From Genes to Protein-Ligand Structures: a Work-flow Tuned for Structural Characterization by Solution NMR

ZoBio
Tools for fragment based drug discovery

Masakazu Kobayashi, Willem-Jan Waterrus, Johan Hollander, Dipen Shah, Eiso AB, Gregg Siegal

ZoBio BV, PO Box 9502, 2300RA, Leiden, The Netherlands. Contact: mkobayashi@zobio.com or gsiegal@zobio.com

Introduction

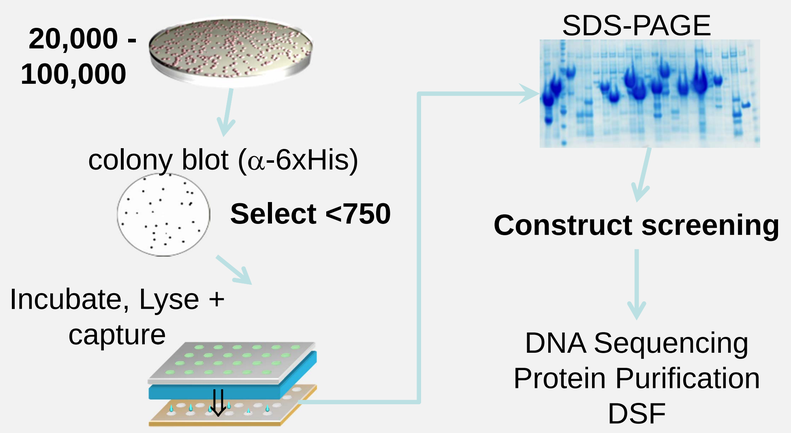
Structure-based drug design (SBDD) is dependent on the ready availability of 3D structure information, typically derived from X-ray crystallography. However, in many cases this is not successful. NMR is an alternative approach that can provide information on protein-ligand complexes at atomic resolution. In order to support NMR-driven SBDD, ZoBio has established a work flow that allows rapid assessment of soluble protein constructs and conditions that are suitable for executing structure elucidation of protein-ligand complexes by NMR. The typical steps taken in a "gene to protein-ligand structure" project at ZoBio are presented in this poster.

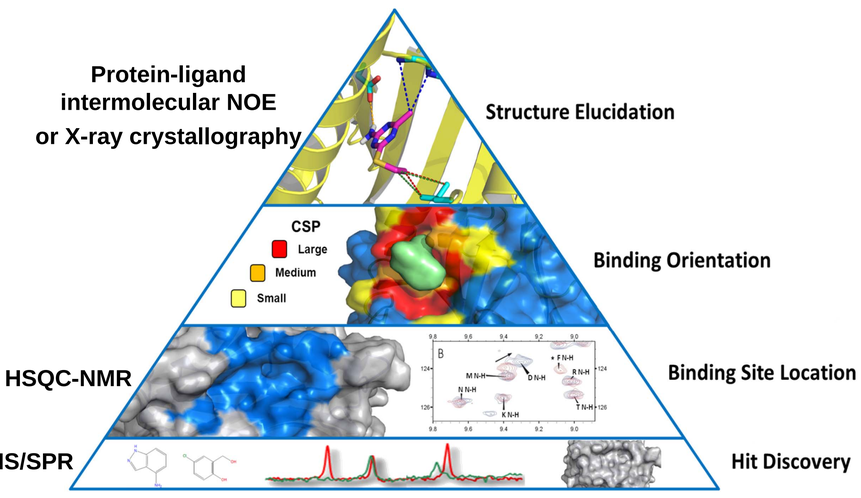
1. Identification of optimal protein constructs for structural biology by Combinatorial Domain Hunting (CDH) from Domainex (UK).

Often large multi-subunit proteins are difficult to express and purified in large quantity. Using a proprietary technique, CDH from Domainex, domains that are soluble, stable and functional are identified from a large set of truncation variants (20,000 – 100,000) that are expressed in *E.coli*.

domainex
the faster route to drug discovery

Gene Fragmentation with random truncations at both C- and N-termini → cloned into a *E. coli* expression vector

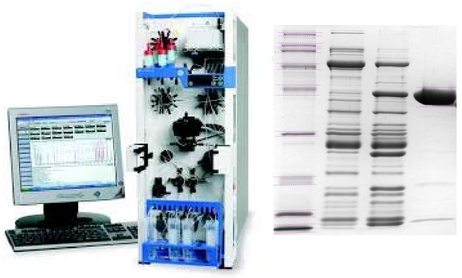




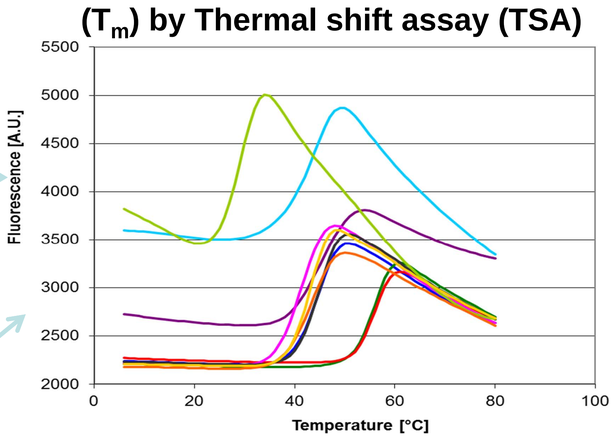
2. Identification of optimal buffer conditions for NMR-based structural biology.

NMR samples must remain as stable monomers (or oligomers) at high concentration (> 250 µ) for at least 1 week. NMR buffer conditions are optimized by iterative cycles of protein stability and oligomeric state assessment respectively by thermal shift assay (TSA) and by Size-Exclusion Chromatography with a Multi-Angle Laser Light Scattering (SEC-MALLS). The optimized conditions are also used to improve protein purification. **Example below.** 45 kDa protein formed a 450 kDa decamer in the Initial buffer condition and was therefore too large to be studied by NMR. Buffer scouting identified conditions in which the protein remained monomeric and displayed well dispersed amide resonances in a 2D [^{15}N,^{1}H]-TROSY spectrum.

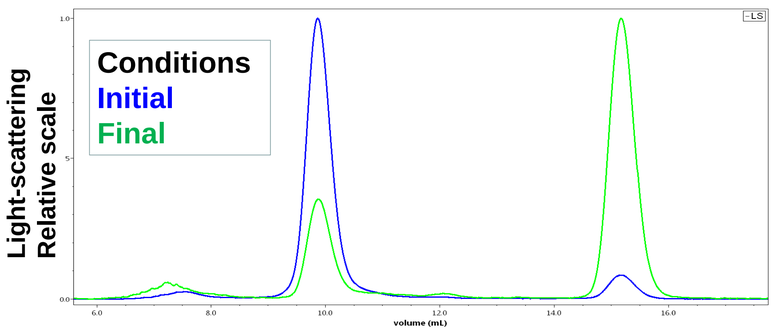
Optimization of purification



Assessment of melting temperature (T_m) by Thermal shift assay (TSA)



Oligomer (~450kDa) **Monomer (45 kDa)**

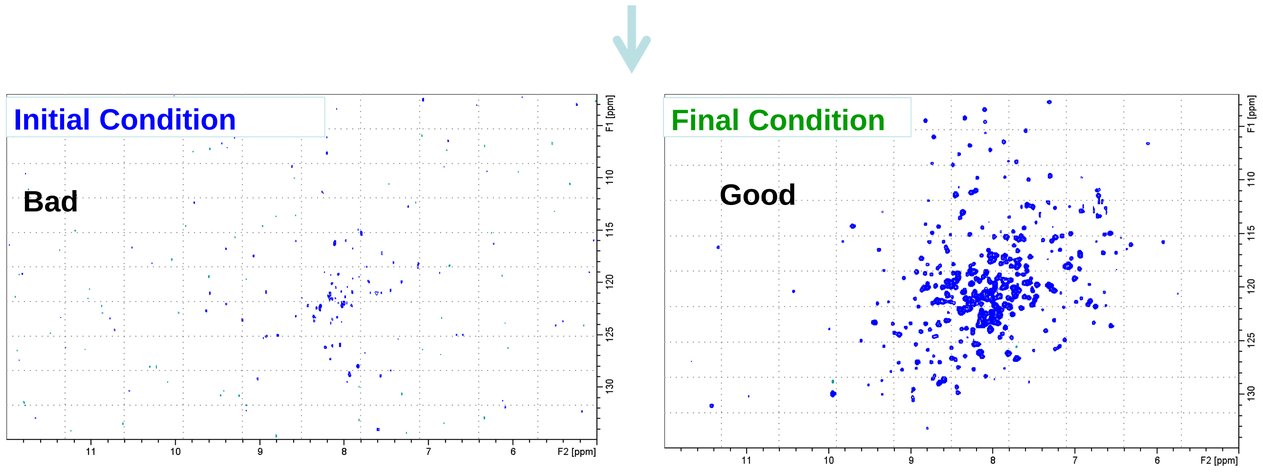


Volume (ml)

Condition with the highest T_m is optimal for protein stability. Investigated conditions include:

- Hampton slice pH (pH3.5 – 9.6)
- Phosphate buffer (pH6.5 – 8.5)
- Hampton Research Solubility & Stability screen kit (94 conditions)
- ZoBio's in–house additives (50 conditions)

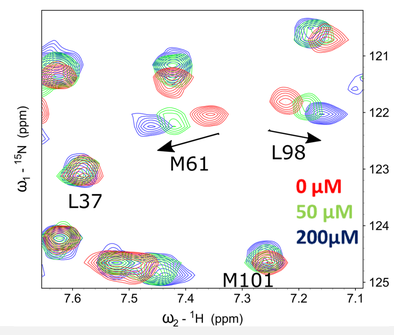
Assessment of oligomeric states by SEC-MALLS. Buffer conditions that favor monomeric (45 kDa) over oligomeric protein (450 kDa) were identified by SEC-MALLS. 90% of the protein in the **Final condition** is monomeric while in the **Initial condition** only 40% was. Note: Light scattering in Y-axis does not directly represent the fraction of the corresponding species.

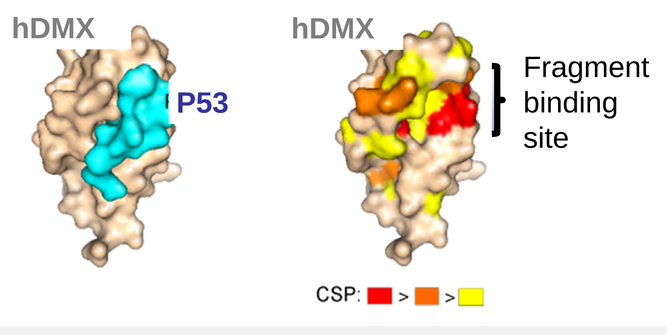


2D [^{15}N, ^{1}H]-TROSY of 270 µM uniformly ^{15}N-labeled 45 kDa protein. Left. In the initial buffer condition protein forms 450 kDa decamer and is too large to be detected by NMR. **Right.** In the optimized final buffer condition the protein remained monomeric (45 kDa) enabling structural study by NMR.

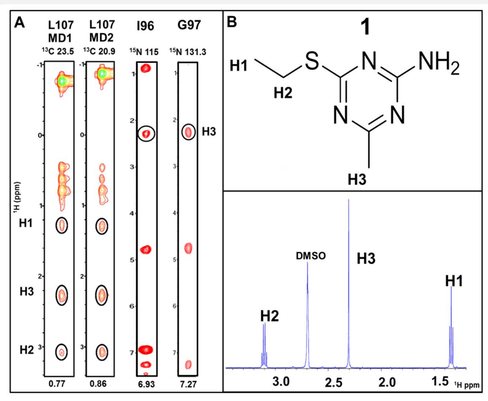
3. Structural characterization of ligand-protein complexes by NMR.

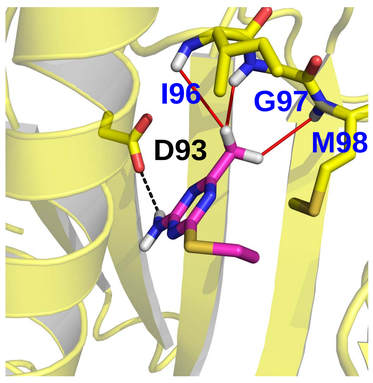
A. Binding of a ligand to a protein induces chemical shift perturbation (**CSP**) of NMR resonances of residues in the vicinity of the binding site. By mapping the residues experiencing a CSP, ligand binding sites can be readily identified. **B & C**. Specific isotope labeling of the methyl groups of e.g. ILV residues allows the measurement of intermolecular distances between the ligand and the interacting residues. The high resolution structure of a protein-ligand complex can then be calculated using the NMR derived distances.



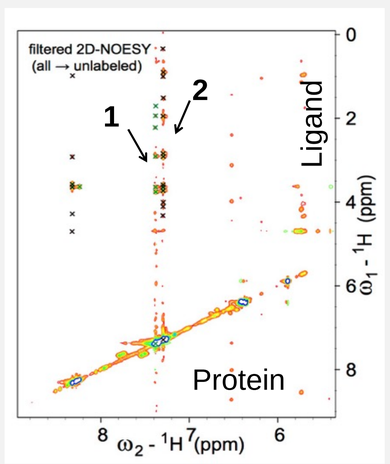


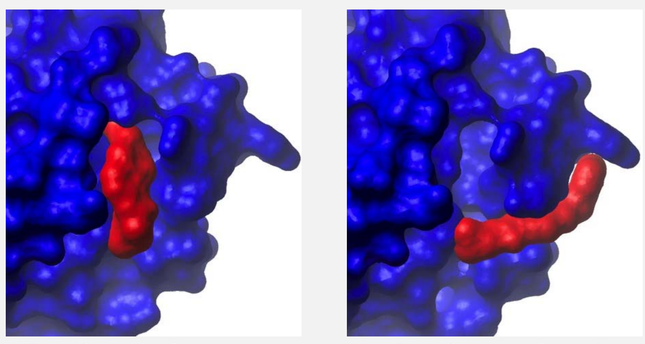
A. Rapid mapping of ligand binding sites. Left. [^{15}N,^{1}H]-HSQC of the p53 binding domain of hDMX in the presence of a ligand (K_D = 30 µM). The resonances experiencing CSPs are indicated by arrows. **Right.** Residues are colored based on the magnitude of the CSP: large (**red**), medium (**orange**) and small (**yellow**) and none (off-white).





B. NMR Structure of HSP90 NBD (28 kDa) complexed with a 170 Da fragment. Left. Intermolecular NOEs between methyl protons (L107) and H1, H3 and H2 and amide protons (I96 and G97) and H3 of 1 were observed in 3D ^{13}C- and ^{15}N- edited NOESY-HSQC spectra. **Right.** Fragment **1** was docked to the NBD using 21 NOE-based distant constraints [1].





C. NMR Structure of a protein (28 kDa) complexed with multiple conformations of a ligand. Filtered 2D-NOESY detected two distinct conformations of a ligand bound to the target. Conformation 1 was only detected by NMR while Conformation 2 was detected by both NMR and X-ray crystallography.

References

1. D. M. Shah et al., J. Med. Chem., 2012, 55 (23), pp 10786–10790