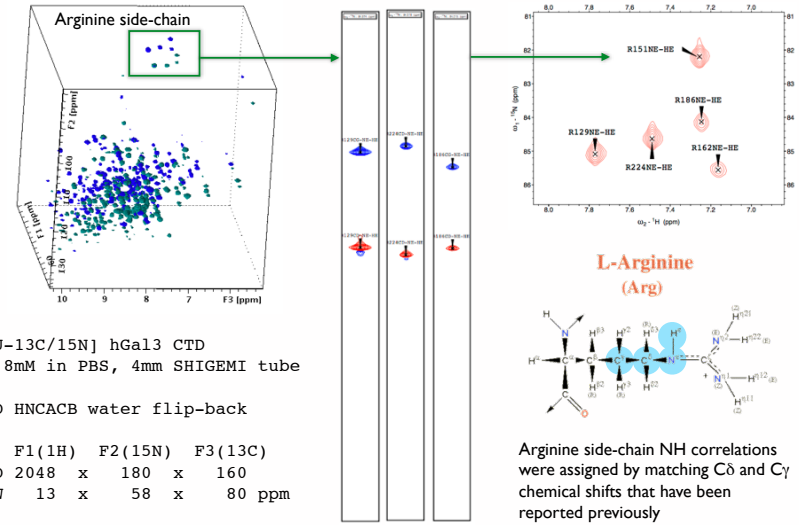
D1=0.4s SW(15N)=56ppm O3P=80ppm

Refocussing shapes
p57:sp40 - Reburps (1.3ms@-17.78dB/AVIII600)

Inversions pulses
p56:sp39 - Bip720,50,20.1 (500us@-13.12dB/AVIII600)

Bernhard Brutscher@IBS






[U-13C/15N] hGal3 CTD
1.8mM in PBS, 4mm SHIGEMI tube

3D HNCACB water flip-back

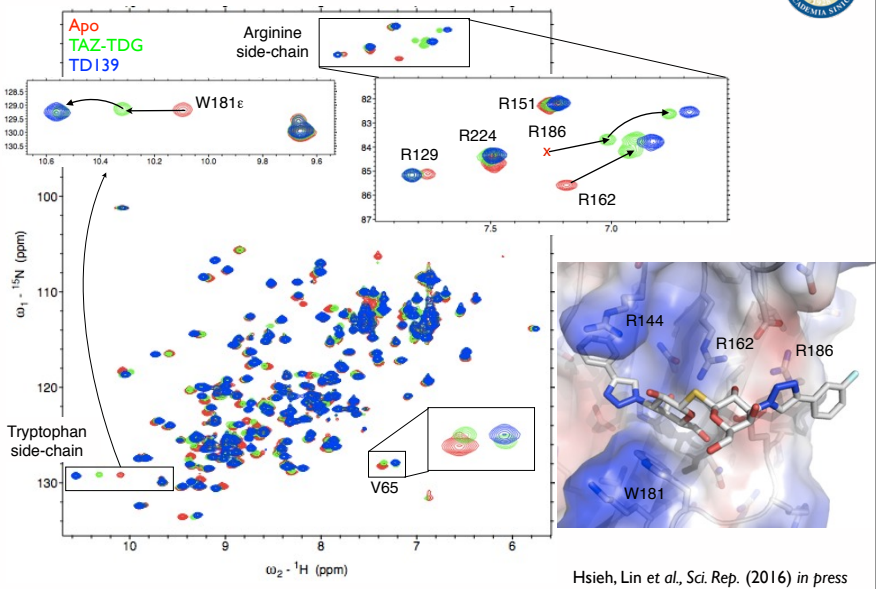
| | F1(1H) | F2(15N) | F3(13C) |
|---|---|---|---|
| TD | 2048 | 180 | 160 |
| SW | 13 | 58 | 80 ppm |

NS=48 **NUS 10% (in 46 hours)**

Arginine side-chain NH correlations were assigned by matching Cδ and Cγ chemical shifts that have been reported previously



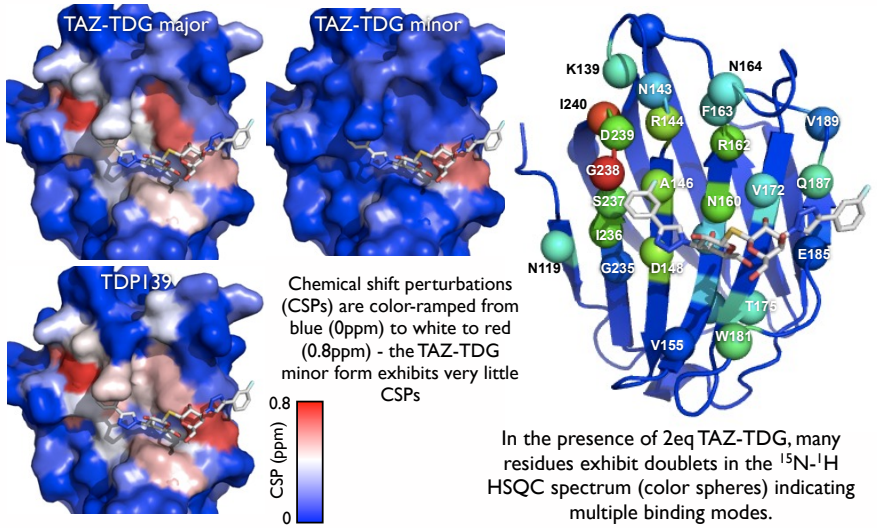
# NMR analyses of TAZ-TDG/TD139 binding to human galectin-3




Hsieh, Lin *et al., Sci. Rep.* (2016) *in press*



# Chemical shift perturbation mapping of galectin-3




Chemical shift perturbations (CSPs) are color-ramped from blue (0ppm) to white to red (0.8ppm) - the TAZ-TDG minor form exhibits very little CSPs

spectrum b, blue_white_red, minimum=0, maximum=0.8

In the presence of 2eq TAZ-TDG, many residues exhibit doublets in the $^{15}$N-$^{1}$H HSQC spectrum (color spheres) indicating multiple binding modes.

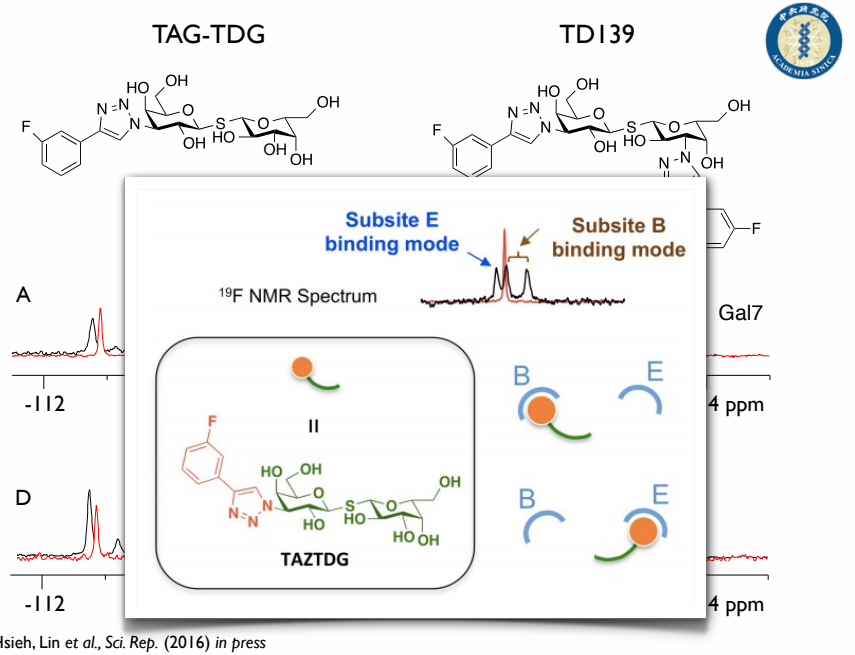
Hsieh, Lin *et al., Sci. Rep.* (2016) *in press*

# Properties of selected nuclei

| Nucleus | I | Υ (T · s)$^{-1}$ | Υ (T · s)$^{-1}$ | Natural abundance (%) |
|---|---|---|---|---|
| $^{1}$H | 1/2 | 3E+08 | 1 | 100 |
| $^{2}$H | 1 | 4E+07 | 0.15 | 0.02 |
| $^{13}$C | 1/2 | 7E+07 | 0.25 | 1.11 |
| $^{14}$N | 1 | 2E+07 | | 100 |
| $^{15}$N | 1/2 | -3E+07 | 0.10 | 0.36 |
| $^{17}$O | 3/2 | -4E+07 | | 0 |
| $^{19}$F | 1/2 | 3E+08 | 0.94 | 100 |
| $^{23}$Na | 3/2 | 7E+07 | | 100 |
| $^{31}$P | 1/2 | 1E+08 | 0.42 | 100.00 |
| $^{113}$Cd | 1/2 | 6E+07 | | 12.26 |

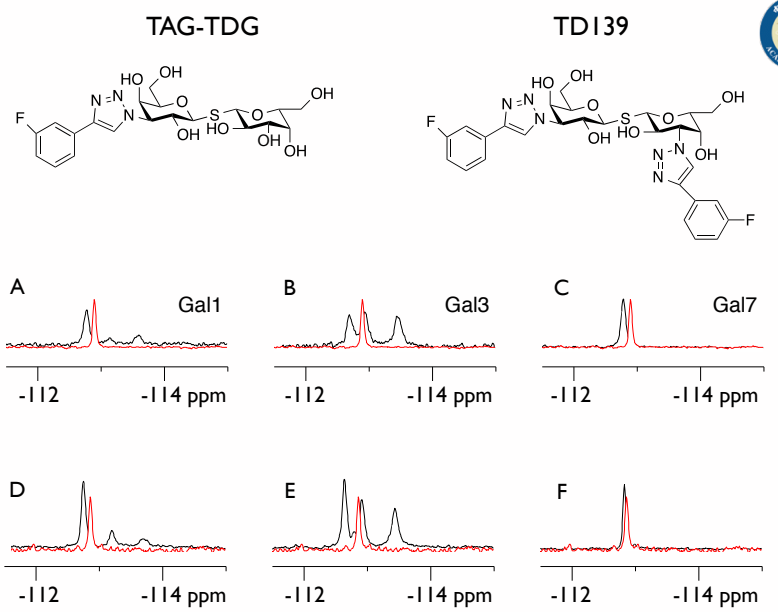
Taken from Cavanagh *et al.* "Protein NMR spectroscopy", 2nd edition, Wiley






Hsieh, Lin *et al., Sci. Rep.* (2016) *in press*

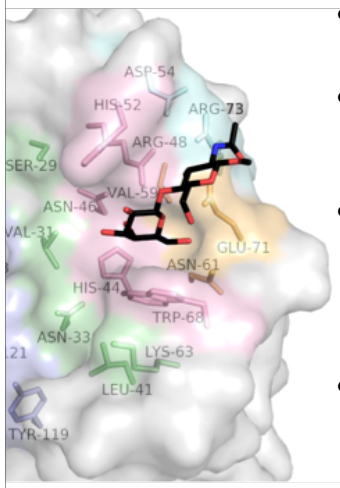





Hsieh, Lin *et al., Sci. Rep.* (2016) *in press*



# Summary




- Loop compositions of different galectins provide substrate specificity
- Arginines, in particular, can be targeted for obtaining inhibitors with higher affinity and selectivity
- $^{15}$N-$^{1}$H correlations of arginine side-chain guanidine groups serve as an ideal structural probe for NMR studies on galectin-substrate interactions around the canonical carbohydrate binding site(s)
- $^{19}$F NMR revealed a due binding mode for TD139 and TAZ-TDG to bind to galectin-1, -3 and -7 over a large dynamic range

# Automated NMR fragment screening at Monash Institute of Pharmaceutical Sciences


Customized liquid handler for sample preparation mixing protein, fragment, buffer, etc.


A sampleJet can hold 5 racks of 96 tubes

Bruker AVANCE III HD console affords optimal stability required for water suppression

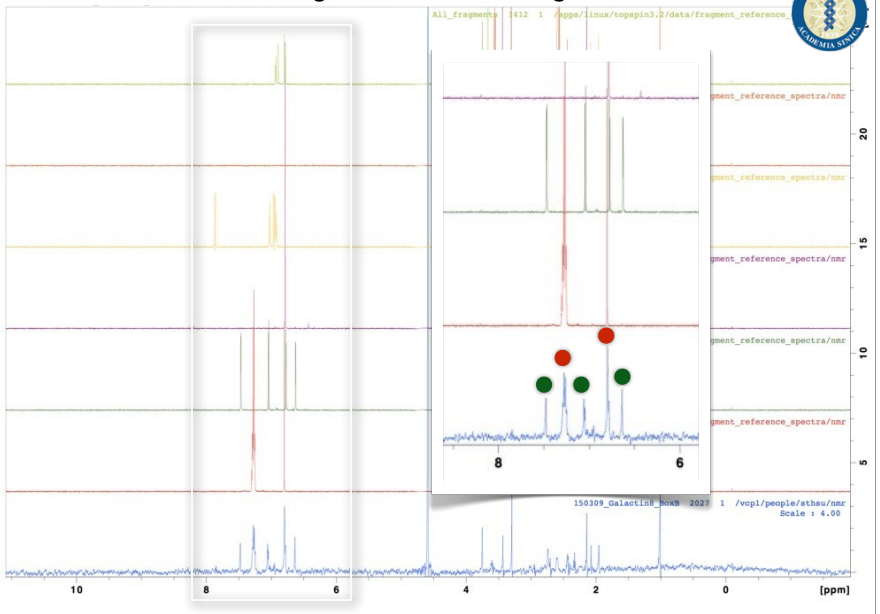


# Fragment screening strategy

- STD-NMR two racks of 96 well mixtures of 6 fragments
- STD/CPMG on singletons
- SPR on singletons with known competitors
- $^{15}N$-$^{1}H$ HSQC CSP mapping
- Crystallisation of fragments in complex with target

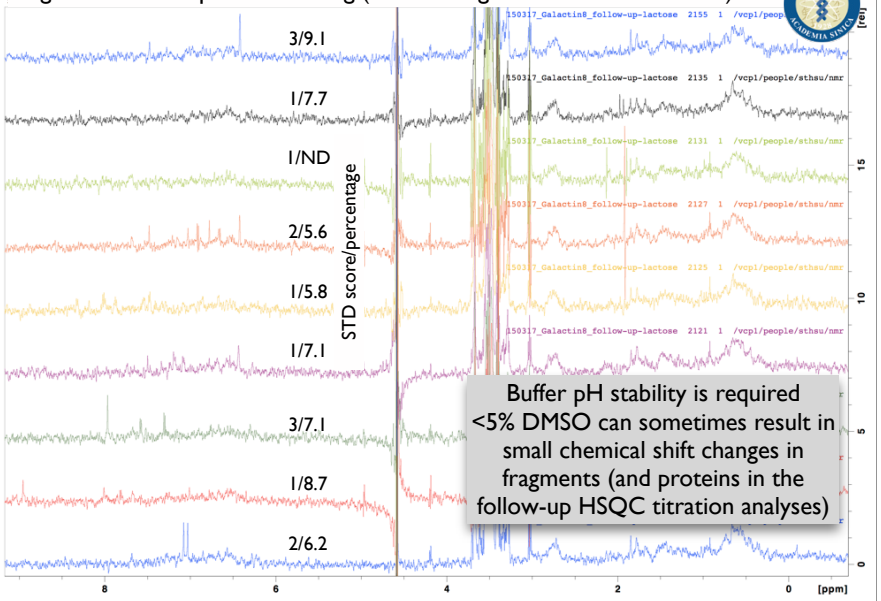


# Initial screen - aromatic regions to search for significant STD



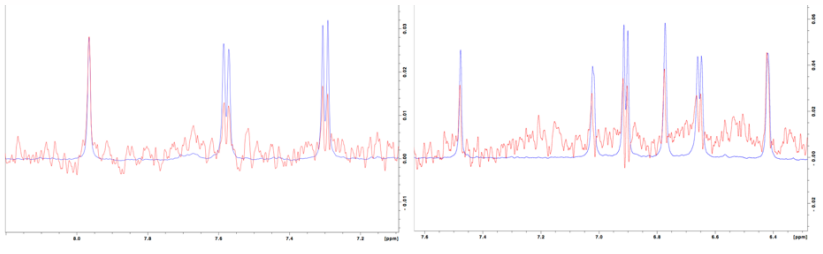



# Singletons for competition binding (0.2 mM fragment with 5mM lactose)




Buffer pH stability is required
<5% DMSO can sometimes result in small chemical shift changes in fragments (and proteins in the follow-up HSQC titration analyses)

## Second screen on singletons - STD percentage > 5%

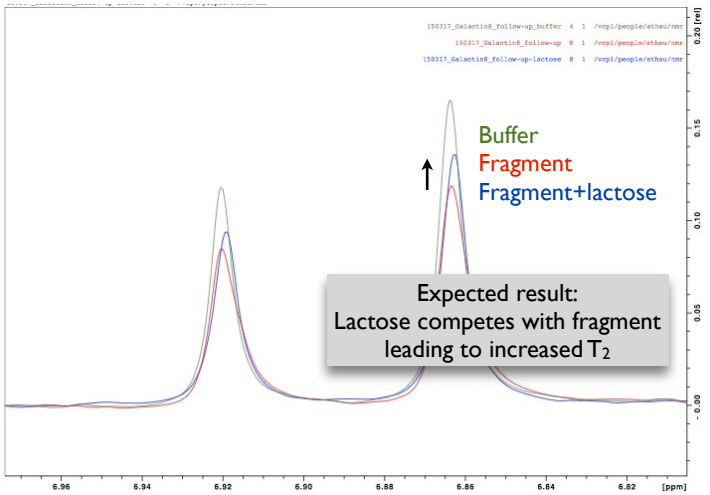


STD:7.1% / score: 3

STD:5.6% / score: 2

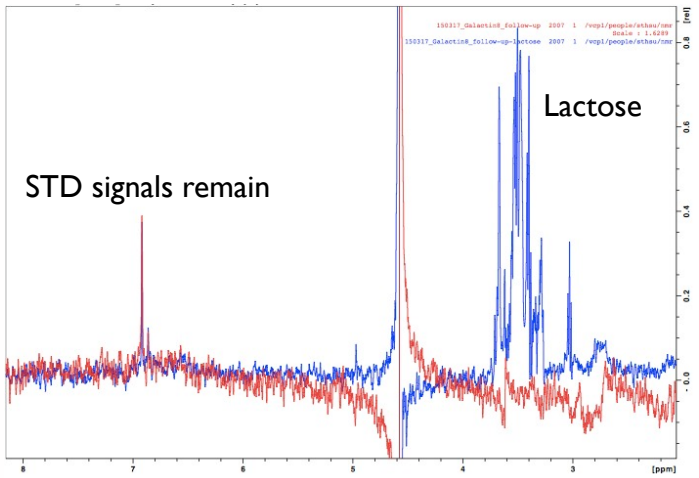


## CPMG without and with competitor



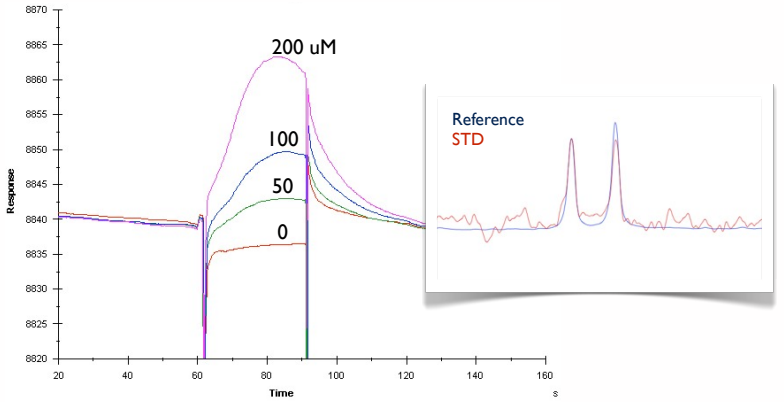



## Second screen on singletons with competitor STD percentage > 5%

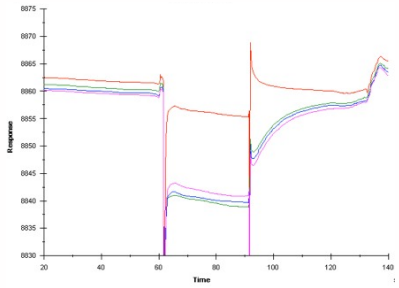


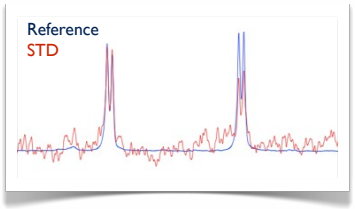


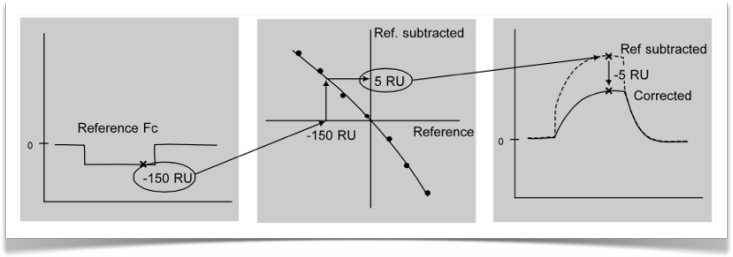

## SPR validation



## Poor solvent correction in DMSO-containing systems



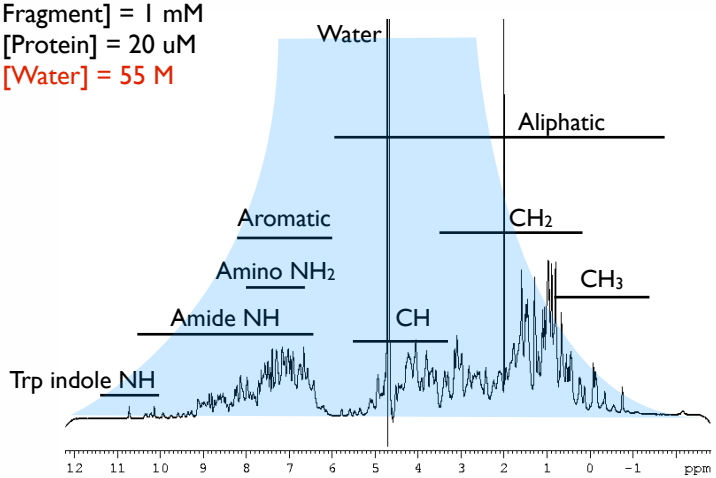





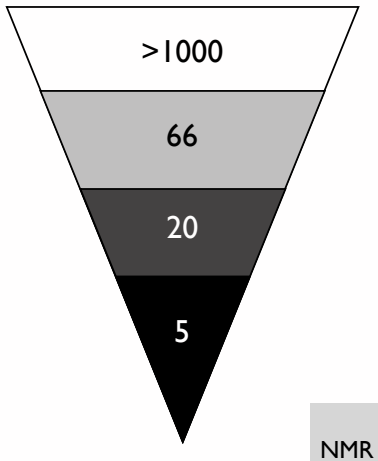

## Stable water suppression is essential

Fragment] = 1 mM
[Protein] = 20 uM
[Water] = 55 M






## Galectin inhibitor screening




STD-NMR two racks of 96 well mixtures of 6 fragments

STD/CPMG on singletons

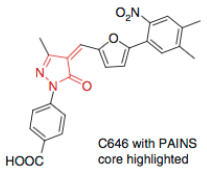
STD/CPMG on singletons with 5mM lactose as competitor and SPR

$^{15}N$-$^{1}H$ HSQC CSP mapping

Selection criteria:
NMR STD percentage > 5% and STD score >= 2



## PAINS
## Pan Assay INterference compoundS




C646 with PAINS core highlighted

37

**Table 1. Apparent activity (half maximal inhibitory concentration, $IC_{50}$) of 37 (C646) against a standard internal panel of enzymes and a control for signal interference (TruHits)**

| Enzyme[A] | $IC_{50}$ of 37 [μM] |
|---|---|
| TruHits (control) | 15 |
| FAK | 4.8 |
| Flt4 | 1.5 |
| KDR | 2.5 |
| cFMS | 0.8 |
| RET | 2.8 |
| SRC | 0.77 |

[A]See Experimental section for details.

Baell *et al.*, *Australian Journal of Chemistry* 66(12):1483 (2013)

# Acknowledgement


Iren Wang 王怡人

Manoj Kumar Sriramoju 庫瑪

Szu-Yu (Kate) Chen 陳思妤

Mei-Yi Chen 陳美意

Shih-Yun (Joanne) Chen 陳思勻

Yun-Tzai (Cloud) Li 李耘在

Chi-Jing Yang 楊子靖

**Institute of Biological Chemistry**

Chun-Hung (Hans) Lin 林俊宏

Tung-Ju Hsieh 謝東儒

Hsien-Ya Lin 林仙雅

Zhijay Tu 杜志傑

**Institute of Biomedical Sciences**

Fu-Tung Liu 劉扶東

**Monash Institute of Pharmaceutical Science - Melbourne / Australia**

Ray Norton

Martin Scanlon

Steve Heady

**Institut de Biologie Structurale - Grenoble / France**

Bernhard Brutscher

Adrien Favier




科技部

Ministry of Science and Technology