



UNC CHARLOTTE

Structural Elucidation of Sleeping Beauty Transposon via NMR Spectroscopy

Zach D. Merino, Dr. Irina V. Nesmelova

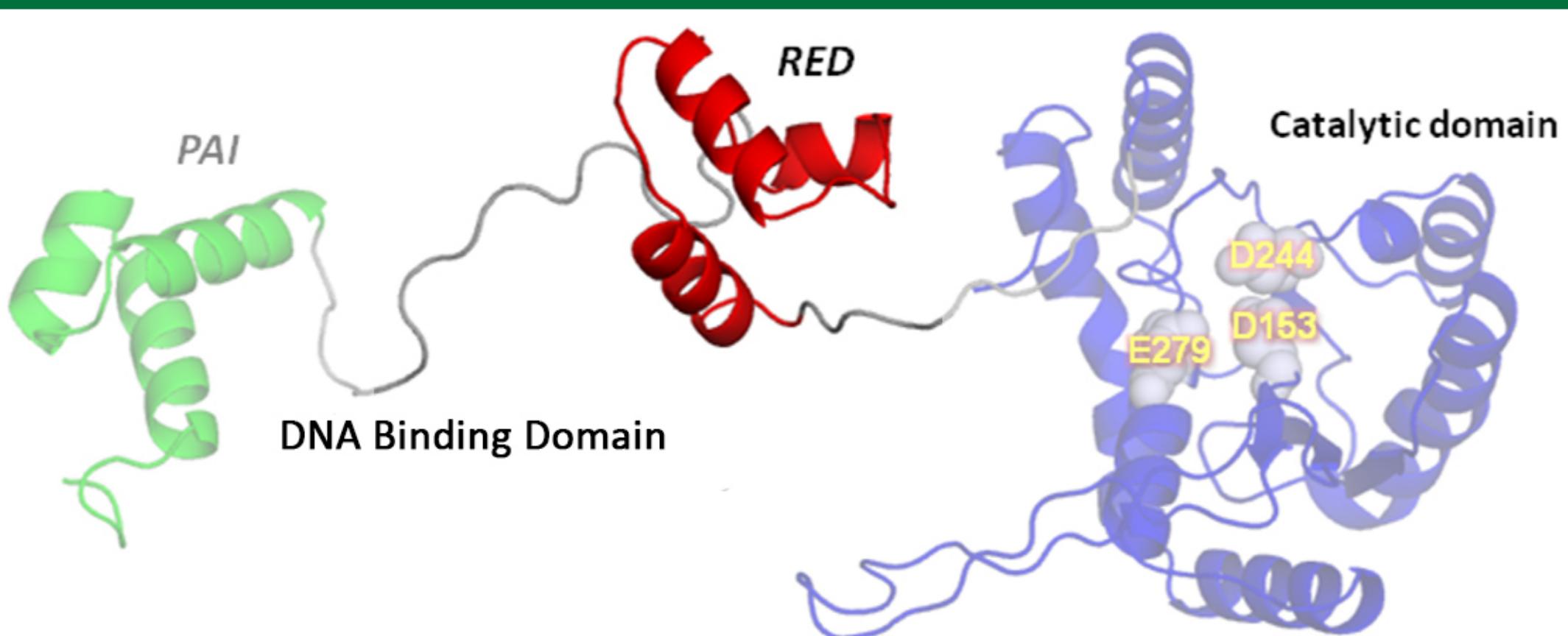
Department of Physics and Optical Science, University of North Carolina, Charlotte, NC, 28223

COLLEGE
RESEARCH
SCHOLARS

ABSTRACT

The Sleeping Beauty (SB) Transposon system is the most widely used DNA transposon and is currently the only DNA transposon in clinical trials for human gene therapy. The evolution of the SB transposon has been dependent on homology and biochemical data. The goal here has been to establish a more mechanistic approach to understanding the SB transposon. The DNA-binding domain of the SB transposase consist of two paired subdomains, PAI and RED. Much is understood about the PAI subdomain's function, however, there is very little understood about the RED subdomain's function. Here structural elucidation is being conducted to gain insight into the function of the RED subdomain of the SB transposon via NMR spectroscopy data analysis. NMR data of the SB transposon in solution allowed for ^1H - ^{15}N HSQC spectra to be utilized for the identification of NH groups in the primary structure of the SB transposon. CBCA(CO)NH and HNCACB spectral data was used to link the NH groups with their $\text{C}\alpha$, $\text{C}\alpha$ -1, $\text{C}\beta$, $\text{C}\beta$ -1. After backbone assignment the ^{15}N -TOCSY spectral data was used for side chain assignment to determine the $\text{H}\alpha$, $\text{H}\beta$, and $\text{H}\gamma$, $\text{H}\delta$, if present in the particular amino acid. Lastly, the ^{13}C -NOESY-HSQC and ^{15}N -NOESY-HSQC spectral data was used to find structural restraints for the SB transposon. Understanding the mechanics of this DNA transposon opens the door for applications of the SB transposon to areas of molecular biology, cell biology, and genome engineering; to list a few.

DNA BINDING DOMAIN

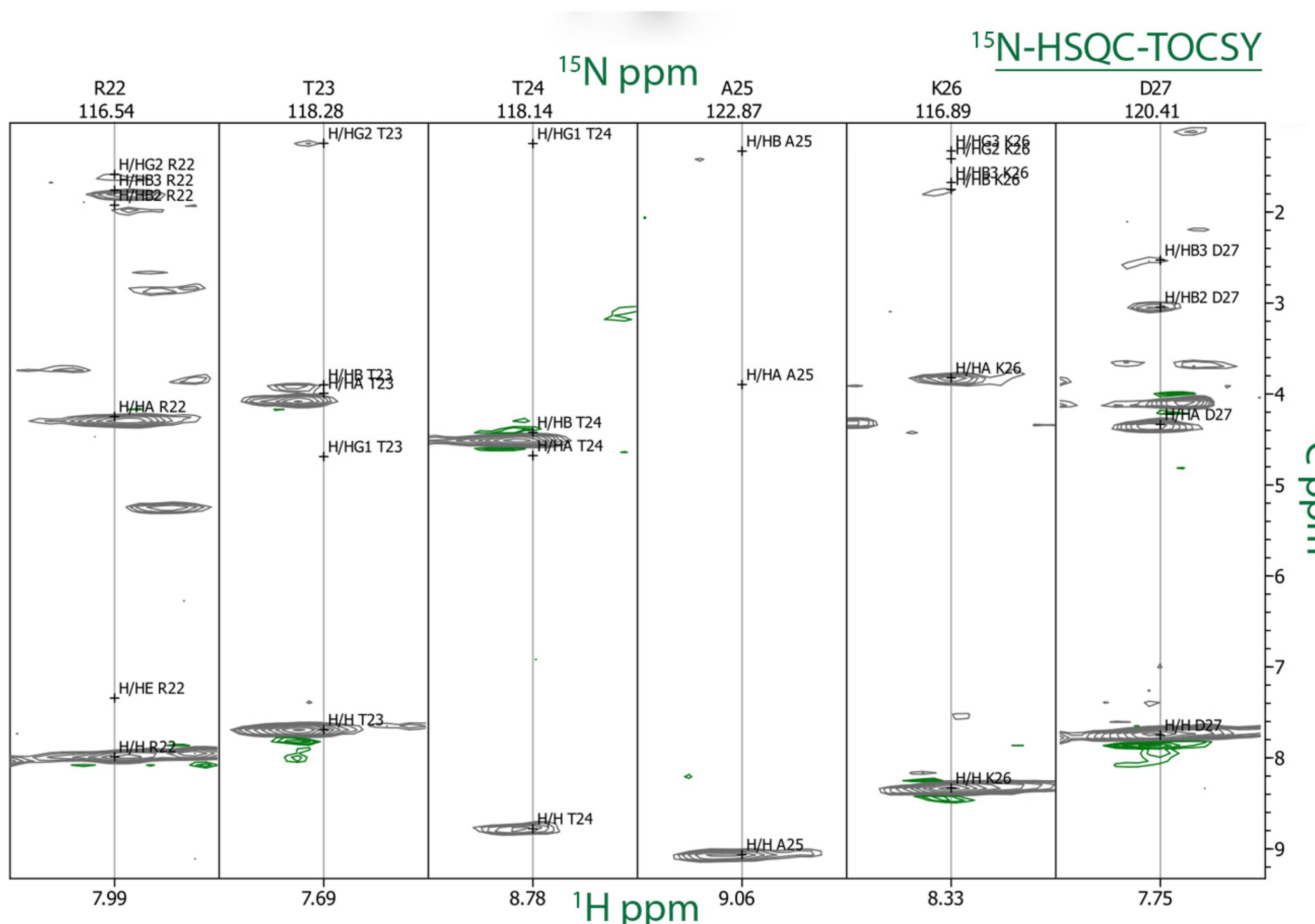
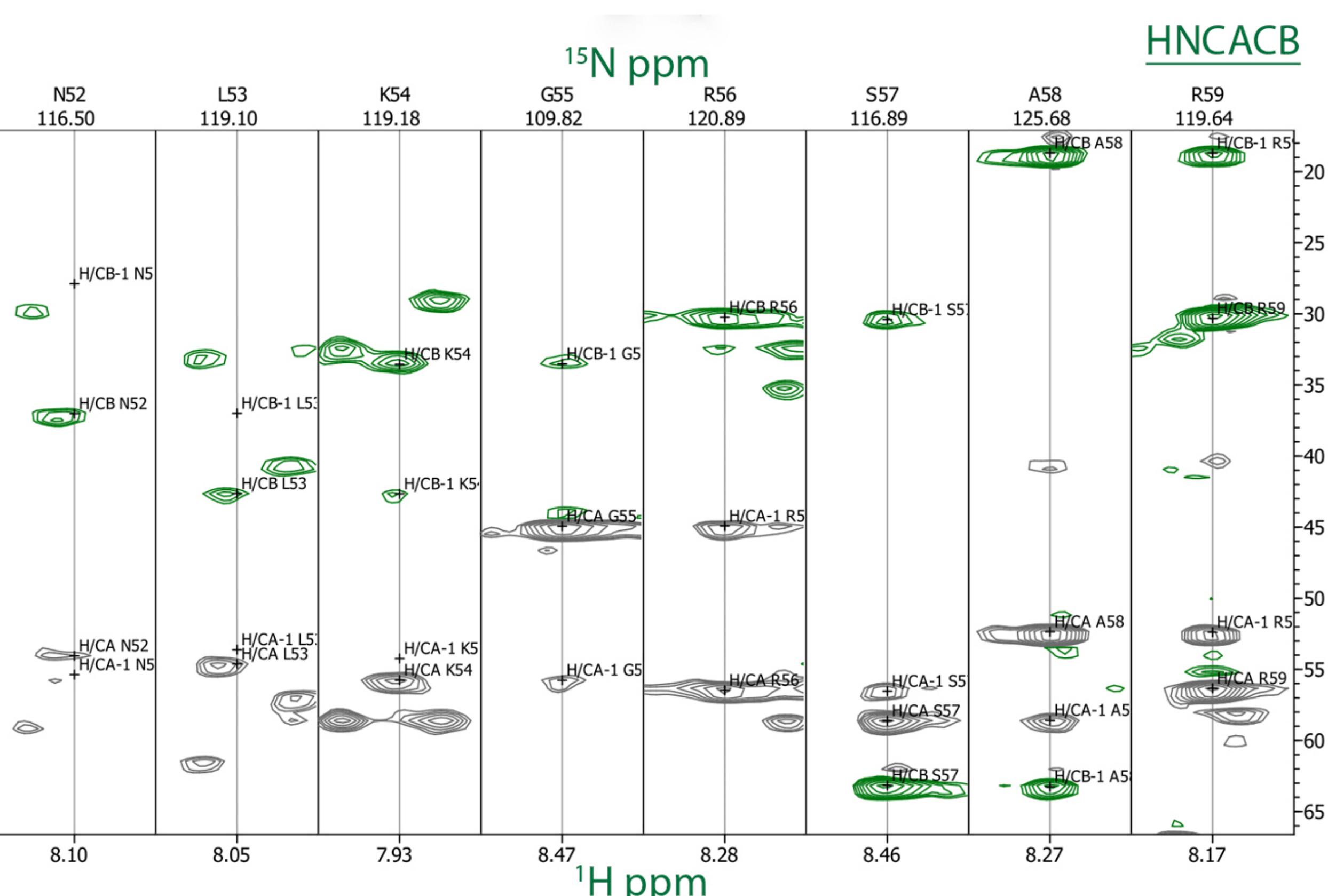
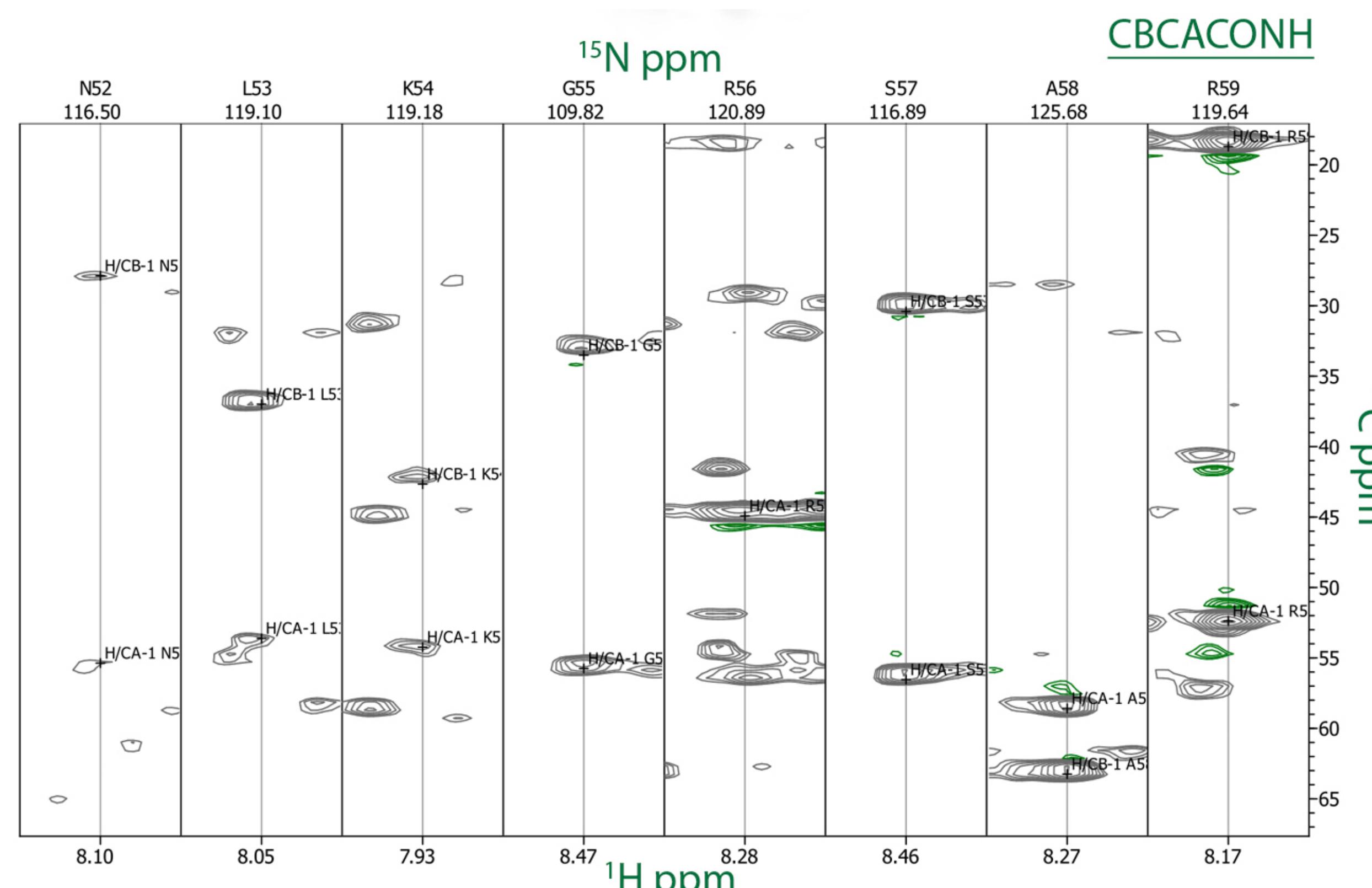


- Two subdomains in the DNA binding domain: PAI and RED
- RED domain's function is unclear
- The goal is to understand the RED subdomain's role in the DNA binding domain

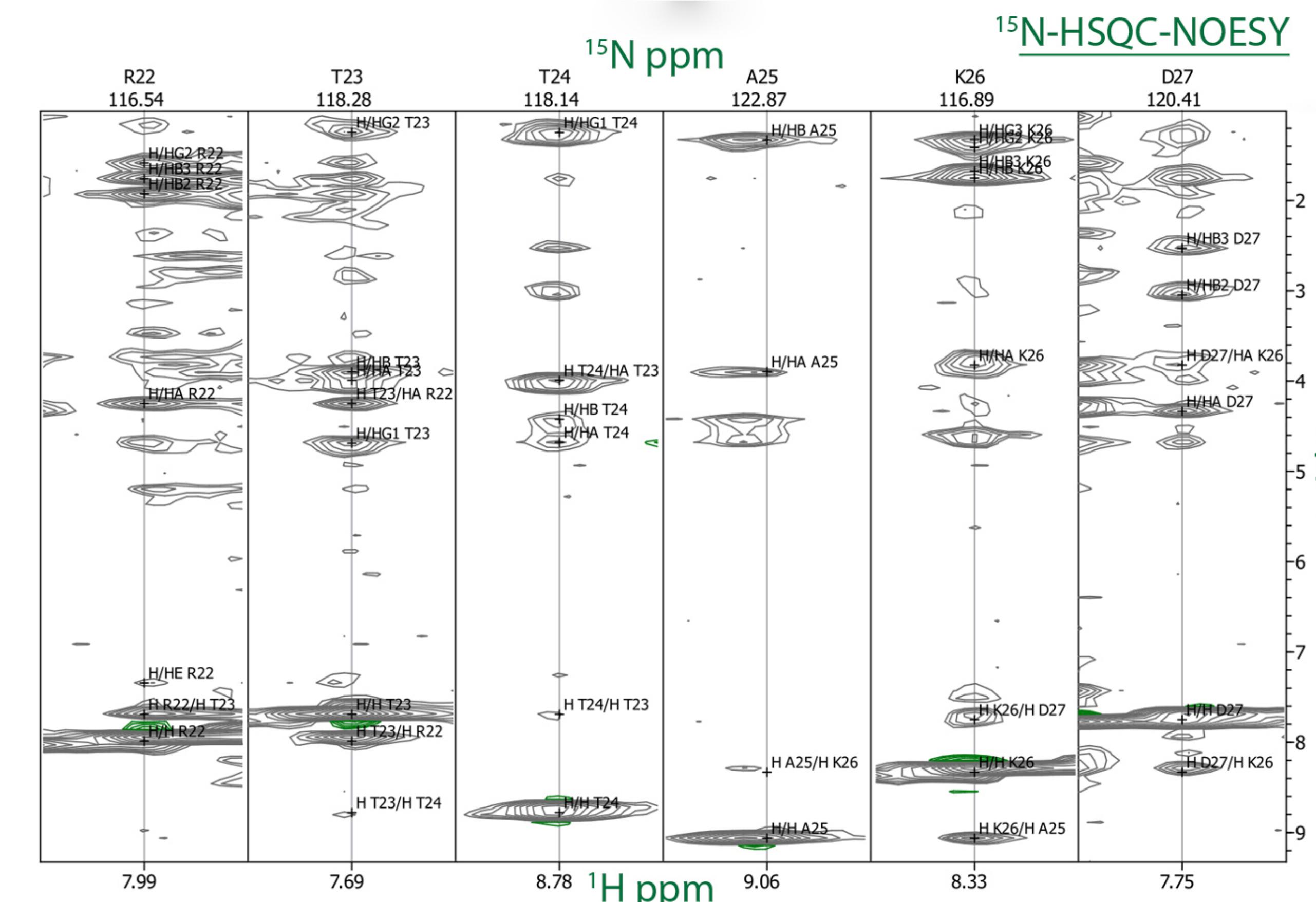
METHODS

- The program C.A.R.A. (computer aided resonance analysis) was used to process the different spectrums collected from the NMR experiments.
- The CBCACONH spectrum correlates ^1H and ^{15}N amide resonances of one residue to the $\text{C}\alpha$ and $\text{C}\beta$ resonances in the predecessor residue via $J(\text{NH})$, $J(\text{N},\text{CO})$, $J(\text{CA},\text{CO})$ and $J(\text{CA},\text{CB})$ coupling constants. Useful in identifying the $\text{C}\alpha$ -1 and $\text{C}\beta$ -1 resonances of a specific residue
- The HNCACB spectrum correlates ^1H and ^{15}N amide resonances of one residue to the inter- or intra-residue's $\text{C}\alpha$ and $\text{C}\beta$ via (NH) , $J(\text{N},\text{CA})$ and (CA,CB) coupling constants. Aided in the sequencing SB transposon and identifying $\text{C}\alpha$, $\text{C}\alpha$ -1, $\text{C}\beta$, and $\text{C}\beta$ -1 resonances.
- The 3D HSQC-TOCSY spectrum correlates ^{15}N resonances of one residue with HN , $\text{H}\alpha$, $\text{H}\beta$, $\text{H}\gamma$, and $\text{H}\delta$ (if applicable for the given amino acid residue) resonances within the residue. Aided in the sequencing of the primary structure of the SB transposon when ambiguities arose in the HNCACB spectrum.
- The 3D HSQC-NOESY spectrum correlates ^{15}N , HN , $\text{H}\alpha$, $\text{H}\beta$, and $\text{H}\gamma$ resonances of one residue to the HN , $\text{H}\alpha$, $\text{H}\beta$, and $\text{H}\gamma$ resonances of the predecessor or successor residue. Aided in the sequencing of the primary structure and side chain assignment of the SB transposon when ambiguities arose in the HNCACB and 3D HSQC-TOCSY spectrum.

RESULTS



RESULTS CONTINUED



Amino Acid sequence of RED Domain

ASMVLSDLRDERTLVRKVQINPRTTAKDLV**KMLEETGT**
KVSISTVKRVLYRHNLKGRSARKKLEHHHHHH

- Each Letter Represents one of the 20 common amino acids.
- The Highlighted letters display how much of the RED domain has been sequenced thus far.

CONCLUSIONS

- The RED domain at 20 degrees celcius in a solution of Na_2SO_4 at 5.0 PH seems to be well structured and mostly folded.
- More than two thirds of the RED domain's amino acid sequence was assigned.
- The $\text{C}\alpha$, $\text{C}\alpha$ -1, $\text{C}\beta$, and $\text{C}\beta$ -1 resonances were assigned to all identified spin systems
- All side-change assignments of HN , $\text{H}\alpha$, and $\text{H}\beta$ have been identified
- Several very distinct side-chains of $\text{H}\gamma$ and $\text{H}\delta$ have been identified thus far

FUTURE DIRECTION

- Finish amino acid assignments and calculate 3D representation of the RED domain system
- Understand more about the DNA binding domain and the role the RED subdomain plays.

ACKNOWLEDGEMENTS

I would like to thank Dr. Nesmelova for mentoring me over the past nine weeks. Thank you to the CRS program for giving me this great opportunity to jump feet first into the field of research and data analysis.