Differential Epitope Mapping by STD NMR Spectroscopy To Reveal the Nature of Protein–Ligand Contacts

Serena Monaco, Louise E. Tailford, Nathalie Juge, and Jesus Angulo*

Abstract: Saturation transfer difference (STD) NMR spectroscopy is extensively used to obtain epitope maps of ligands binding to protein receptors, thereby revealing structural details of the interaction, which is key to direct lead optimization efforts in drug discovery. However, it does not give information about the nature of the amino acids surrounding the ligand in the binding pocket. Herein, we report the development of the novel method differential epitope mapping by STD NMR (DEEP-STD NMR) for identifying the type of protein residues contacting the ligand. The method produces differential epitope maps through 1) differential frequency STD NMR and/or 2) differential solvent (D₂O/H₂O) STD NMR experiments. The two approaches provide different complementary information on the binding pocket. We demonstrate that DEEP-STD NMR can be used to readily obtain pharmacophore information on the protein. Furthermore, if the 3D structure of the protein is known, this information also helps in orienting the ligand in the binding pocket.

Saturation transfer difference (STD) NMR spectroscopy is a powerful NMR technique for ligand screening and gaining quantitative structural information from biologically relevant protein–ligand complexes.[1] The approach is appropriate for small-molecule binders of medium-weak affinity (high mM to low mM), there is no upper limit for protein size,[2] and labelling is not required. The technique is highly versatile and popular in the context of hit identification in drug discovery.[3]

STD NMR is based on the generation of saturation on a selected group of protein protons through selective “saturating” irradiation. For large proteins, spin diffusion spreads the saturation over the entire macromolecule, ultimately leading to intermolecular NOEs with protons of the ligand in the binding pocket. Mapping the STD NMR intensities on the ligand structure allows identification of “all” the ligand contacts with the protein in the complex (group epitope mapping).[4] However, it fails to provide information about the ligand location/orientation in the bound state. For this, competitive STD NMR experiments with “spy” ligands of known binding location must be carried out.[5]

Importantly, spin diffusion does not always ensure homogeneous saturation over the receptor, and differences in ligand epitope maps can be observed if STDs are acquired at different saturating frequencies.[5-6] This is prominent in the case of nucleic acids, where these differences are exploited to distinguish between major- and minor-groove binders.[6] For proteins, no methods have been described to exploit these differences to gain additional structural information. Although for proteins, these differences are typically small, they can be magnified and mapped into the ligand. Notably, ligand protons close to directly irradiated protein protons will show a relative increase in STDs compared to conditions where they are not directly irradiated (Figure 1a). Since the frequencies of irradiation can be chosen, we can select what types of protein protons will be “directly irradiated”, so that the differences will highlight parts of the ligand contacting those types of protein residues in the bound state (Figure 1a).

Another source of minor differences in epitope maps is the solvent. In D₂O, the polar side chains in the binding pocket have their exchangeable protons replaced by ²H, which is inefficient for transferring saturation.[7] In H₂O, these protons can contribute to an additional transfer of saturation compared to D₂O.[8] This process depends on their exchange rate with bulk water, with slowly exchanging polar protons being expected to produce the largest variations.[8] Therefore, ligand protons contacting these polar residues will show a relative increase in STDs compared to the binding epitopes in D₂O (Figure 1b).

In this work, we proposed to exploit these epitope differences to get information on the types of protein protons (aromatic, polar, or apolar) the ligand is contacting, thereby allowing assessment of the pharmacophore of the protein. To detect and quantify these differences, we have designed a simple method consisting of running pairs of STD NMR experiments under two experimental conditions (Figure 1) and quantifying the differences in relative STDs. Each pair consists of experiment-1 (exp1) and experiment-2 (exp2), performed under two conditions, that is, two different frequencies or two different solvents (D₂O or H₂O). The result provides a “Differential Epitope Map” of the ligand, which is a map of epitope differences from each pair of STD NMR experiments. We call the method differential epitope mapping STD NMR spectroscopy (DEEP-STD NMR).
and understanding the binding of 2,7-anhydro-Neu5Ac as:

\[ \exp^2 \exp^1 \]

exp factors, ideally close to 0. Differential Epitope Mapping (0.6/6.55 ppm) of 2,7-anhydro-Neu5Ac to RgNanH-GH33. For DEEP-STD NMR, the selection of frequencies to irradiate different types of protein protons can be based on the spectral properties of the protein (if chemical shifts are assigned), NMR databases (e.g., BMRB[12]), or predictions using a 3D model of the protein[13]. We ran two STD NMR experiments irradiating (0.5 s) at 0.60 ppm (exp1) and 6.55 ppm (exp2). These frequencies are known to be centered in the aliphatic and aromatic protein spectral regions, respectively.[12]

Several \( \Delta \text{STD} \) factors were observed (Figure 2a), confirming changes in the ligand binding epitope map under the two different irradiations. The \( \Delta \text{STD} \) factors (0.60/6.55 ppm) are shown in Figure 2a.

Positive \( \Delta \text{STD} \) factors report relative STD increases with irradiation at 0.60 ppm (aliphatic residues), whereas negative

\[ \Delta \text{STD}_{i} = \frac{\text{STD}_{\exp1}}{\text{STD}_{\exp2}} - \frac{1}{n} \sum_{j} \left( \frac{\text{STD}_{\exp1}}{\text{STD}_{\exp2}} \right) \]  

(1)

To obtain a consistent scale of \( \Delta \text{STD} \) factors, exp1 must be the experiment showing larger total ligand saturation. The term \((1/n) \sum_{j} (\text{STD}_{\exp1} / \text{STD}_{\exp2})\) accounts for intrinsic differences in saturation levels under the changes in experimental conditions. Ligand protons not affected by the changes in experimental conditions should show low \( \Delta \text{STD} \) factors, ideally close to 0. The differential epitope map of the ligand is obtained by mapping the \( \Delta \text{STD} \)s onto its structure. Notably, since \( \Delta \text{STD} \)s reflect contacts to specific types of amino acids, if the 3D structure of the protein is known, the method can also reveal the orientation of the ligand in the binding pocket.

As a proof of principle, we studied two biologically relevant protein–ligand complexes for which high-resolution X-ray structures are available: 1) 2,7-anhydro-Neu5Ac with RgNanH-GH33, the catalytic domain (belonging to glycoside family GH33) of the intramolecular trans-sialidase from the gut symbiont Ruminococcus gnavus,[19] and 2) 3-nitrophenyl-\( \alpha \)-galactopyranoside (3NPG) with Cholera toxin subunit B (CTB).[20] Understanding the binding of 2,7-anhydro-Neu5Ac by sialidases is important to unveil mechanisms of gut microbiota adaptation.[21,11] 3NPG is a well-known inhibitor of CTB and is frequently chosen as a reference for the design of novel inhibitors.

We first tested the method by analyzing the effect of different irradiation frequencies on the binding of 2,7-anhydro-Neu5Ac to RgNanH-GH33. For DEEP-STD NMR, the selection of frequencies to irradiate different types of protein protons can be based on the spectral properties of the protein (if chemical shifts are assigned), NMR databases (e.g., BMRB[12]), or predictions using a 3D model of the protein[13]. We ran two STD NMR experiments irradiating (0.5 s) at 0.60 ppm (exp1) and 6.55 ppm (exp2). These frequencies are known to be centered in the aliphatic and aromatic protein spectral regions, respectively.[12] Several \( \Delta \text{STD} \) factors were observed (Figure 2a), confirming changes in the ligand binding epitope map under the two different irradiations. The \( \Delta \text{STD} \) factors (0.60/6.55 ppm) are shown in Figure 2a.

Positive \( \Delta \text{STD} \) factors report relative STD increases with irradiation at 0.60 ppm (aliphatic residues), whereas negative

\[ \Delta \text{STD}_{i} = \frac{\text{STD}_{\exp1}}{\text{STD}_{\exp2}} - \frac{1}{n} \sum_{j} \left( \frac{\text{STD}_{\exp1}}{\text{STD}_{\exp2}} \right) \]  

(1)

To obtain a consistent scale of \( \Delta \text{STD} \) factors, exp1 must be the experiment showing larger total ligand saturation. The term \((1/n) \sum_{j} (\text{STD}_{\exp1} / \text{STD}_{\exp2})\) accounts for intrinsic differences in saturation levels under the changes in experimental conditions. Ligand protons not affected by the changes in experimental conditions should show low \( \Delta \text{STD} \) factors, ideally close to 0. The differential epitope map of the ligand is obtained by mapping the \( \Delta \text{STD} \)s onto its structure. Notably, since \( \Delta \text{STD} \)s reflect contacts to specific types of amino acids, if the 3D structure of the protein is known, the method can also reveal the orientation of the ligand in the binding pocket.

As a proof of principle, we studied two biologically relevant protein–ligand complexes for which high-resolution X-ray structures are available: 1) 2,7-anhydro-Neu5Ac with RgNanH-GH33, the catalytic domain (belonging to glycoside family GH33) of the intramolecular trans-sialidase from the gut symbiont Ruminococcus gnavus,[19] and 2) 3-nitrophenyl-\( \alpha \)-galactopyranoside (3NPG) with Cholera toxin subunit B (CTB).[20] Understanding the binding of 2,7-anhydro-Neu5Ac by sialidases is important to unveil mechanisms of gut microbiota adaptation.[21,11] 3NPG is a well-known inhibitor of CTB and is frequently chosen as a reference for the design of novel inhibitors.

We first tested the method by analyzing the effect of different irradiation frequencies on the binding of 2,7-anhydro-Neu5Ac to RgNanH-GH33. For DEEP-STD NMR, the selection of frequencies to irradiate different types of protein protons can be based on the spectral properties of the protein (if chemical shifts are assigned), NMR databases (e.g., BMRB[12]), or predictions using a 3D model of the protein[13]. We ran two STD NMR experiments irradiating (0.5 s) at 0.60 ppm (exp1) and 6.55 ppm (exp2). These frequencies are known to be centered in the aliphatic and aromatic protein spectral regions, respectively.[12] Several \( \Delta \text{STD} \) factors were observed (Figure 2a), confirming changes in the ligand binding epitope map under the two different irradiations. The \( \Delta \text{STD} \) factors (0.60/6.55 ppm) are shown in Figure 2a.

Positive \( \Delta \text{STD} \) factors report relative STD increases with irradiation at 0.60 ppm (aliphatic residues), whereas negative
Differential Epitope Mapping (DEEP-STD NMR) is a robust tool to get information on the nature of the amino acids surrounding the ligand in the binding site in order to assess the pharmacophore of the protein. This information is inaccessible by traditional STD NMR. Additionally, if the protein 3D structure, and the results indicated the E51 and Q56 protons are likely to lack slowly exchanging polar side chains (e.g., Arg). Remarkably, the ligand contains an aromatic moiety, which precludes protein irradiation in this spectral region. However, in DEEP-STD NMR, it is possible to select other groups of protein protons for irradiation, providing that they are in spectral regions devoid of ligand signals. For CTB, we targeted protein resonances at 2.25 ppm. We predicted the chemical shifts of the protons of CTB within 4 Å of the ligand in the X-ray structure, and the results indicated the E51 and Q56 protons as the ones likely to be directly irradiated (Figure 4c, and Table S1 in the Supporting Information).

Experiments conducted with differential frequencies (2.25/0.60 ppm) resulted in positive STD values for protons H4, H5, H6, and H6' on the galactose, thus indicating an increase in relative STDs when irradiating at 2.25 ppm (Figure 4a). In contrast, negligible STD factors were observed for H1, H2, H3, and the aromatic protons at the opposite end of the molecule. The differential epitope map of 3NPG (Figure 4b) was found to be in perfect agreement with the published crystal structure of the complex between 3NPG and CTB (Figure 4c),[10] in which the galactose ring area of H4 to H6 is surrounded by the side chains of E51 and Q56. In contrast, H1, H2, H3, and the aromatic carbon atoms point away from those side chains in the binding pocket (Figure 4c).

Finally, we conducted DEEP-STD NMR experiments with differential D₂O/H₂O conditions on the 3NPG-CTB complex. The STDs of 3NPG were negligible, and no differential epitope was obtained, thus indicating that changing solvent did not significantly affect the STD pattern (Figure S4). This is in agreement with the lack of slowly exchanging polar residues in the CTB binding pocket.[10] This suggests that when no differential epitope is obtained after a change from D₂O to H₂O, the protein binding pocket is likely to lack slowly exchanging polar side chains (e.g., Arg).[8]

In summary, DEEP-STD NMR is a robust tool to get information on the nature of the amino acids surrounding the ligand in the binding site in order to assess the pharmacophore of the protein. This information is inaccessible by traditional STD NMR. Additionally, if the protein 3D structure,a nd the results indicated the E51 and Q56 protons were added using Schrodinger software.[15] The slowly exchangeable protons in the binding pocket are depicted with green surfaces.
The authors declare no conflict of interest.

Keywords: epitope mapping · fragment-based drug design · NMR spectroscopy · pharmacophores · protein–ligand binding

How to cite: Angew. Chem. Int. Ed. 2017, 56, 15289–15293
Angew. Chem. 2017, 129, 15491–15495


Structure is known, the method allows information to be gained from STD NMR on the orientation of the ligand for the first time. In comparison to the SOS-NMR method proposed by Hajduk et al. to reveal ligand orientation,[6] DEEP-STD NMR does not need selective protein deuteration. We envisage that DEEP-STD NMR in combination with classical STD NMR could become a popular approach to characterize in depth the binding of weak ligands to protein receptors. This is of great relevance to accelerate fragment-based drug discovery efforts, an approach of increasing importance in the biopharmaceutical industry for the development of novel therapeutics.

Acknowledgements

This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) through a New Investigator grant awarded to J.A. (BB/P010660/1). S.M. acknowledges a postgraduate studentship from the School of Pharmacy of the University of East Anglia. N.J. and L.E.T. acknowledge funding by the BBSRC Institute Strategic Programme for Gut Health and Food Safety (BB/J004529/1).

Figure 4. Differential Epitope Mapping (2.25 ppm/0.6 ppm) of 3-nitrophenyl-a-d-galactopyranoside (3NPG) in complex with Cholera toxin subunit B (CTB). a) ΔSTD histogram: protons with positive ΔSTDs (above the limit of ±0.75) after irradiation at 2.25 ppm are shown in orange. b) DEEP-STD map of the ligand. Orange surfaces indicate ligand contacts with protein side chains directly irradiated at 2.25 ppm. The ligand polar protons have been omitted. c) Crystal structure of the complex (PDB ID: 1EEI).[15] Protons were added using Schrodinger software.[16] Protein protons directly irradiated at 2.25 ppm are depicted with orange surfaces.


Manuscript received: July 27, 2017
Revised manuscript received: September 8, 2017
Accepted manuscript online: October 4, 2017
Version of record online: October 23, 2017