

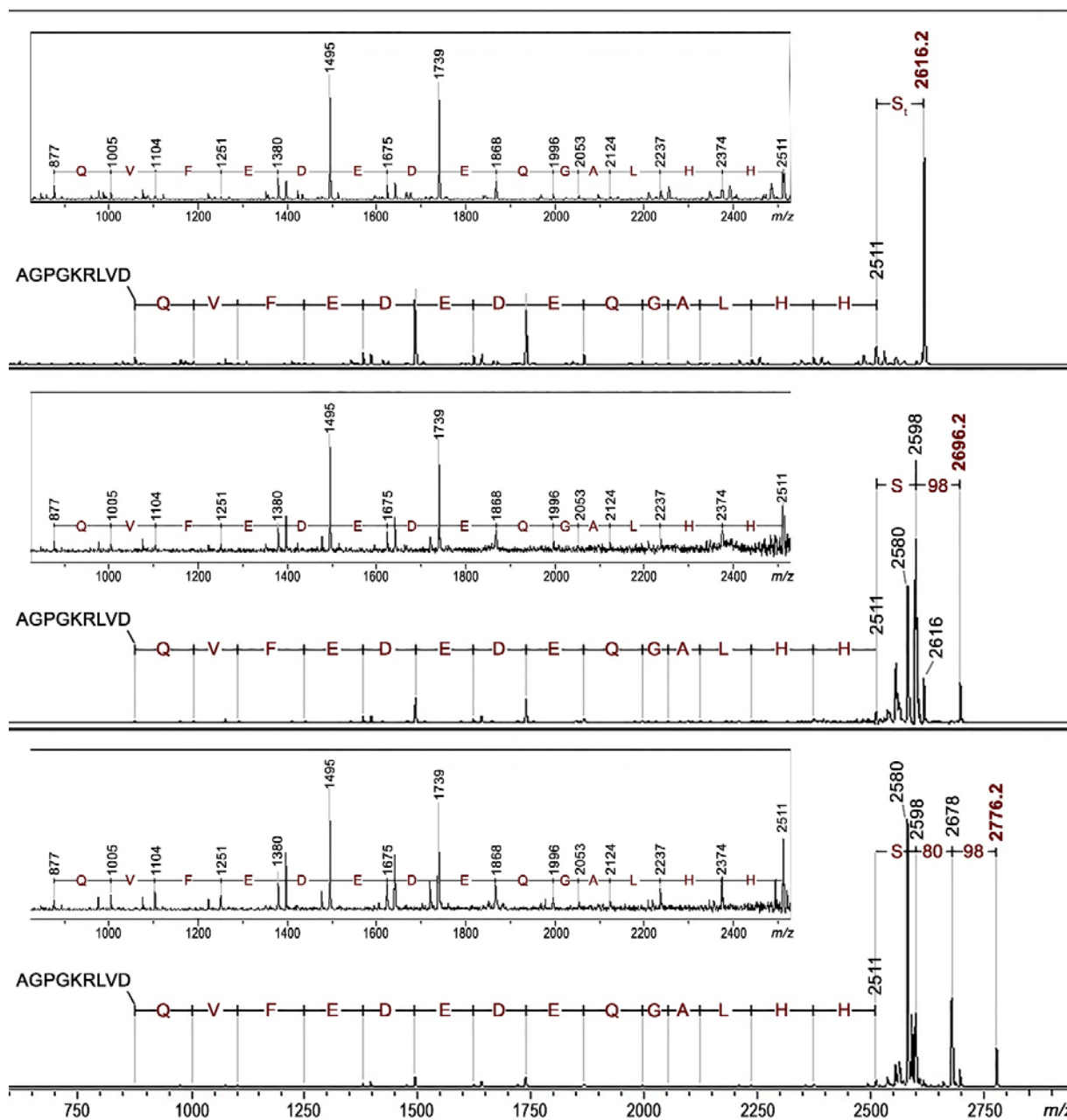
Supplementary Figures

**Efficient *in vivo* Synthesis of Lasso Peptide Pseudomycoidin Proceeds in the Absence of both the Leader and the Leader Peptidase**

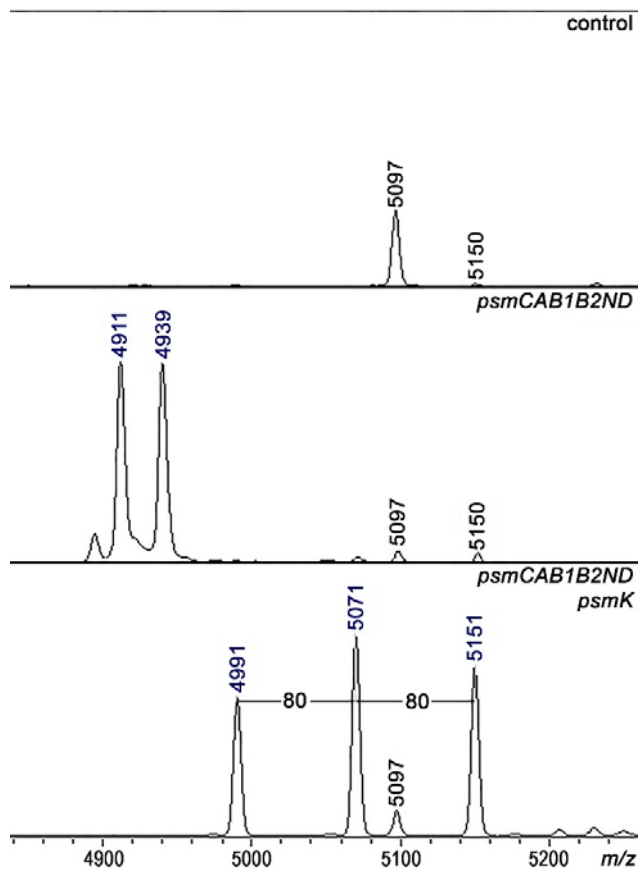
**Tatyana Zyubko, Marina Serebryakova, Julia Andreeva, Mikhail Metelev, Guy Lippens, Svetlana Dubiley, and Konstantin Severinov**

**Figure S1. MALDI MS-MS analysis of unphosphorylated and phosphorylated forms of pseudomycoidin.**

Upper panel – the fragmentation spectrum of unphosphorylated pseudomycoidin (average  $m/z = 2618$ , monoisotopic  $m/z = 2616.2$ ). The peaks corresponding to the C-terminal fragments of the lasso peptide are marked. Low-intensity daughter peaks are magnified in the insert. The daughter peak with  $m/z = 877$  matches the macrocycle formed from the N-terminal AGPGTSTPD peptide. Middle and lower panels – fragmentation spectra of monophosphorylated (average  $m/z = 2698$ , monoisotopic  $m/z = 2696.2$ ) and diphosphorylated (average  $m/z = 2778$ , monoisotopic  $m/z = 2776.2$ ) pseudomycoidin. Fragment analysis shows the presence of phosphate group(s) attached to the C-terminal serine residue. Mass differences of 98 and 80 Da match  $H_3PO_4$  and  $-HPO_3$  groups, respectively.



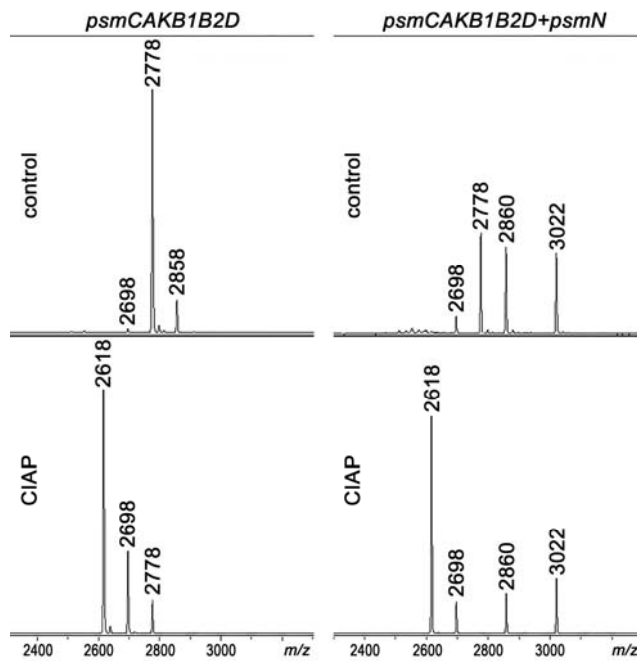
MALDI-MS analysis of the *E. coli* BL21(DE3) cells containing control plasmids (upper panel), plasmids expressing *psmCABIBIND* (middle panel), or plasmids expressing complete *psm* cluster (lower panel). Average  $m/z$  = 4911 and 4939 match the PsmA precursor peptide and its formylated form, respectively. Average  $m/z$  = 4991, 5071, and 5151 match mono-, di-, and triphosphorylated PsmA precursor peptide, respectively. The  $m/z$  values of mass-peaks specific for cells carrying the *psm* genes are marked with blue-color font.





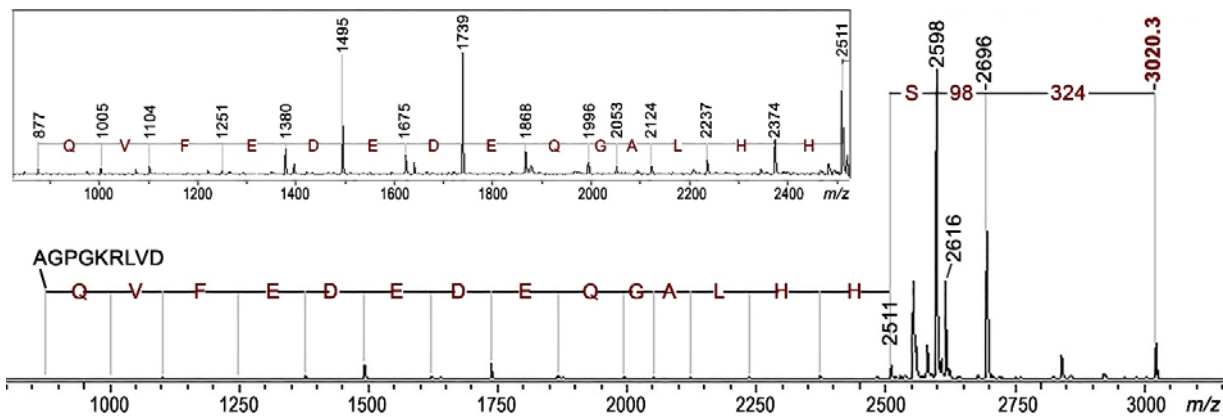
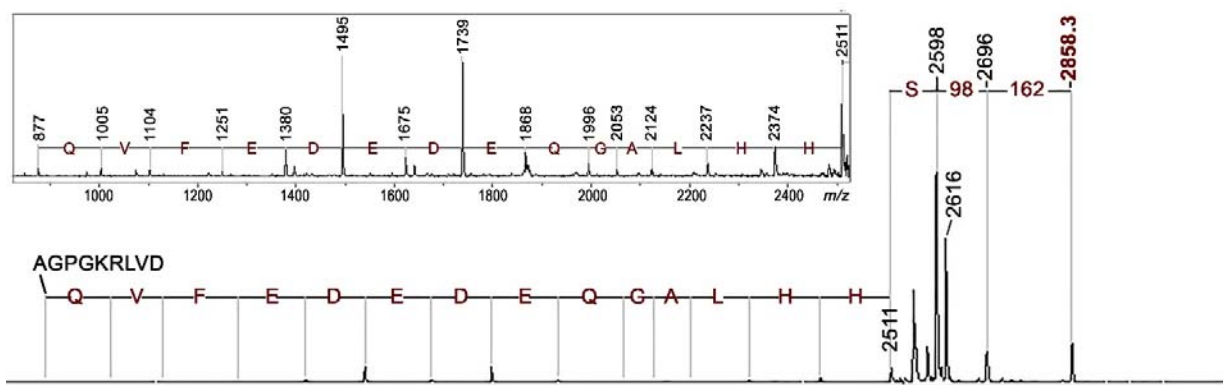
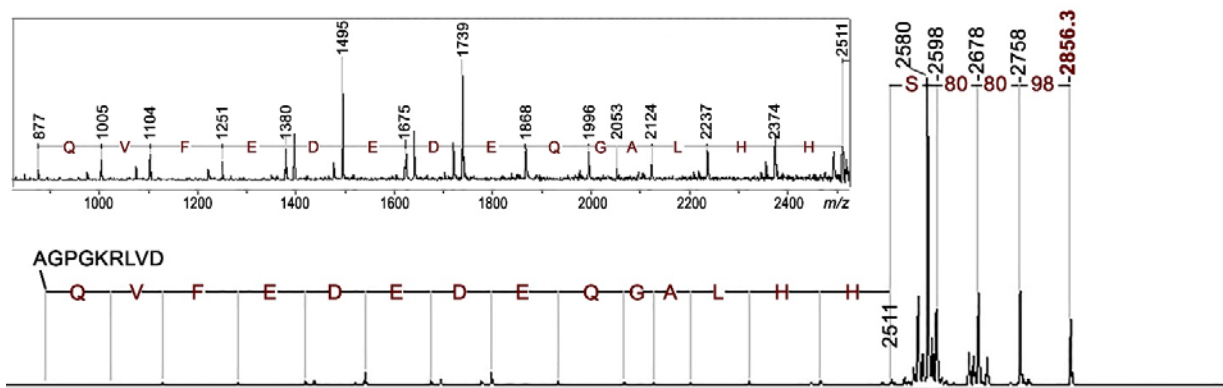
**Figure S3. The +162 Da PsmN-dependent modification protects phosphorylated pseudomycoidin from enzymatic dephosphorylation.**

MALDI-MS analysis of extracts prepared from *E. coli* cells carrying the *psm* cluster with deleted *psmN* deletion and an additional control (left panels) or *psmN*-overexpressing plasmid (right panels). Where indicated, extracts were treated with CIAP prior to mass-spectrometric analysis.



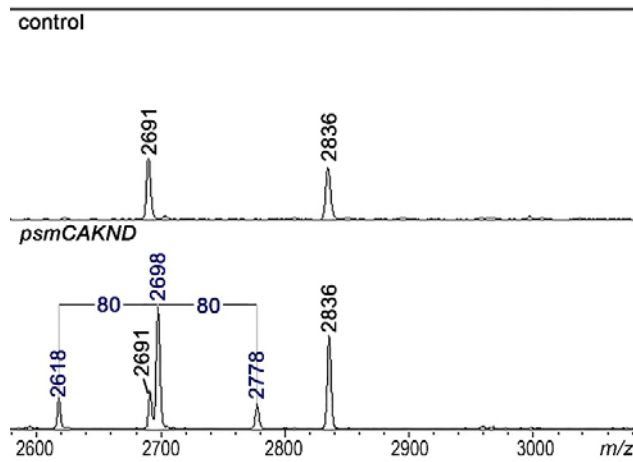
**Figure S4. MALDI MS-MS analysis of different forms of pseudomycoidin.**

Upper panel – fragmentation spectrum of triphosphorylated lasso pseudomycoidin (monoisotopic  $m/z$  = 2856.3, average  $m/z$  = 2858). Peaks corresponding to C-terminal lasso peptide fragments are marked. Low-intensity daughter peaks corresponding to fragments that arise due to breakage of peptide binds inside pseudomycoidin are magnified in the insert. Mass differences of 98 and 80 Da match the  $H_3PO_4$  and  $HPO_3$  groups, respectively. Middle panel – fragmentation spectrum of monoisotopic  $m/z$  = 2858.3 (average  $m/z$  = 2860) pseudomycoidin form. A mass difference of 162 Da between the mother ion and the first daughter ion ( $m/z$  = 2696) corresponds to the addition of a  $C_6H_{10}O_5$  group, which matches an unmodified hexose residue. The mass difference between the first daughter ion and the dehydrated C-terminal Ser pseudomycoidin ( $m/z$  = 2598) equals 98 Da, which matches the  $H_3PO_4$  group, and indicates that the hexose moiety is attached to the C-terminal phosphate. Lower panel – fragmentation spectrum of monoisotopic  $m/z$  = 3020.3 (average  $m/z$  = 3022) pseudomycoidin form. The mass difference of 324 Da between the daughter peak of phosphorylated peptide ( $m/z$  = 2696) and the mother peak matches the dihexose moiety.



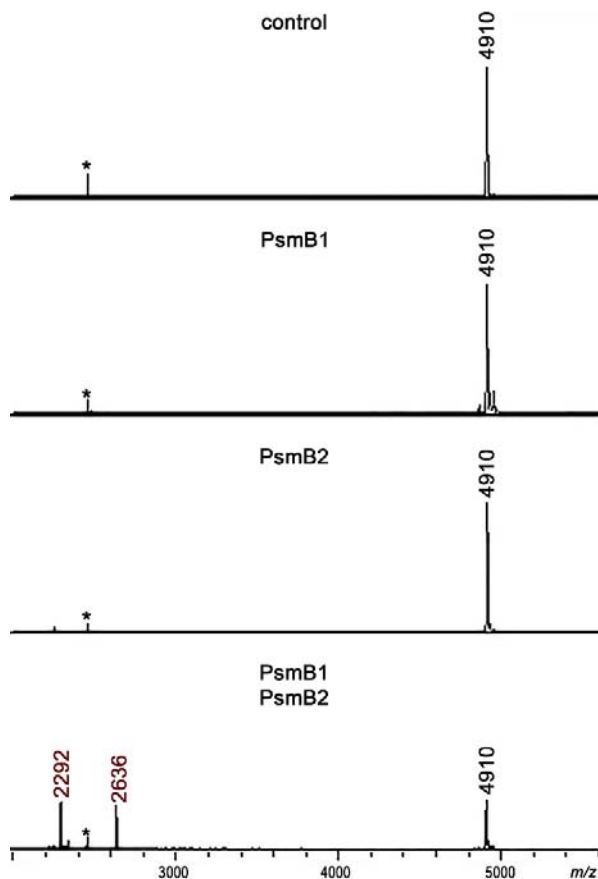
**Figure S5. PsmB1 and PsmB2 are dispensable for pseudomycoidin maturation.**

MALDI MS analysis of the *E. coli* BL21(DE3) cells harboring the *psm* cluster without the *psmB1B2* genes. Mass-peaks corresponding to unmodified lasso peptide (average  $m/z = 2616$ ) and its mono- and diphosphorylated forms (average  $m/z = 2698$  and  $2778$ , respectively), are labeled with the blue color font.



**Figure S6. PsmB1 and PsmB2 cleave the leader peptide of the pseudomycoidin precursor peptide *in vitro*.**

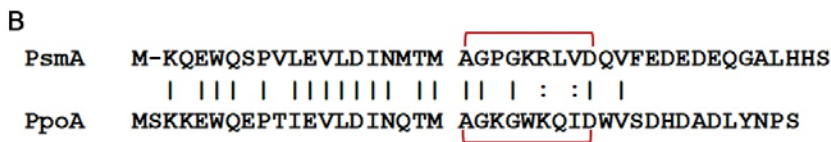
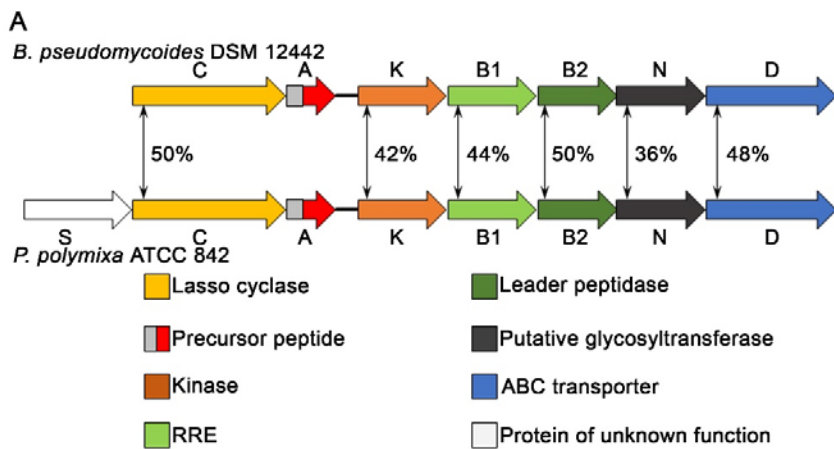
MALDI MS spectra of the products or reaction of incubation of the PsmA propeptide with indicated proteins. The mass peak with average  $m/z = 4910$  corresponds to PsmA; mass peaks with average  $m/z = 2292$  and  $2636$  correspond to the leader and core peptide parts, respectively. The peak marked with asterisk is the  $[MH^{++}]$  ion of full-sized PsmA.



**Figure S7. Comparison of lasso peptide gene clusters from *B. pseudomycoides* DSM 12442 and *P. polymixa* ATCC 842.**

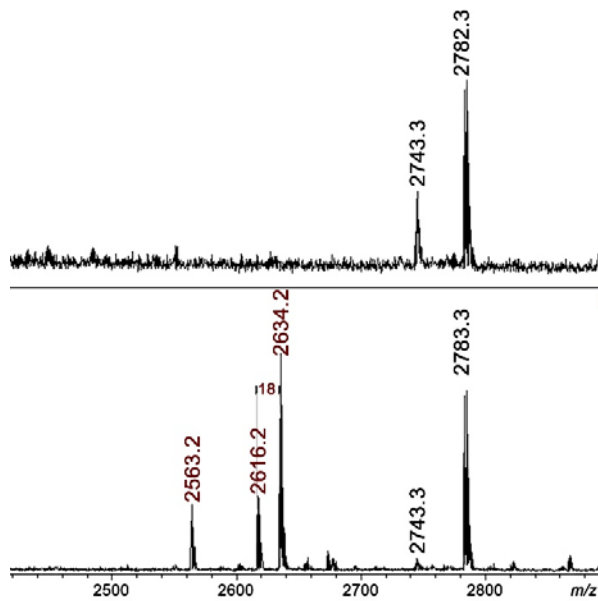
(A) A schematic representation of the gene arrangements in the clusters. Genes are indicated by arrows, and homologous genes are indicated by identical colors. Numbers indicate the percentage of identity between homologous protein sequences. Known or putative gene product functions are listed at the bottom. PpoS has homology to the sulfotransferase family of proteins.

(B) An alignment of precursor peptides from *B. pseudomycoides* DSM 12442 and *P. polymixa* ATCC 842. Amino acids forming the macrolactam ring are linked by red brackets.



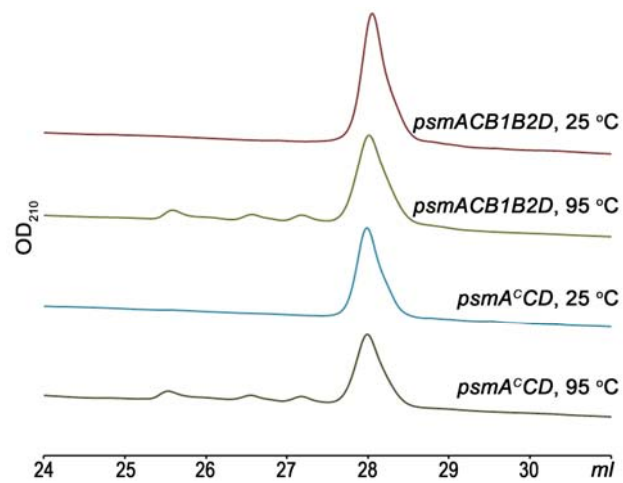
**Figure S8. Pseudomucoidin production in *B. subtilis*.**

MALDI MS spectra of *B. subtilis* cells harboring the control plasmid (upper panel) or a derivative plasmid expressing the *psmCAD* gene set (lower panel). The monoisotopic  $m/z = 2634.2$  peak corresponds to the core part of PsmA propetide (amino acids 1-24); the monoisotopic  $m/z = 2563.2$  peak matches incorrectly processed PsmA (amino acids 2-24); the monoisotopic  $m/z = 2616.2$  corresponds to pseudomucoidin lasso peptide.



**Figure S9. Thermostability of pseudomycoidin.**

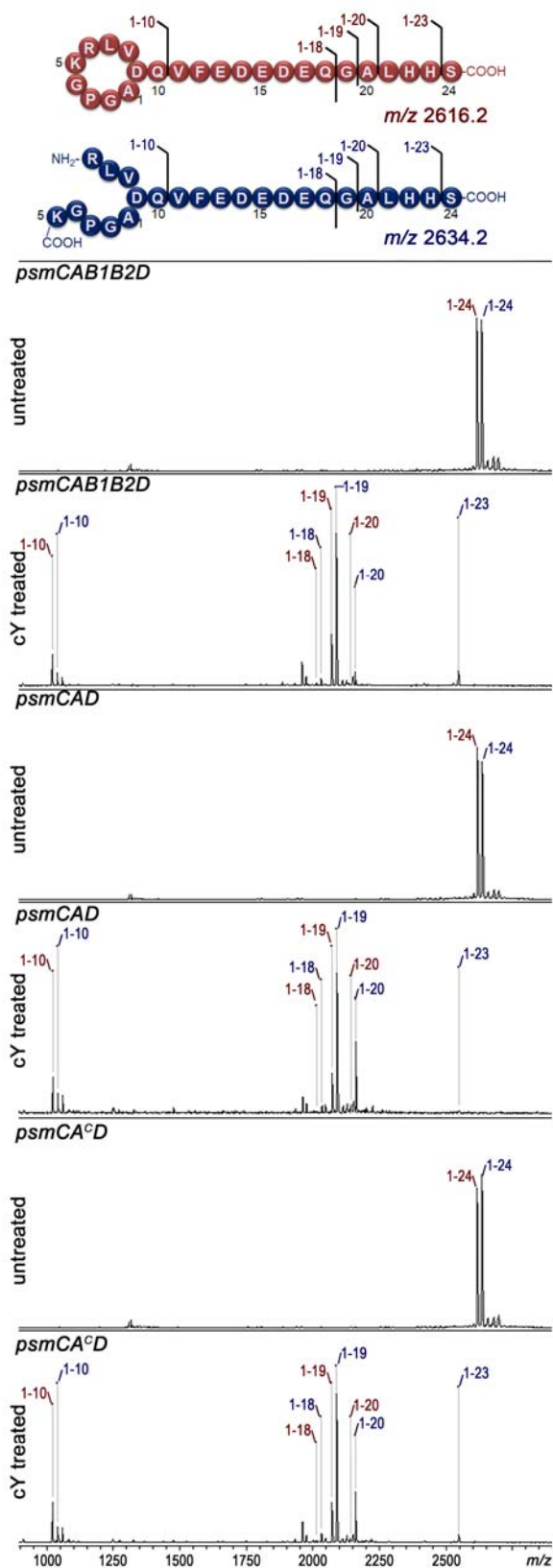
HPLC traces of pseudomycoidin purified from cells harboring *psmACB1B2D* or *psmA<sup>c</sup>CD* gene sets before and after 2-hour incubation at 95 °C.





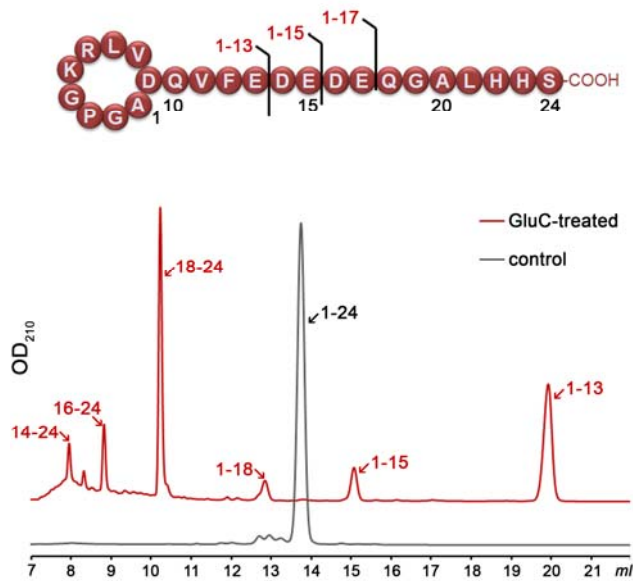
**Figure S10. Digestion of intact and trypsin-treated pseudomycoidin with carboxypeptidase Y.**

Lasso peptides were produced by the *E. coli* cells harboring psm gene sets indicated at top of each pair of panels. The lasso peptides were partially digested with trypsin to open the macrocycle ring and release the potentially trapped C terminus, thus creating the branched peptide with two C-termini, as schematically shown on the top. Monoisotopic  $m/z$  of two forms of the peptide are shown under the schematic structures. Trypsin digestion product is distinguished from intact pseudomycoidin by an 18-Da mass shift. Mixtures of trypsin-treated and intact pseudomycoidin were treated with carboxypeptidase Y and the products of the reactions were analyzed with MALDI MS. The smallest proteolytic fragments identified (1-10) correspond to N-terminal macrocycle extended by one amino acid. Peaks corresponding to branched peptides originating from trypsin-digested pseudomycoidin and intact lasso peptide are marked in red and blue, respectively.



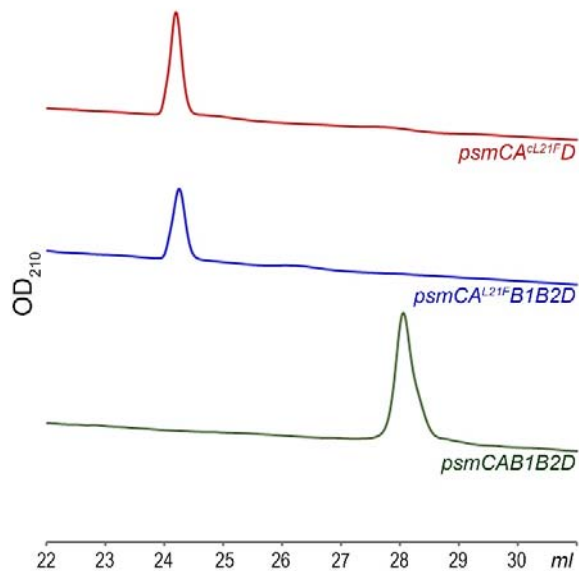
**Figure S11. HPLC traces of pseudomycoidin cleaved with Glu-C protease.**

The Glu-C cleavage sites in pseudomycoidin are schematically shown at the top of the figure. Pseudomycoidin was produced in *E. coli* cells harboring the minimal set *psmCA<sup>c</sup>D* gene set. The products of Glu-C proteolysis were separated by reverse phase HPLC along with untreated control (red and grey traces, respectively). HPLC peaks containing proteolytic fragments were identified using MALDI MS.



**Figure S12. HPLC traces of the wild type and L21F mutant forms of pseudomycoidin.**

Lasso peptides were produced using either complete (*psmCAB1B2D*) or minimal (*psmCA<sup>c</sup>D*) set of pseudomycoidin biosynthesis genes as indicated. Wild-type pseudomycoidin (average  $m/z = 2618$ ) and the L21F mutant (average  $m/z = 2652$ ) were identified in HPLC peaks by MALDI MS.



**Figure S13. Assigned homonuclear TOCSY (blue) and NOESY (red) spectra of wild type pseudomycoidin.**

Sample was a ~3mM solution of peptide dissolved in 200µl of phosphate buffer 100mM pH 6.5. All spectra were acquired on a 800MHz Bruker Avance II spectrometer (Bruker, Fällanden, Switzerland) equipped with a cryogenic triple resonance probe head and operating at 293K. NOESY spectrum was acquired with a mixing time of 200ms, as a complex matrix of 4096 x 400 points, with 16 scans per increment. TOCSY spectrum was acquired with a 70ms mixing time implemented as a dipsi2 pulse train, and was equally acquired as a complex matrix of 4096 x 400 points, with 16 scans per increment. Spectra were Fourier transformed after apodization with a shifted square sine bell in both dimensions using the Topspin3.0 software package (Bruker, Fällanden, Switzerland). Assignment were done as described by Wüthrich (K. Wüthrich, NMR of proteins and Nucleic Acids, Wiley, 1986). The following panels demonstrate the assignment procedure by highlighting one residue per panel. Characteristic NOEs that were used for the assignment are circled, and explicit assignment is given.

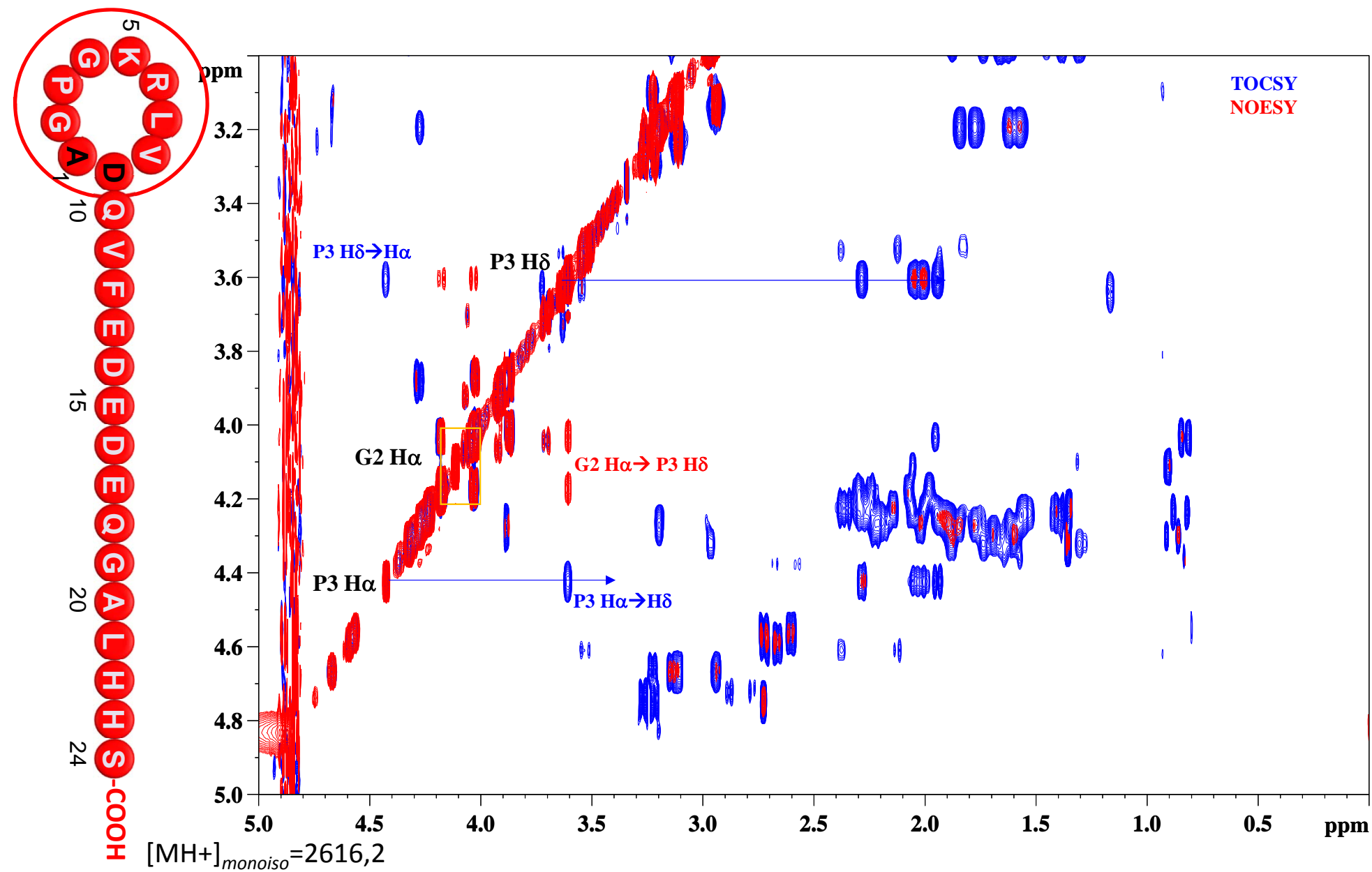


Figure S13  
wt - 1

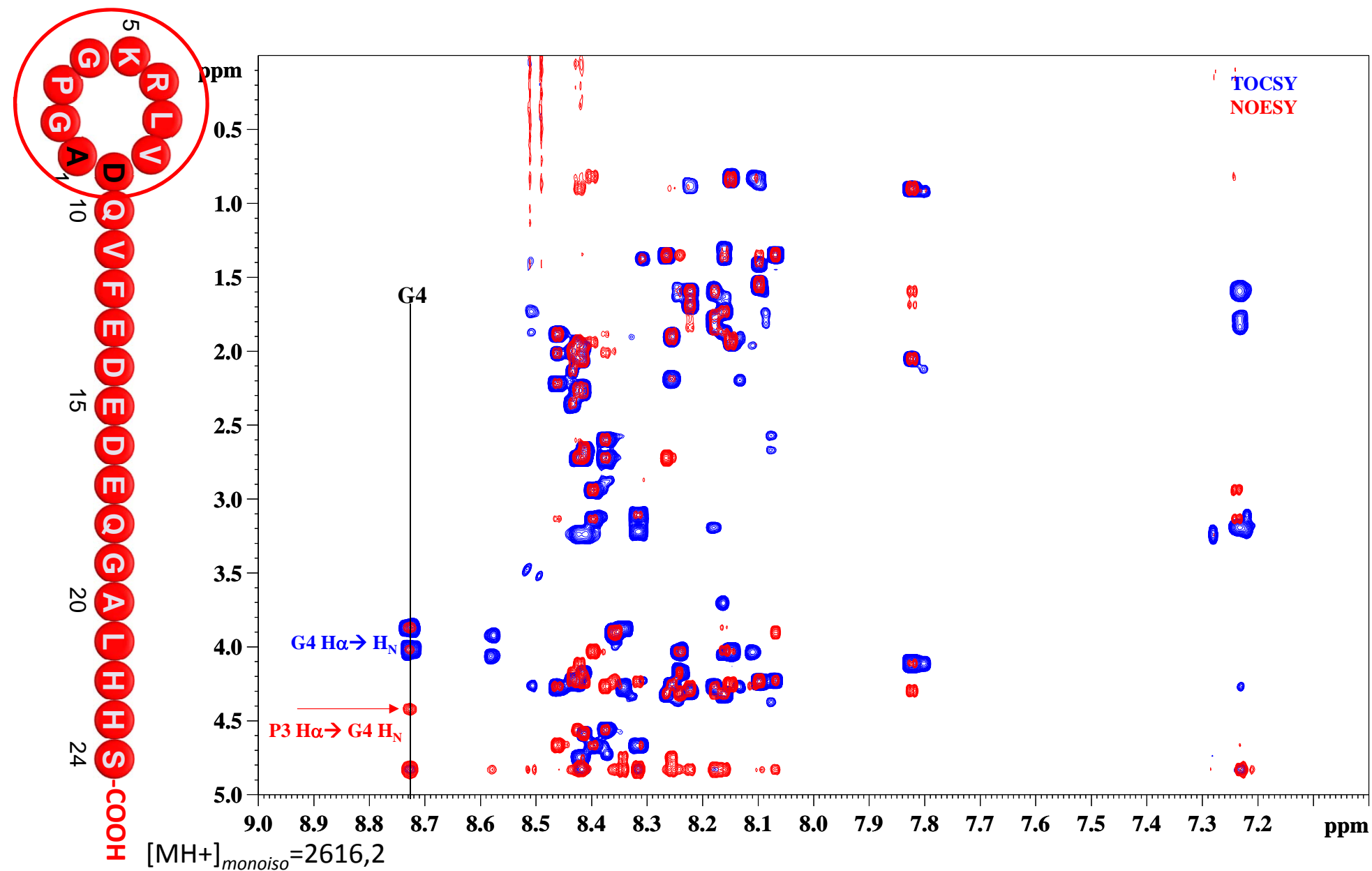


Figure S13  
wt - 2

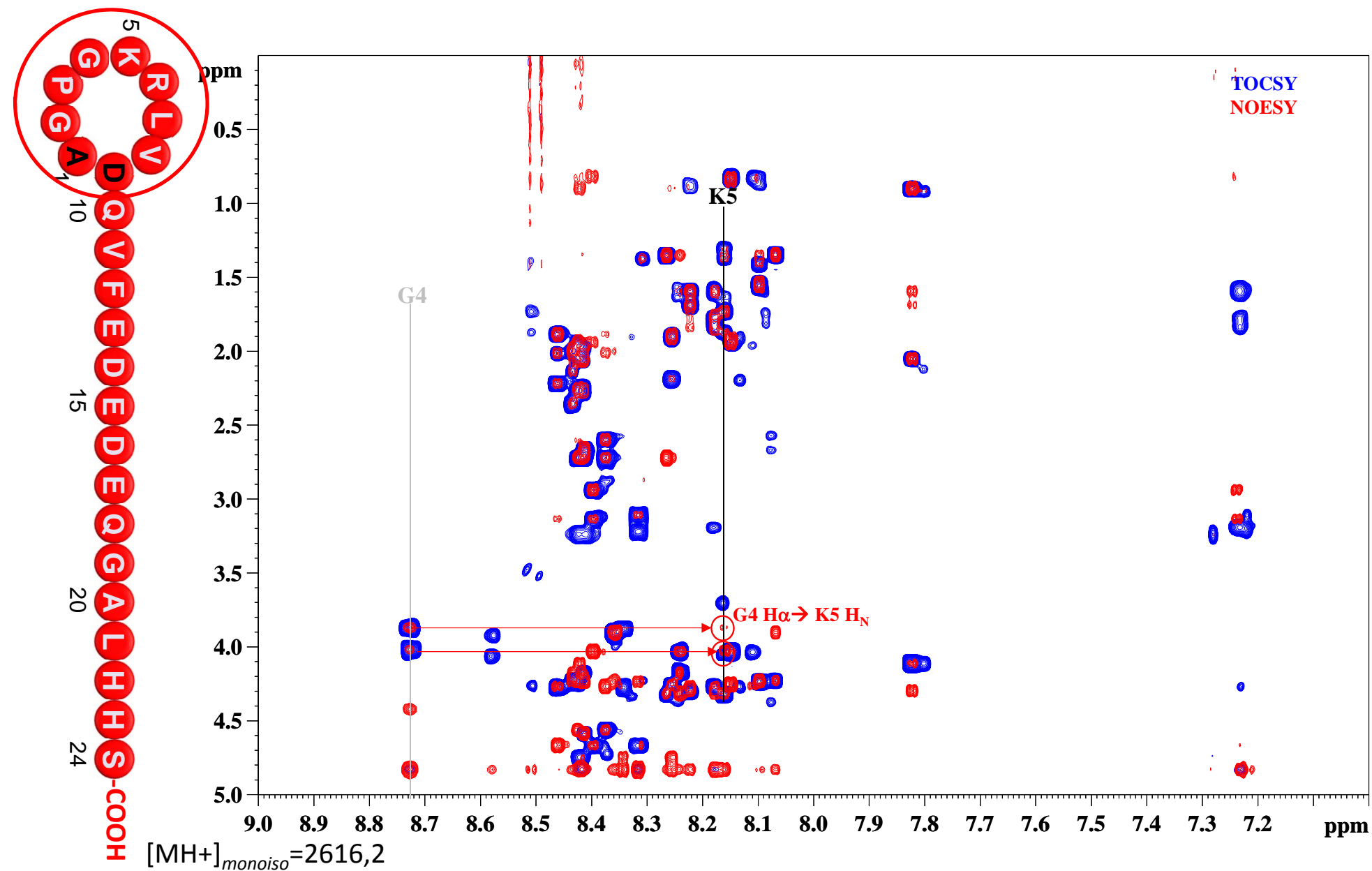
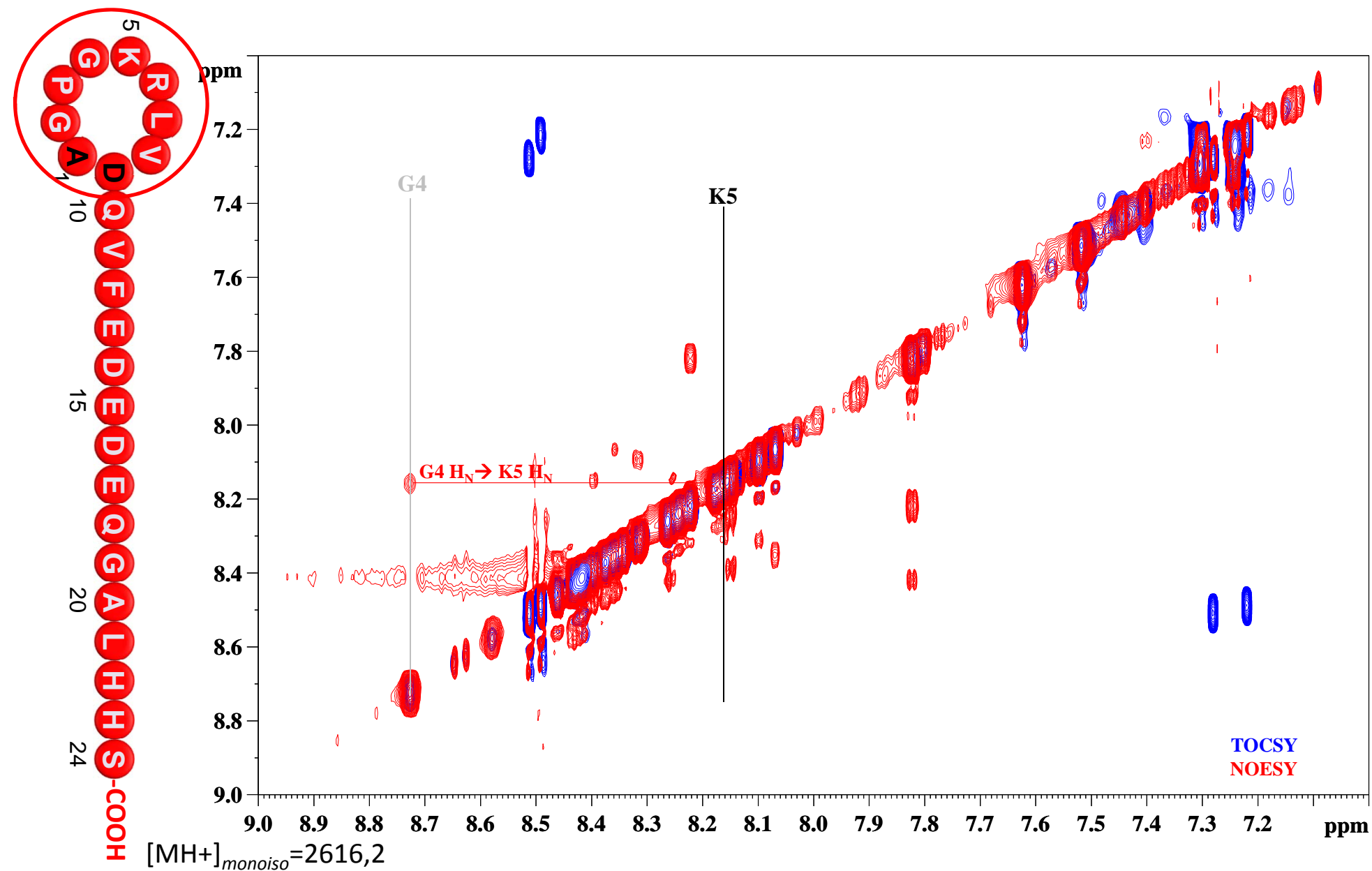


Figure S13  
wt - 3





**Figure S13**  
**wt - 4**

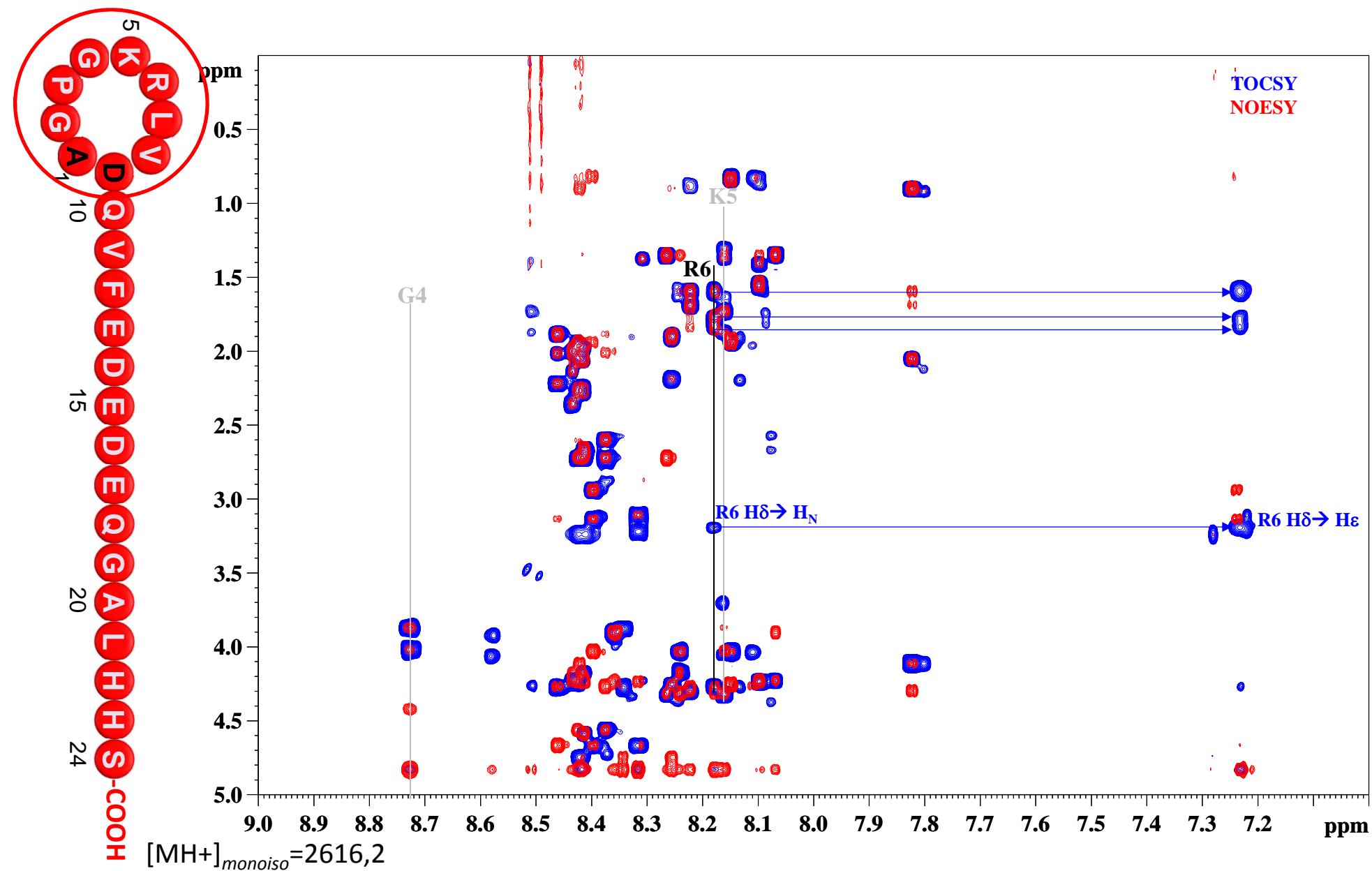


Figure S13  
wt - 5

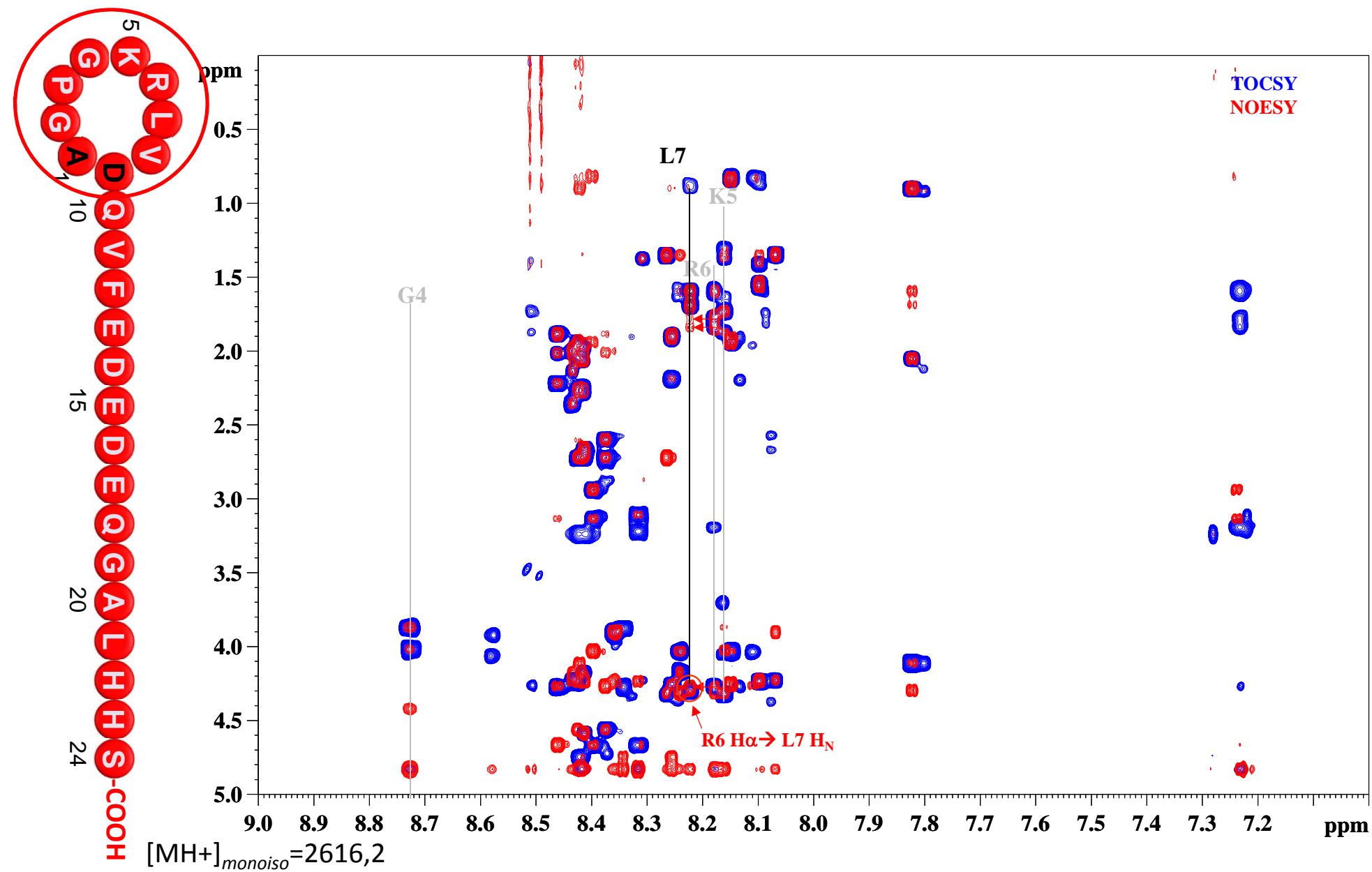


Figure S13  
wt - 6

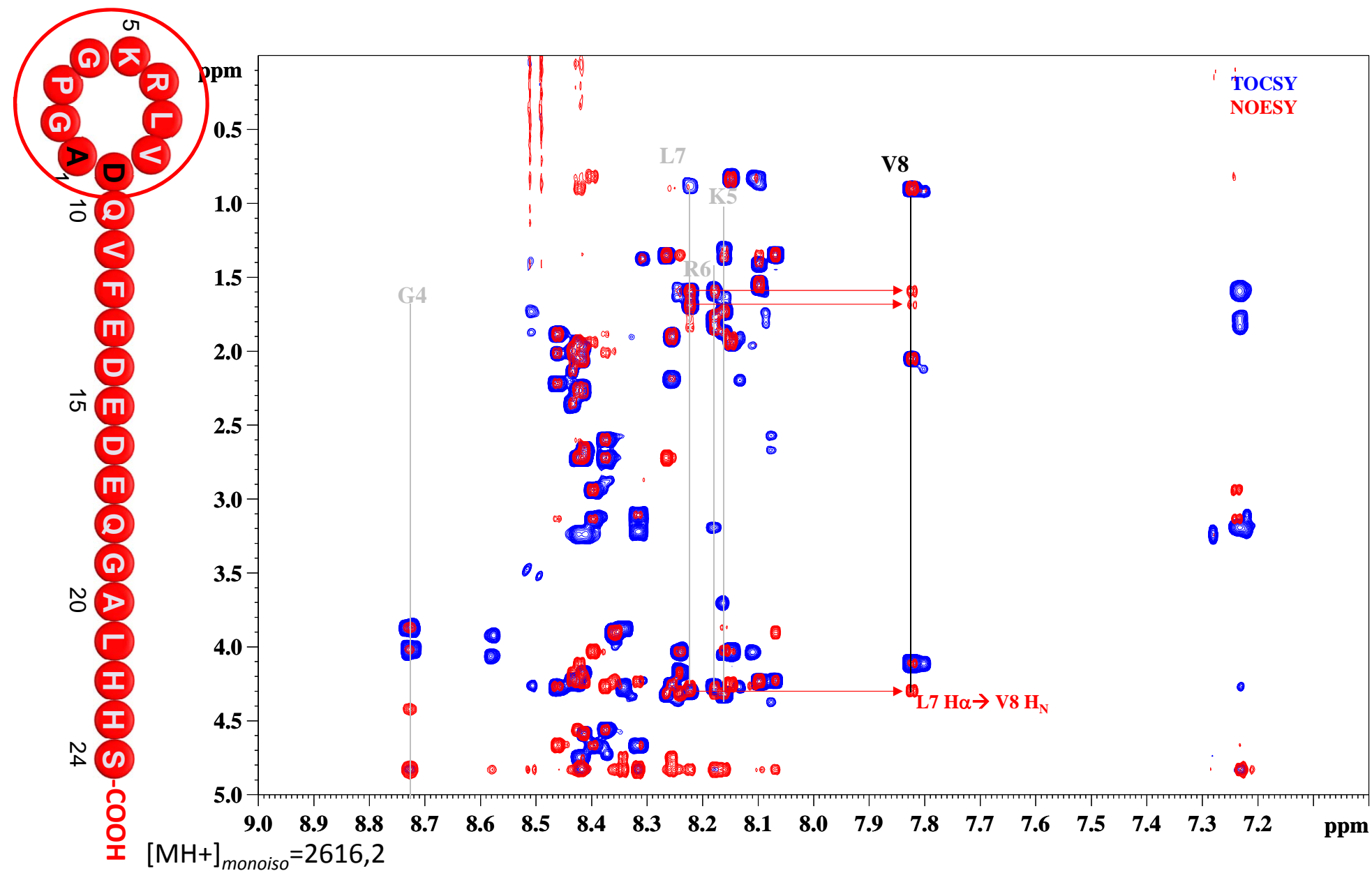


Figure S13  
wt - 7

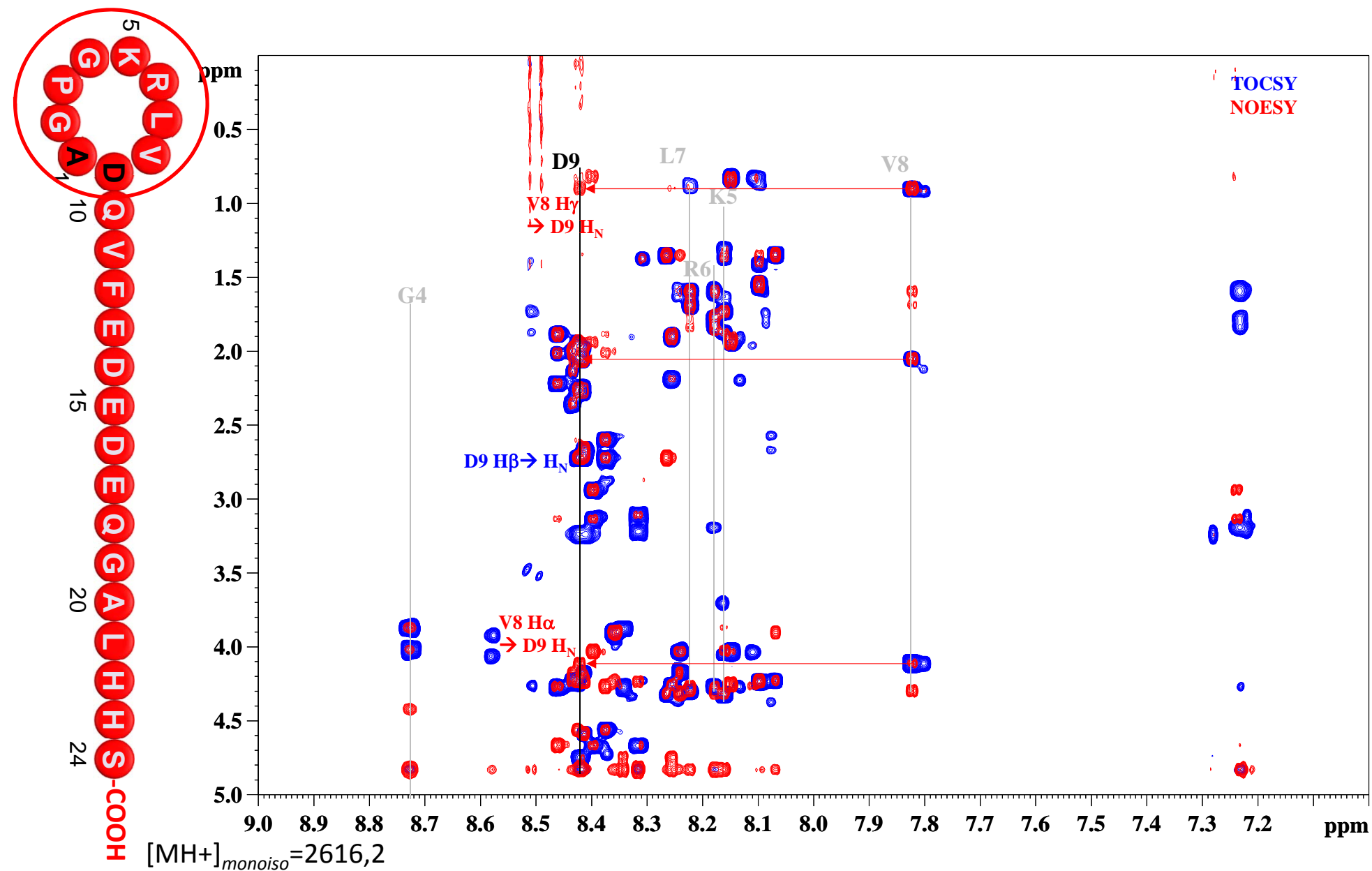
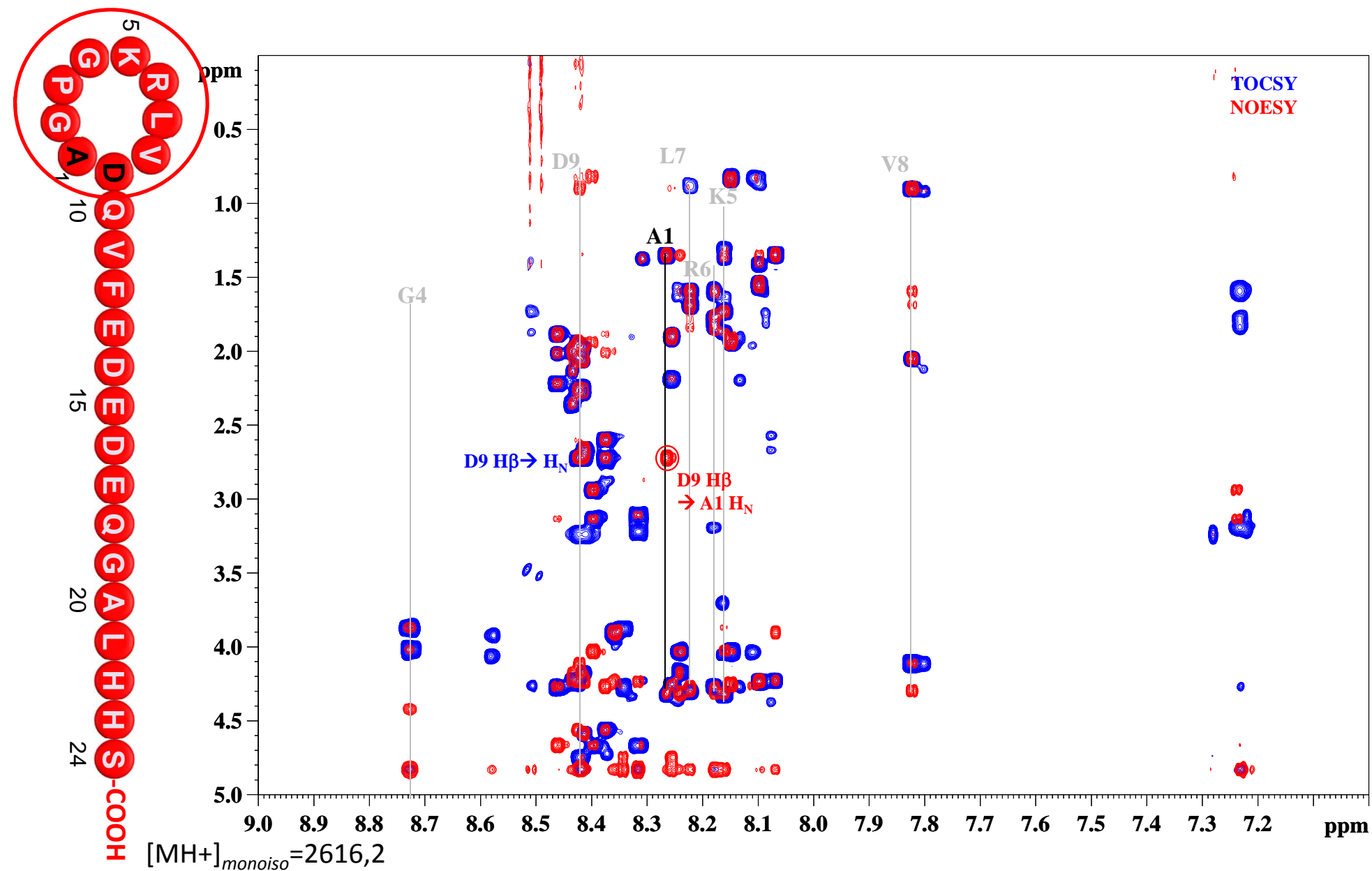


Figure S13  
wt - 8



**Figure S13**  
wt - 9

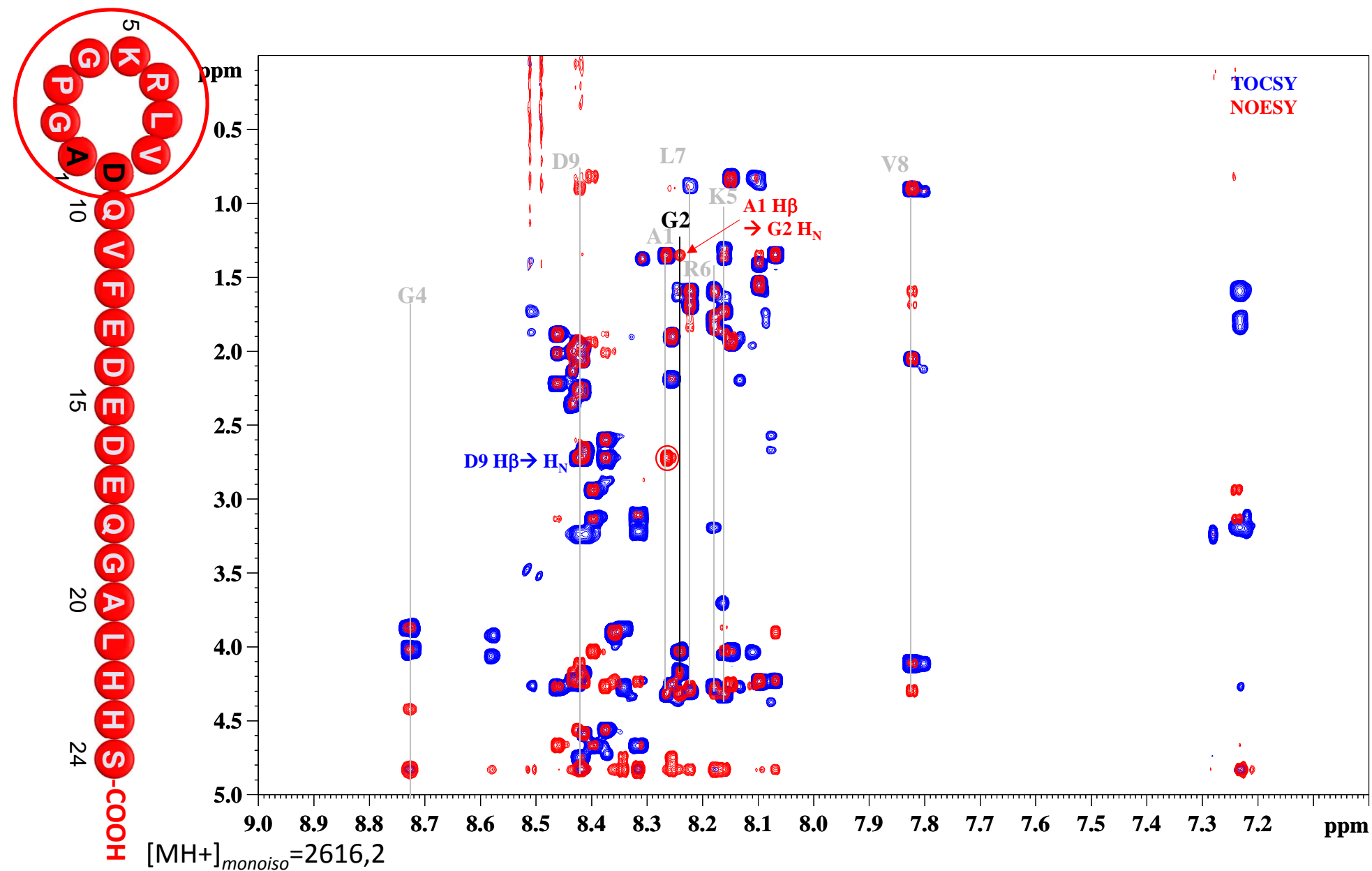


Figure S13  
wt - 10

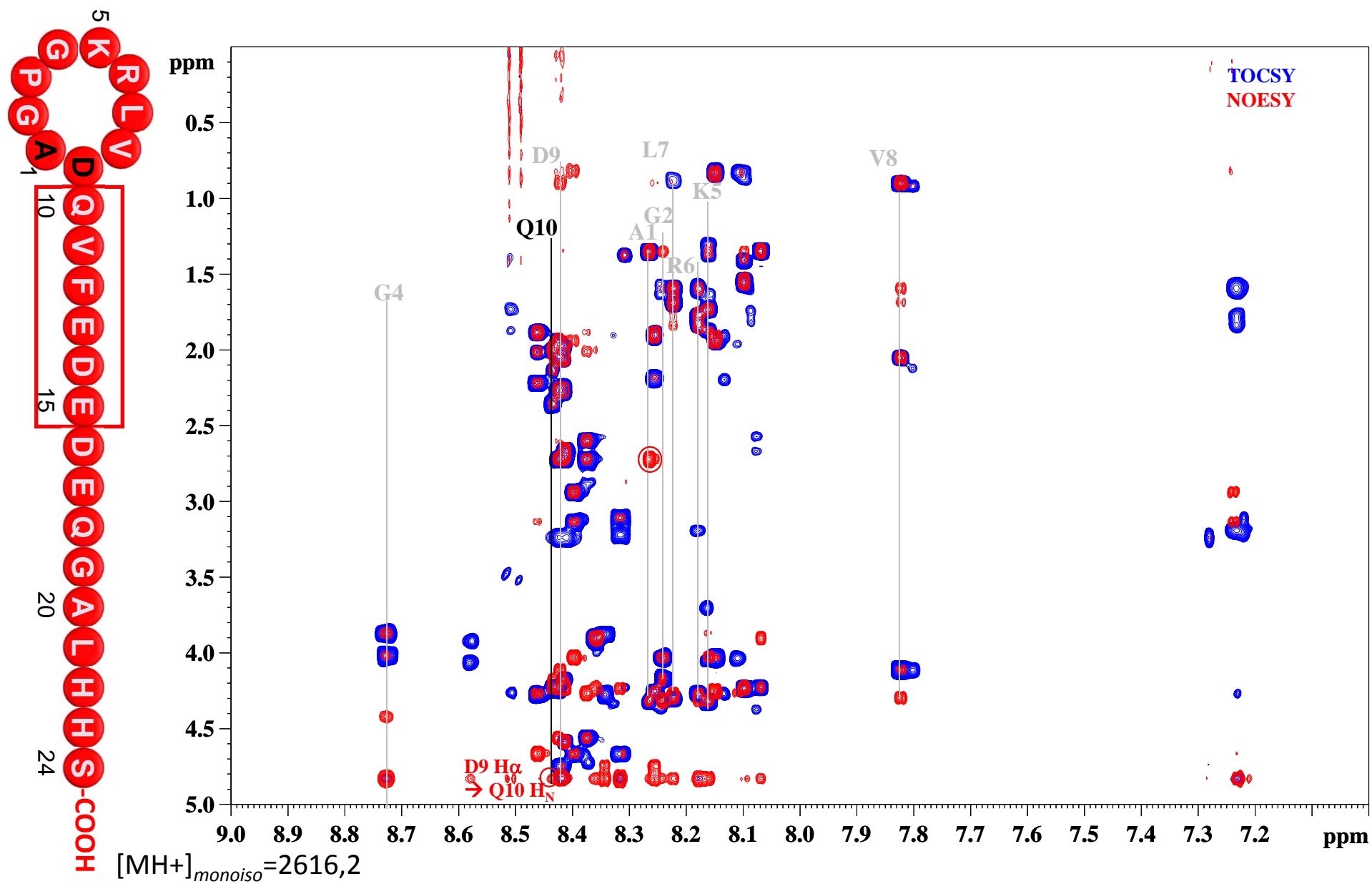
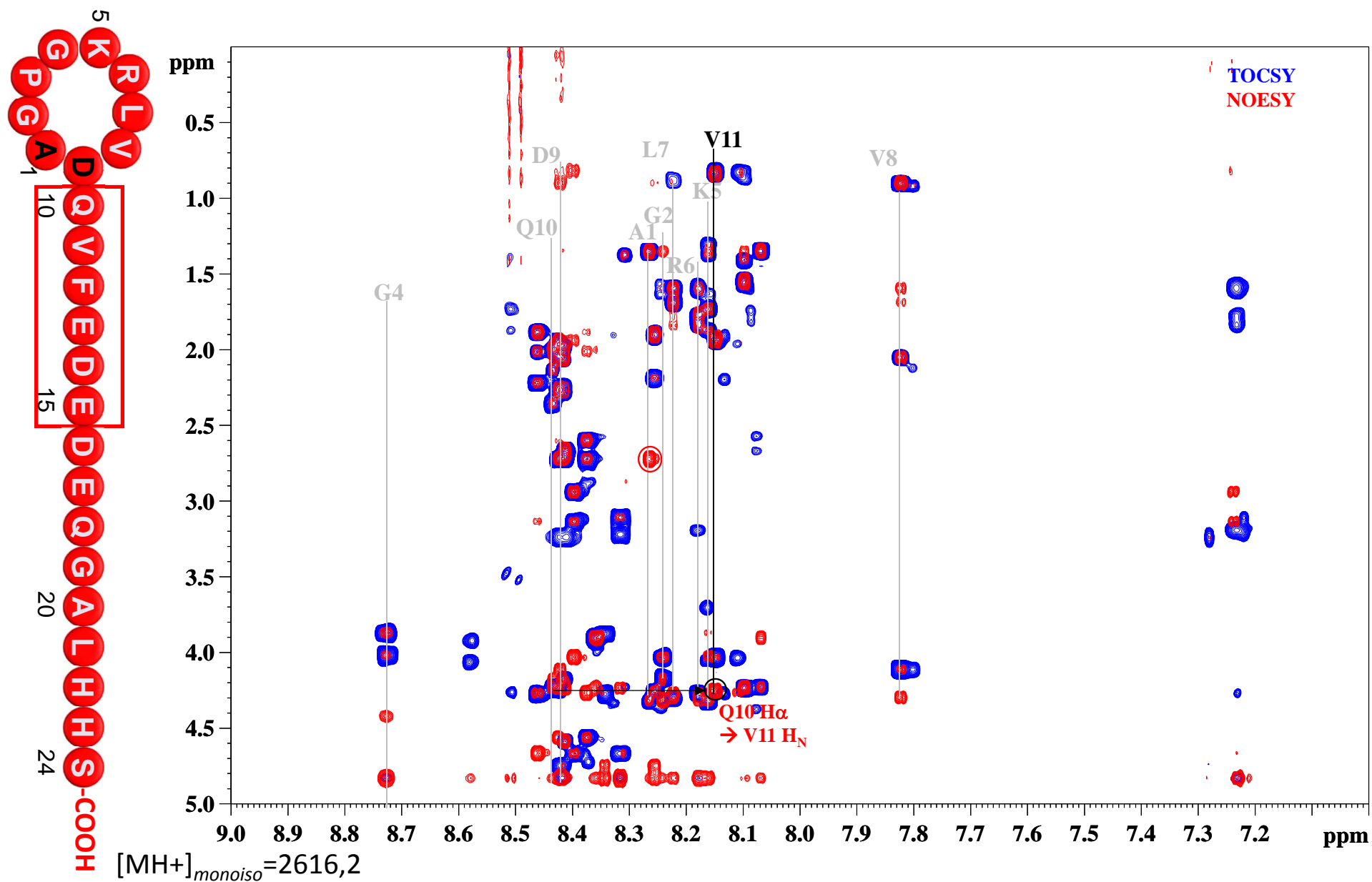


Figure S13  
 wt - 11





**Figure S13**  
 wt - 12

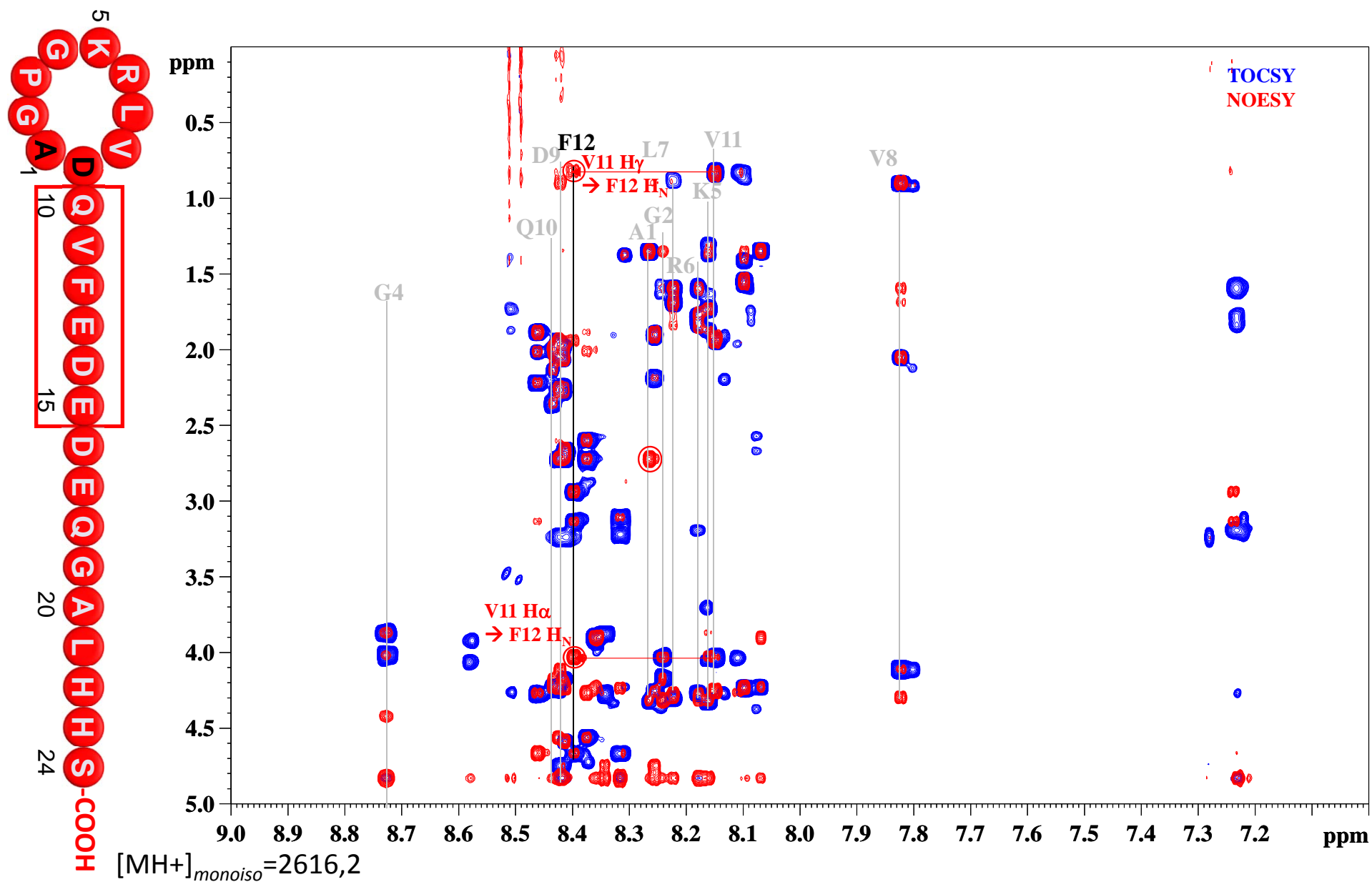


Figure S13  
wt - 13

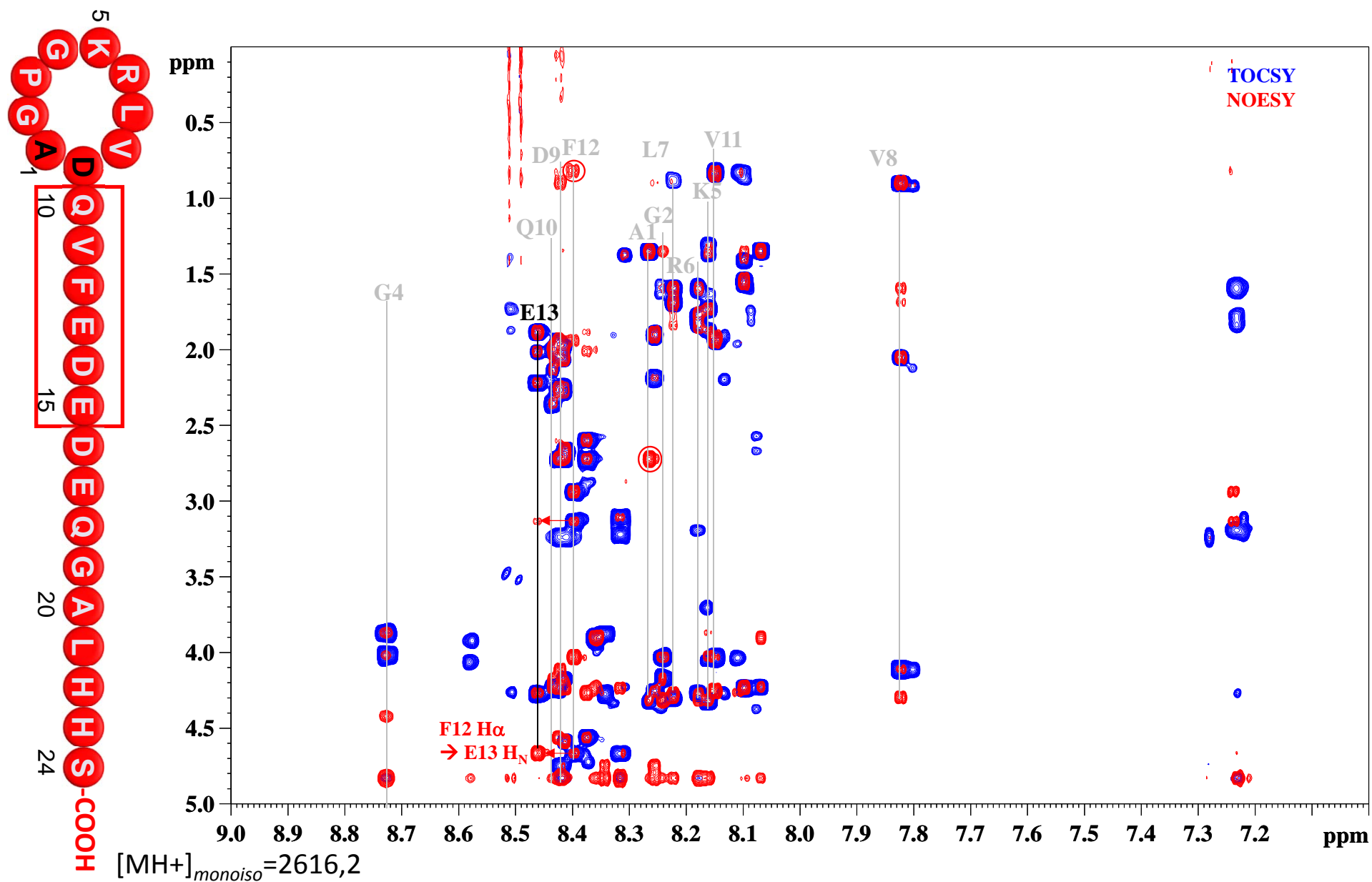


Figure S13  
 wt - 14

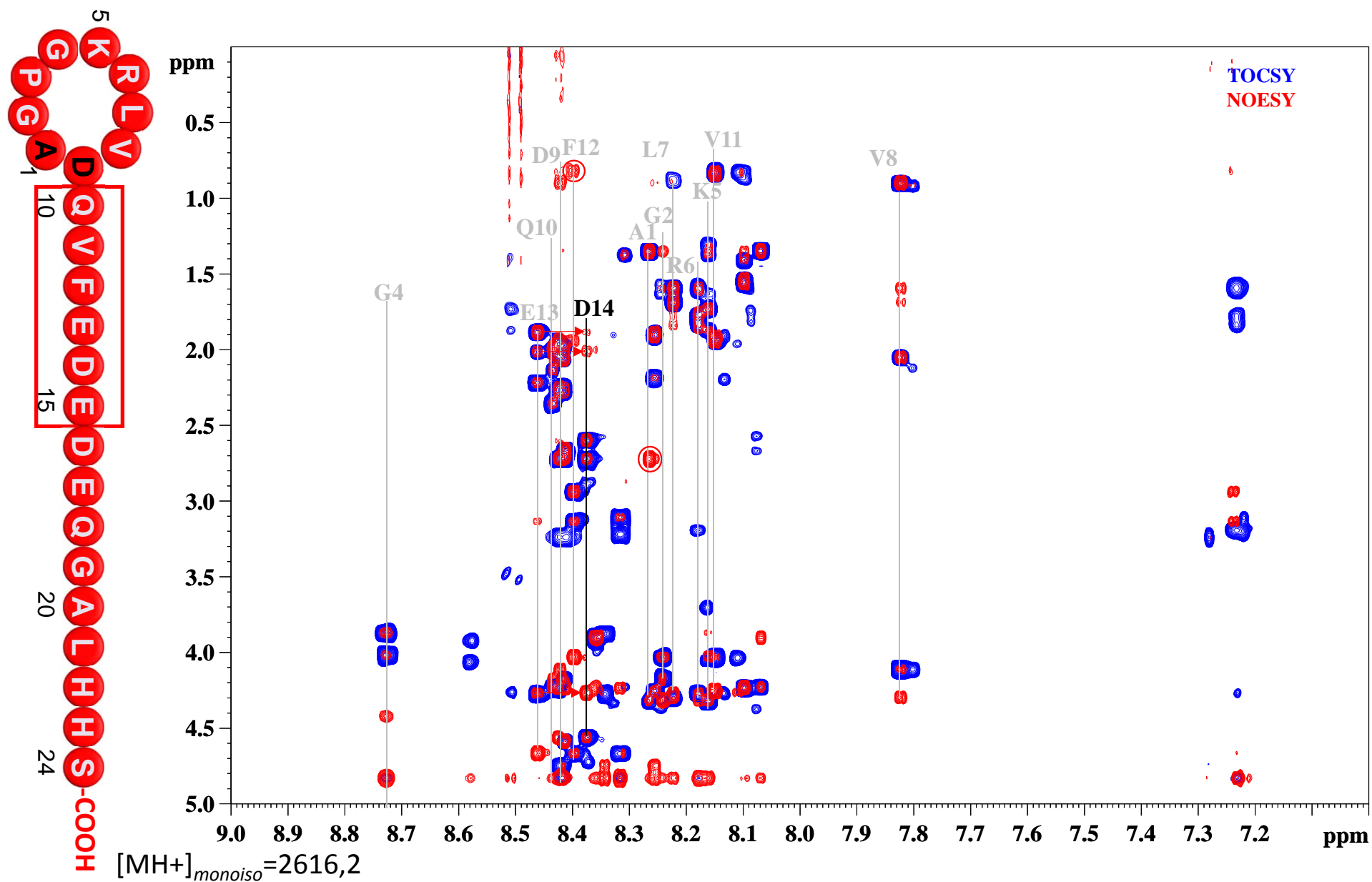
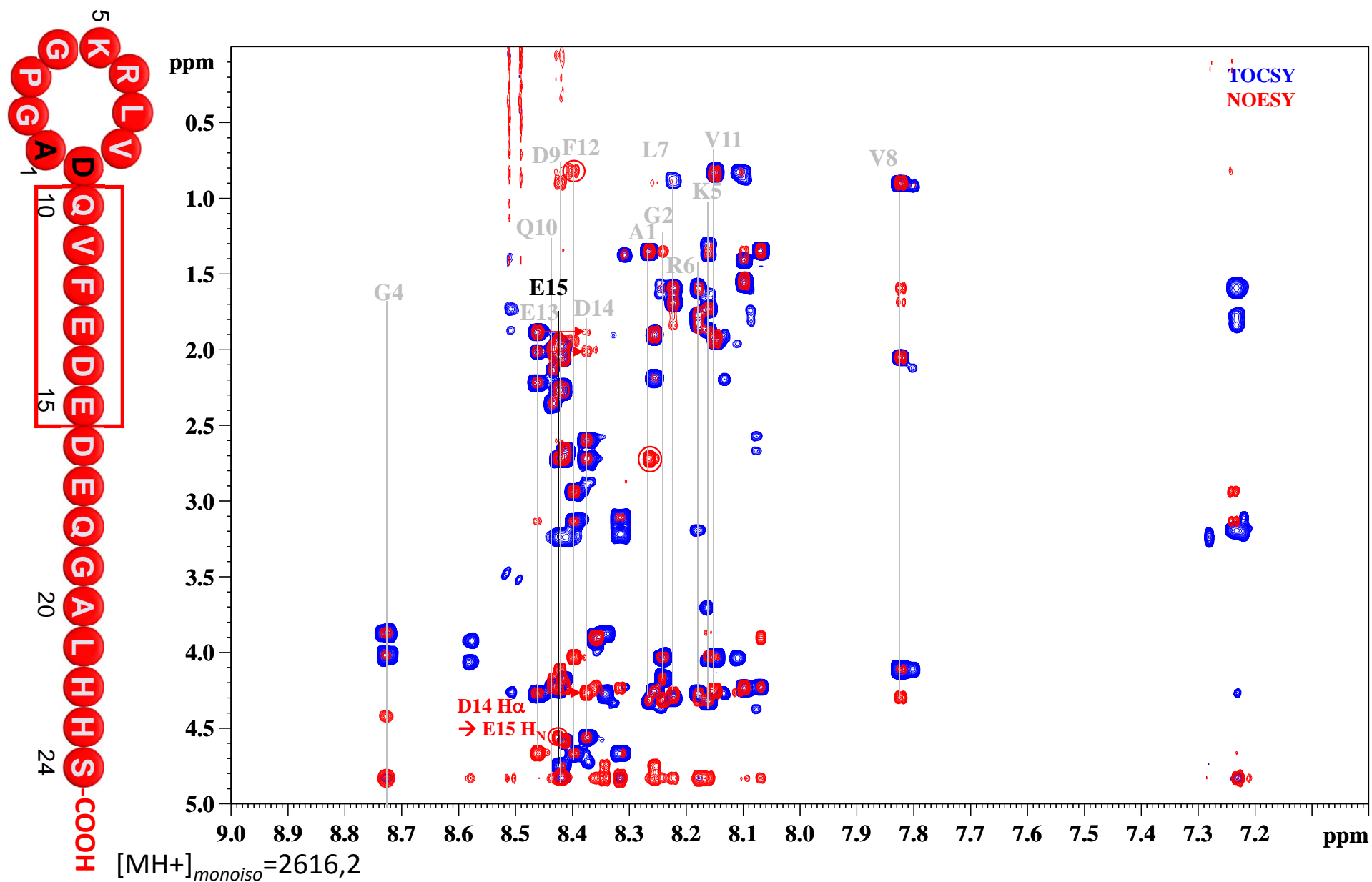


Figure S13  
wt - 15



**Figure S13**  
 wt - 16

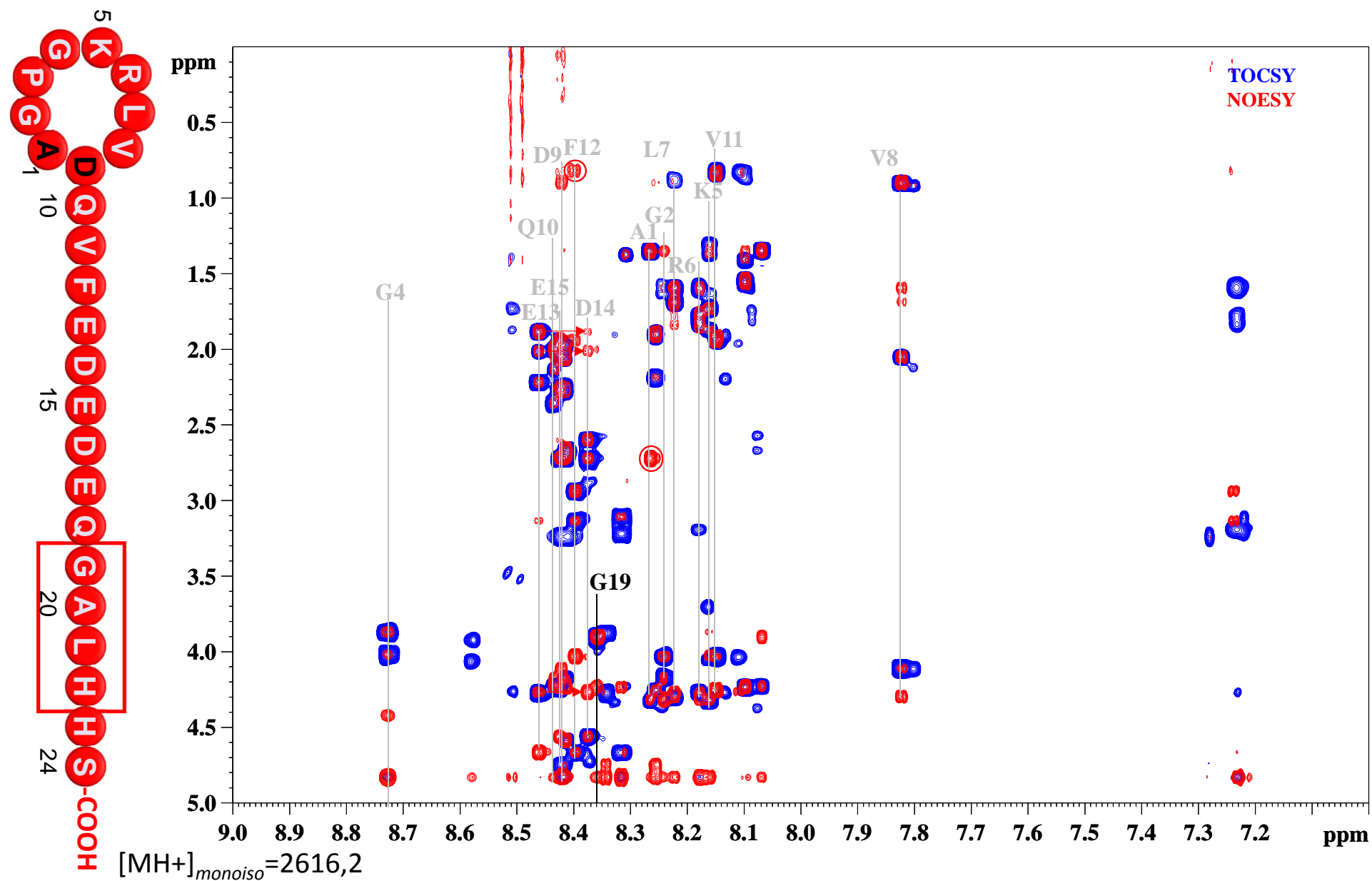


Figure S13  
 wt - 17

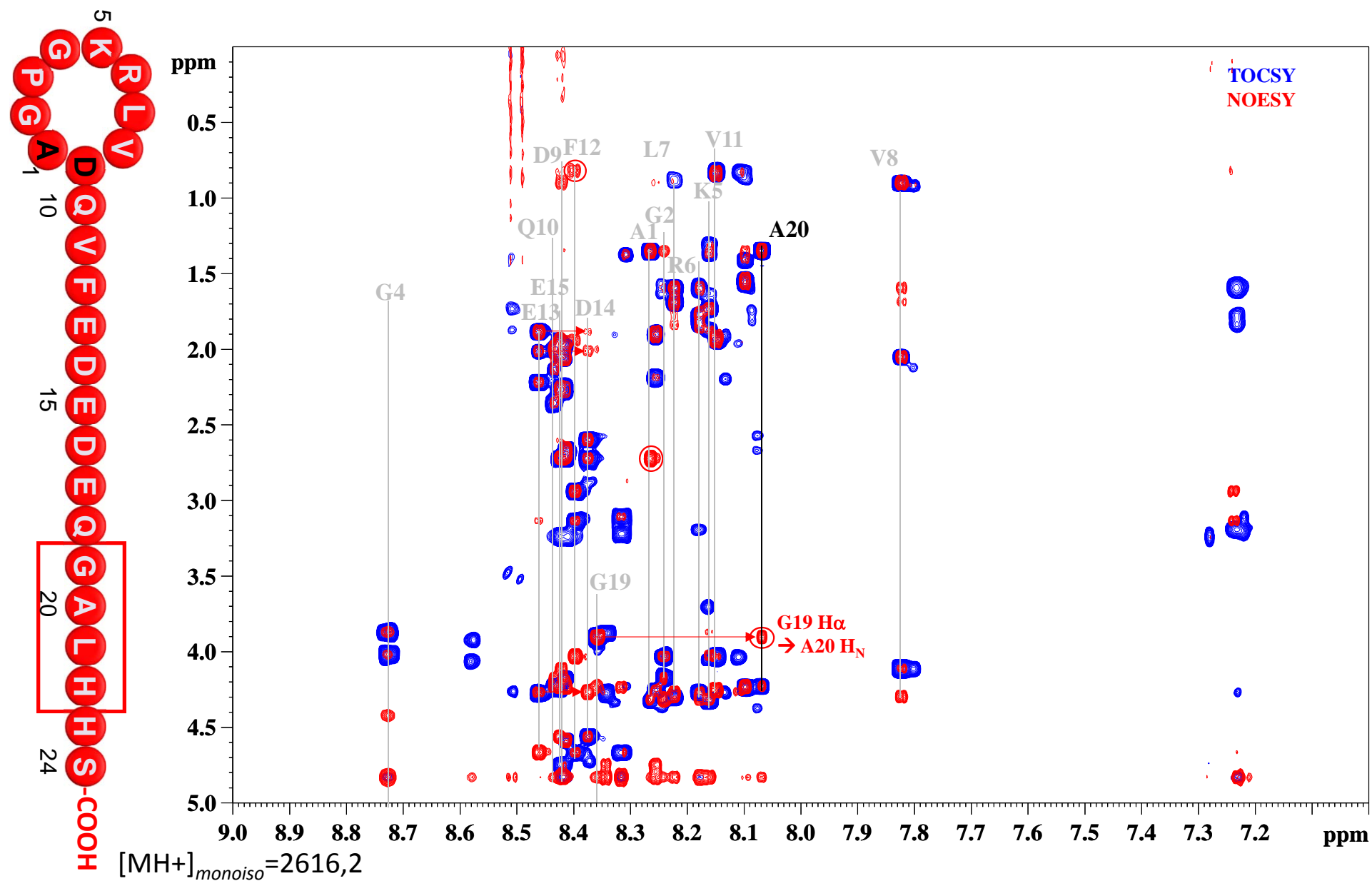


Figure S13  
 wt - 18

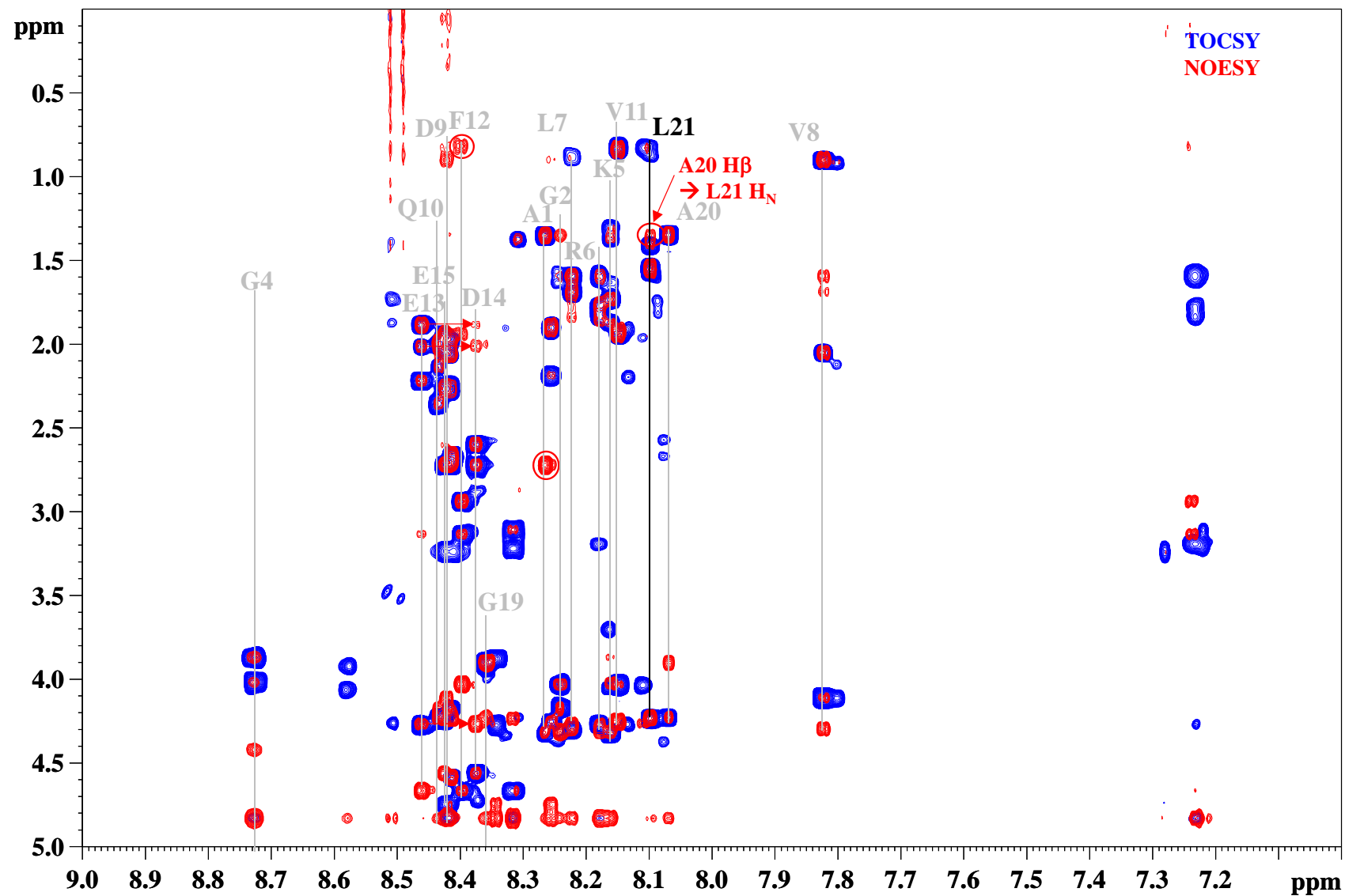


Figure S13  
wt - 19



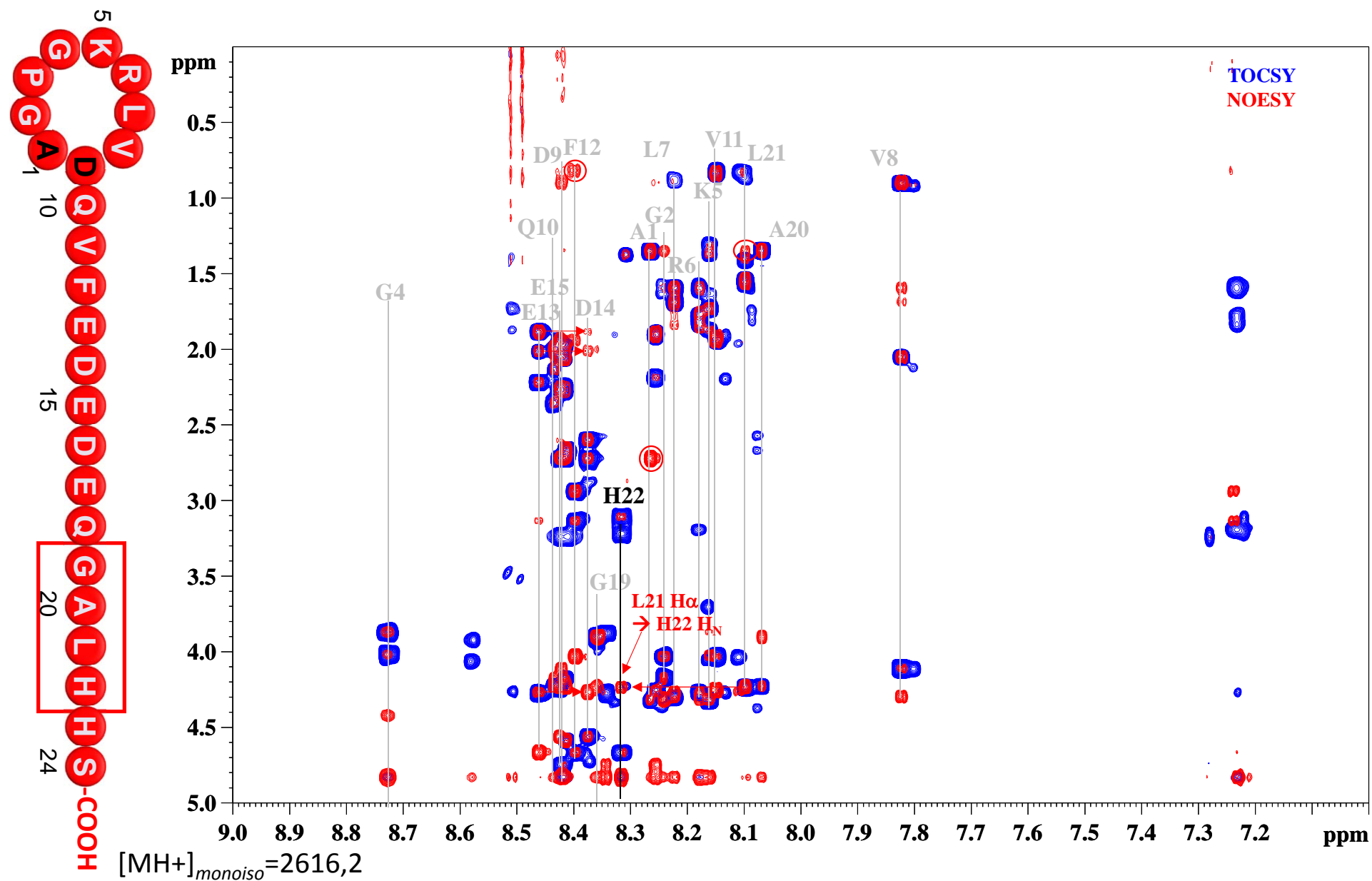


Figure S13  
 wt - 20

**Figure S14. Assigned homonuclear TOCSY (blue) and NOESY (red) spectra of L21F pseudomycoidin.**

Sample was a ~3mM solution of peptide dissolved in 200µl of phosphate buffer 100mM pH 6.5. All spectra were acquired on a 800MHz Bruker Avance II spectrometer (Bruker, Fällanden, Switzerland) equipped with a cryogenic triple resonance probe head and operating at 293K. NOESY spectrum was acquired with a mixing time of 200ms, as a complex matrix of 4096 x 400 points, with 16 scans per increment. TOCSY spectrum was acquired with a 70ms mixing time implemented as a dipsi2 pulse train, and was equally acquired as a complex matrix of 4096 x 400 points, with 16 scans per increment. Spectra were Fourier transformed after apodization with a shifted square sine bell in both dimensions using the Topspin3.0 software package (Bruker, Fällanden, Switzerland). Assignment were done as described by Wüthrich (K. Wüthrich, NMR of proteins and Nucleic Acids, Wiley, 1986). The following panels demonstrate the assignment procedure by highlighting one residue per panel. Characteristic NOEs that were used for the assignment are circled, and explicit assignment is given.

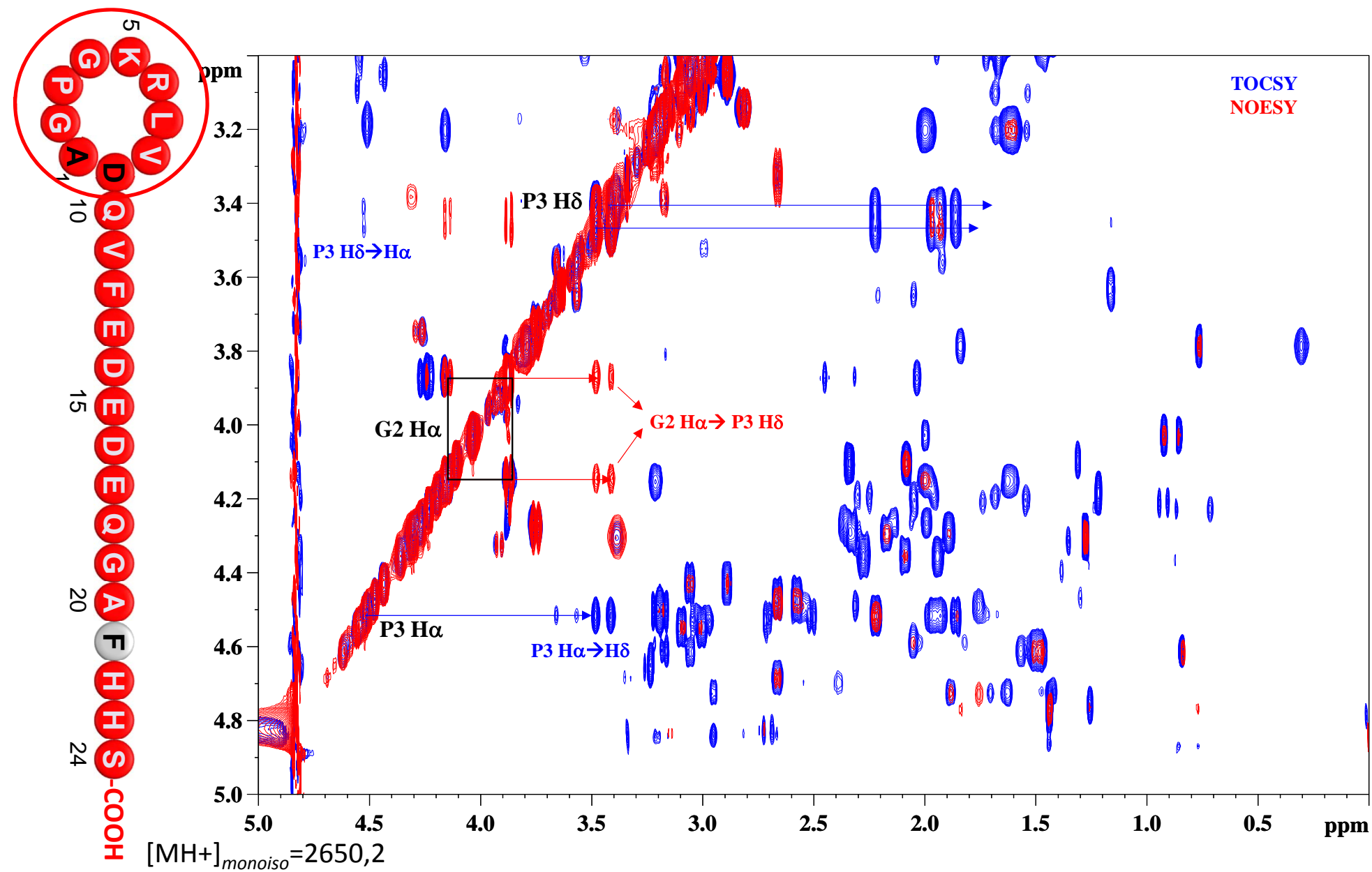


Figure S14  
L21F - 1

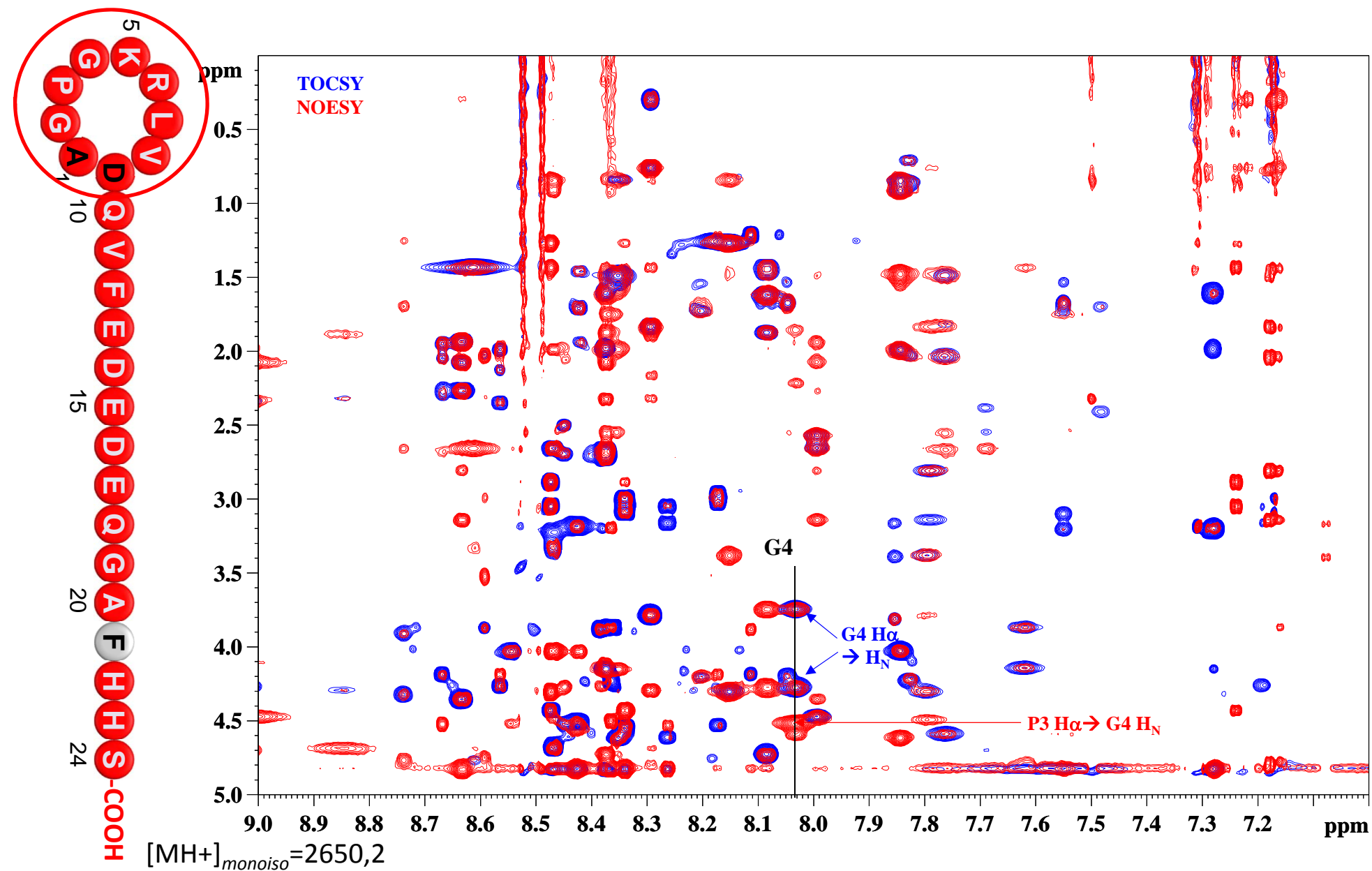


Figure S14  
L21F - 2

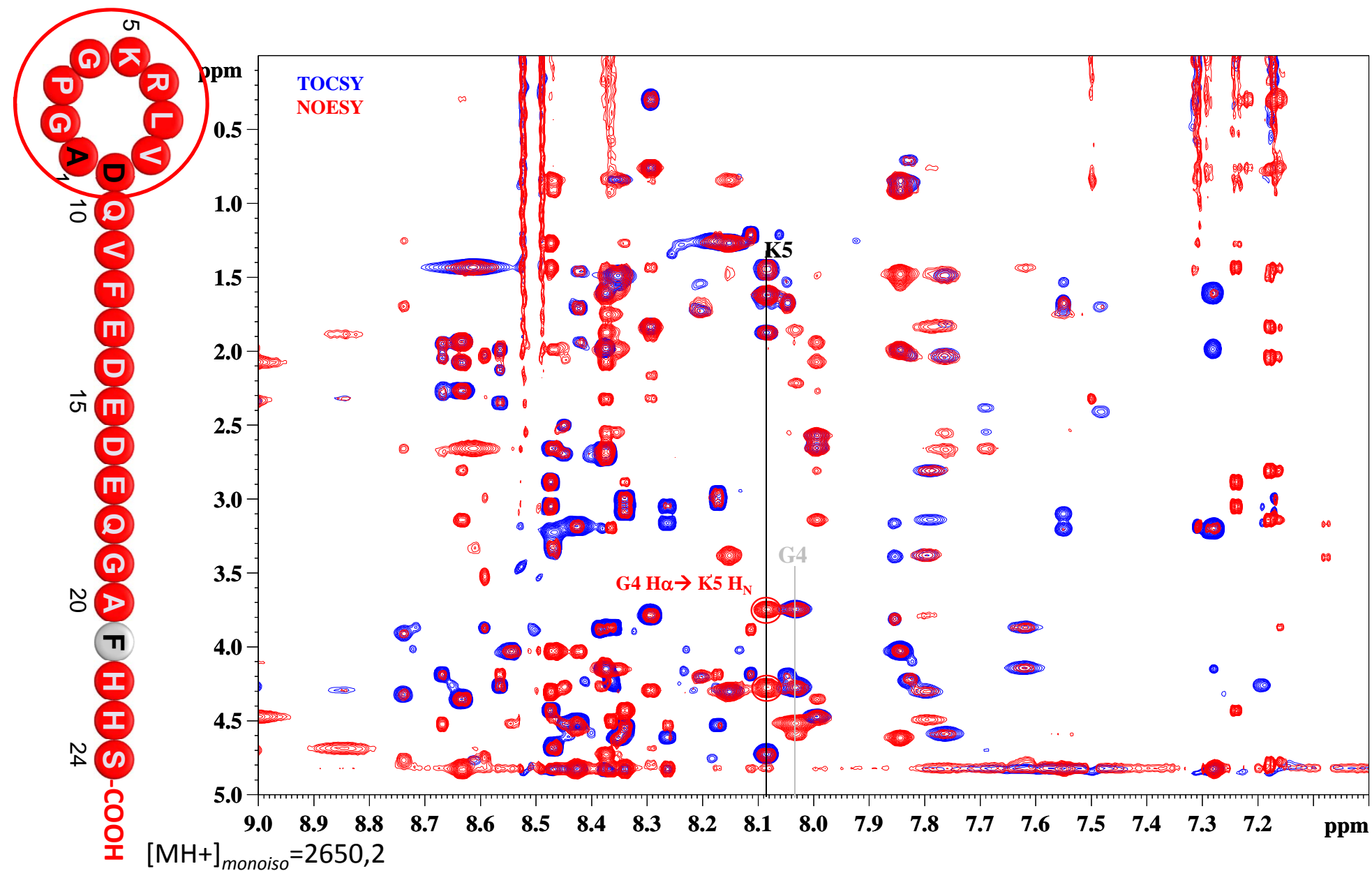


Figure S14  
L21F - 3

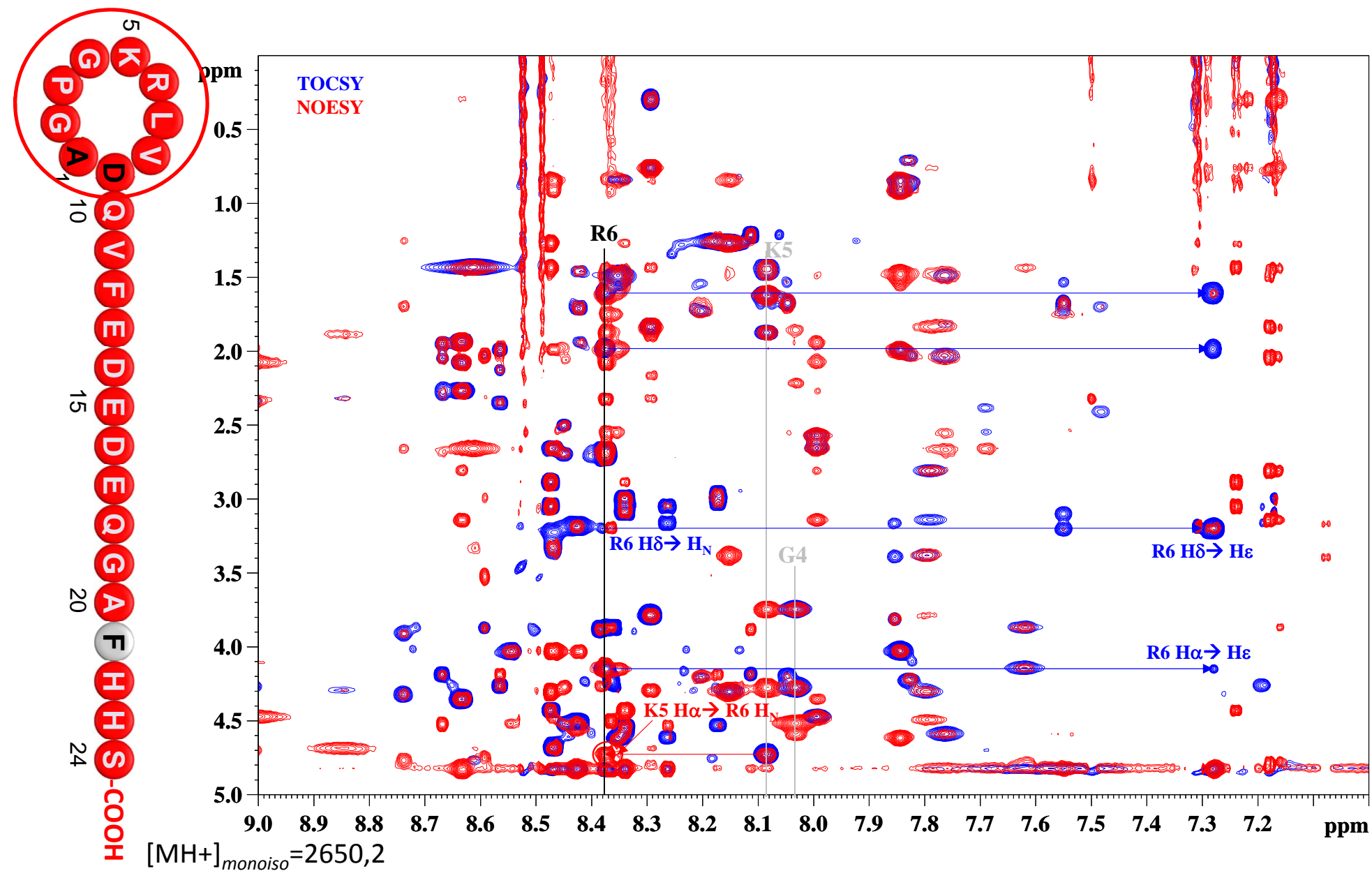


Figure S14  
L21F - 4

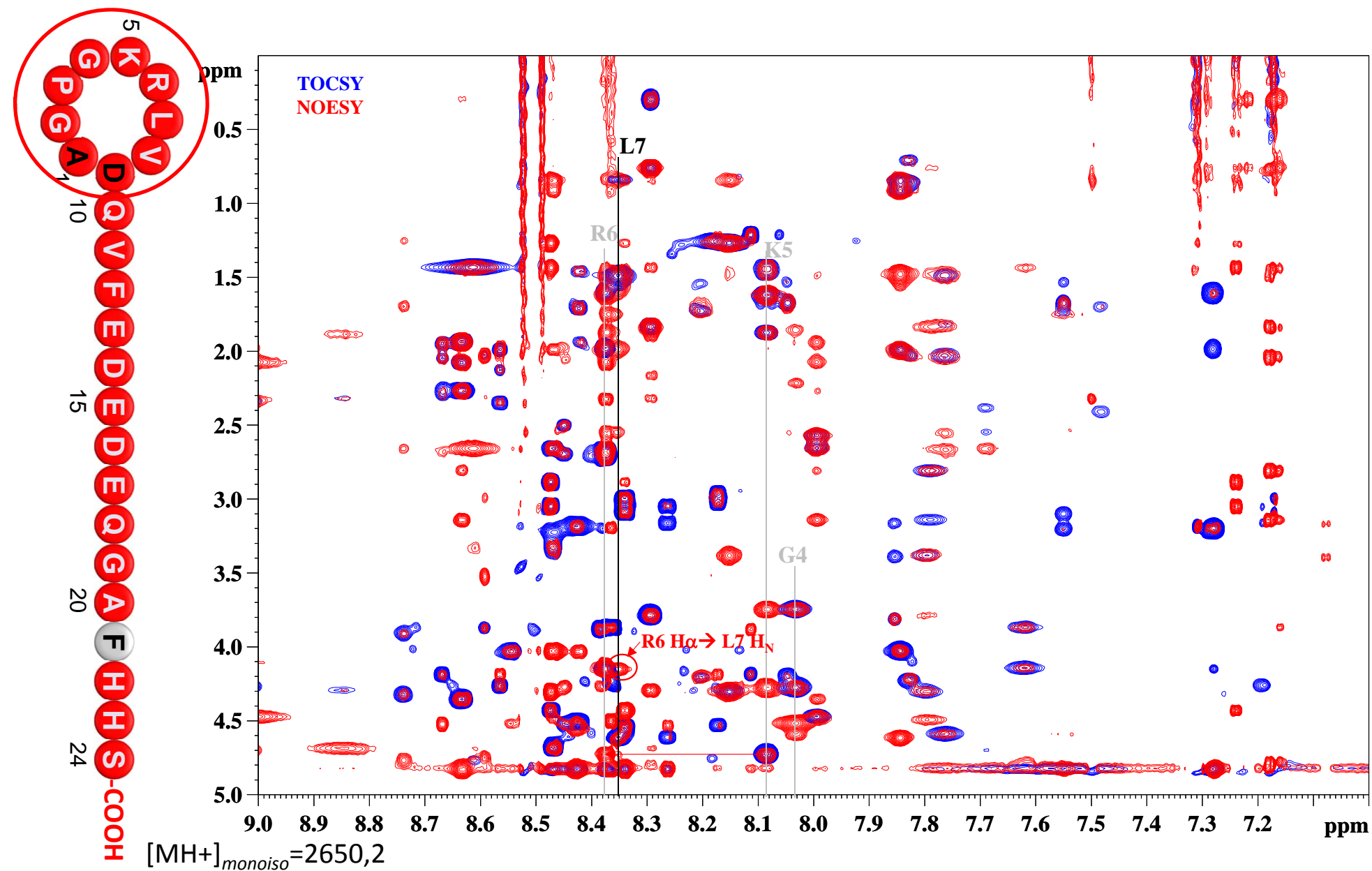


Figure S14  
L21F - 5



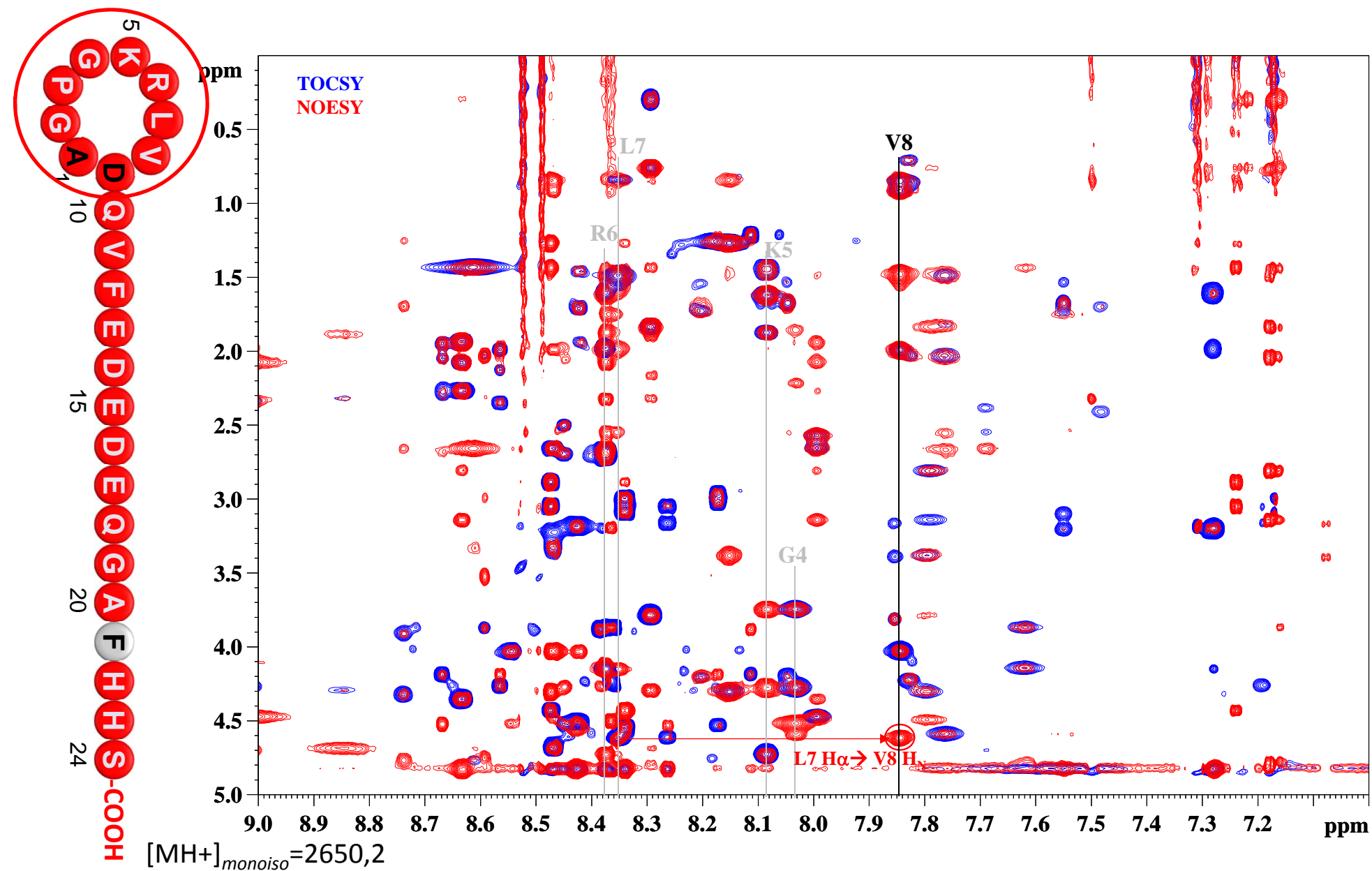
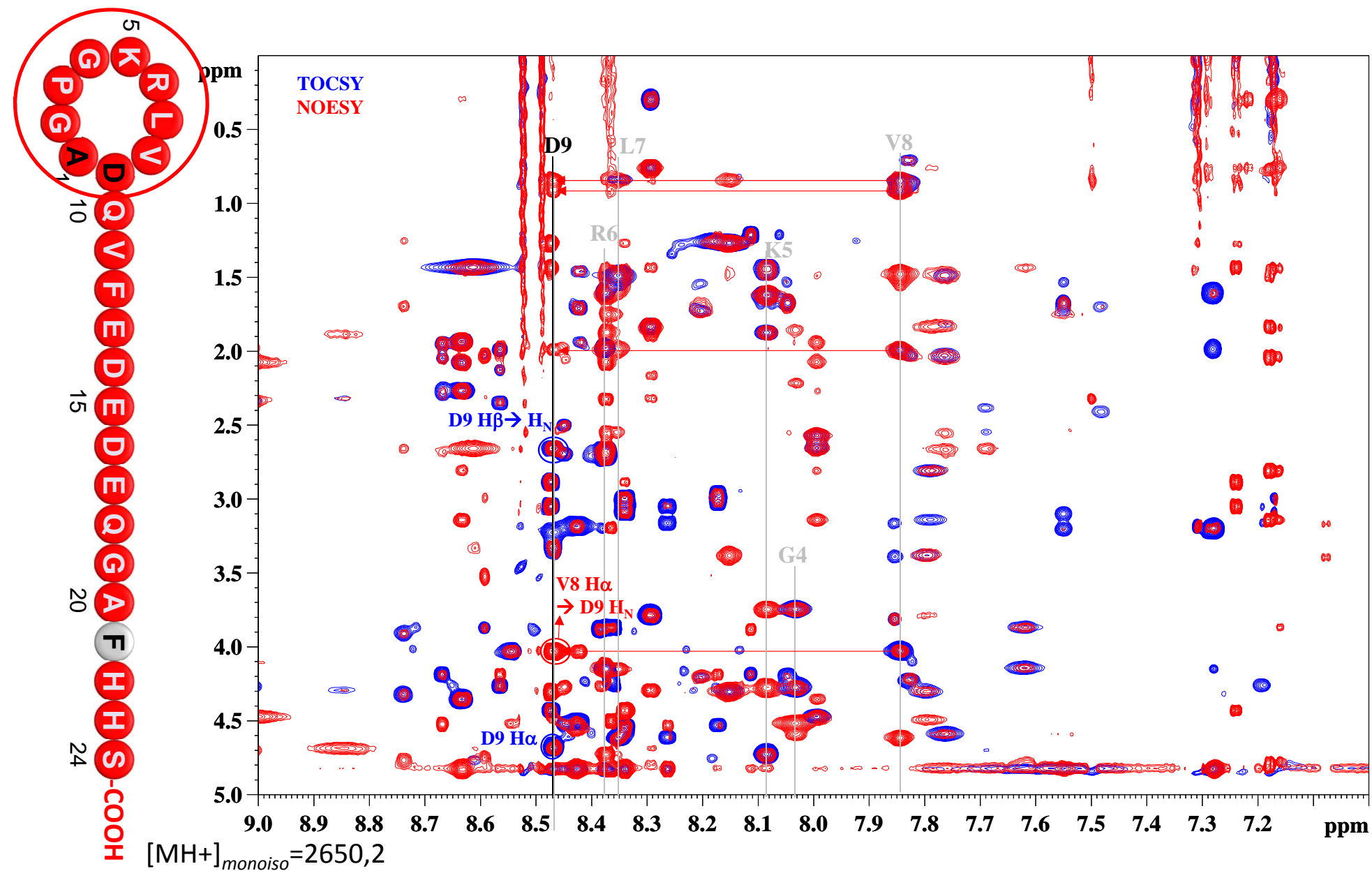


Figure S14  
L21F - 6





**Figure S14**  
**L21F - 7**

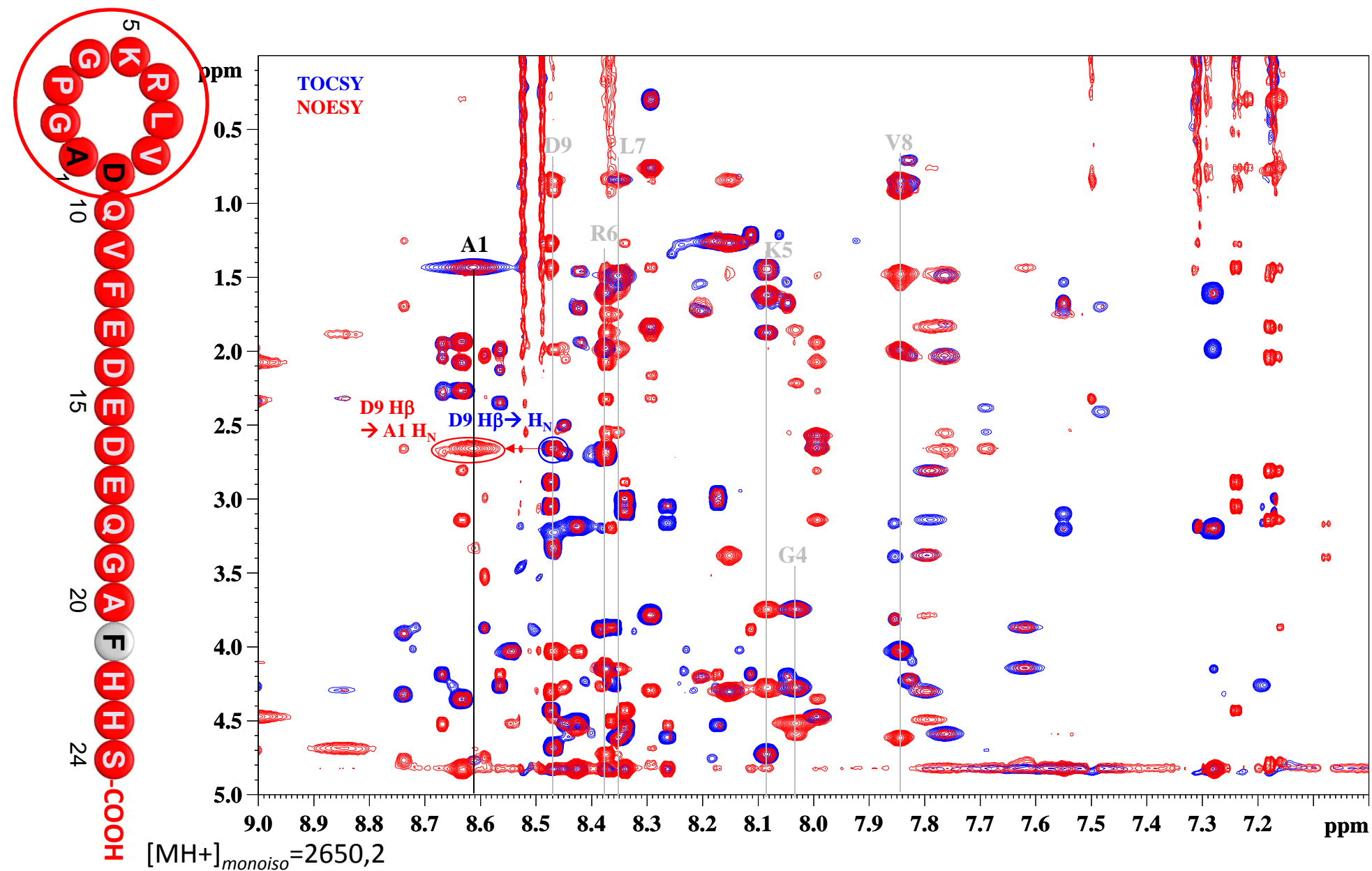


Figure S14  
L21F - 8



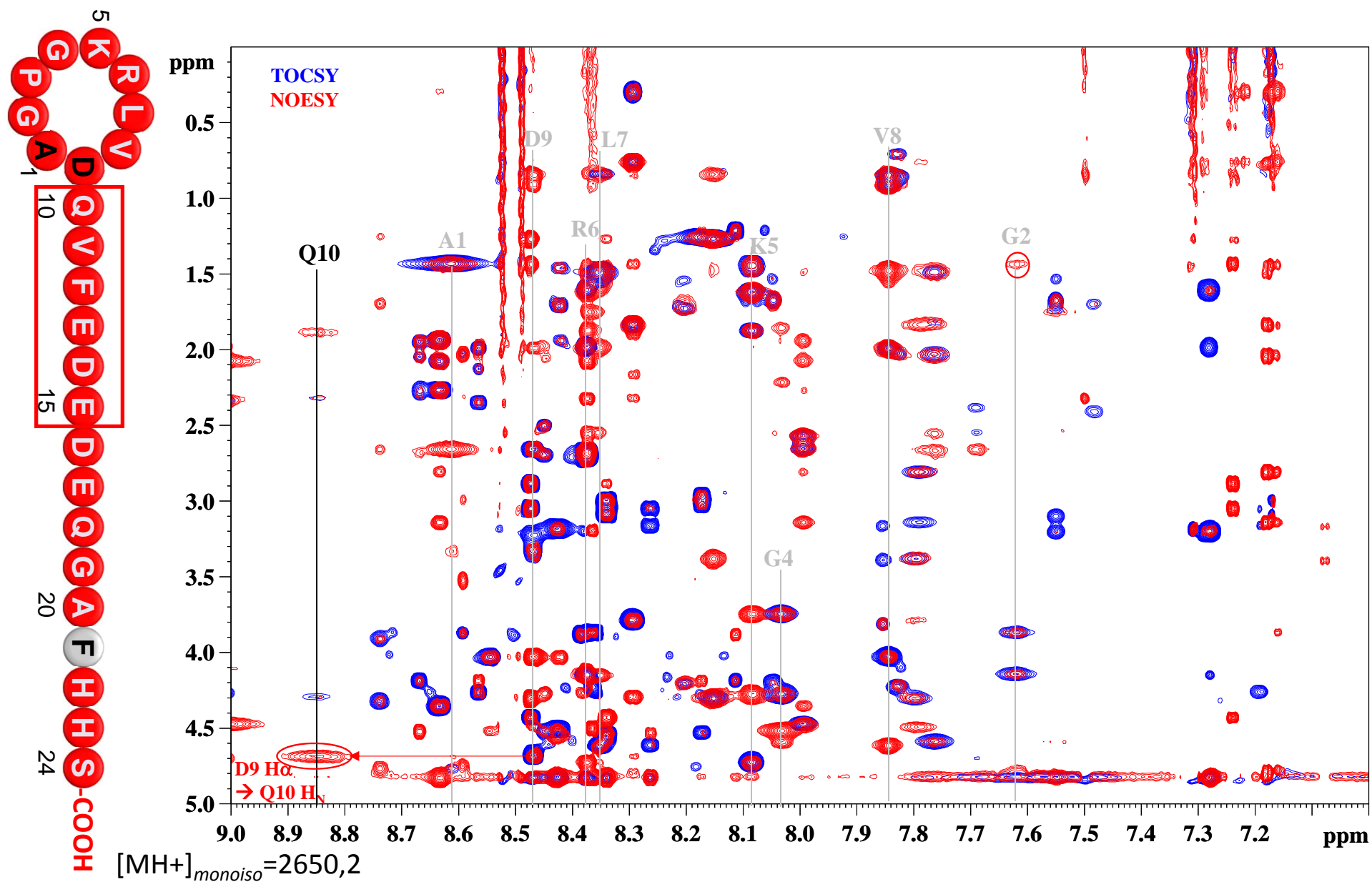


Figure S14  
 L21F - 10

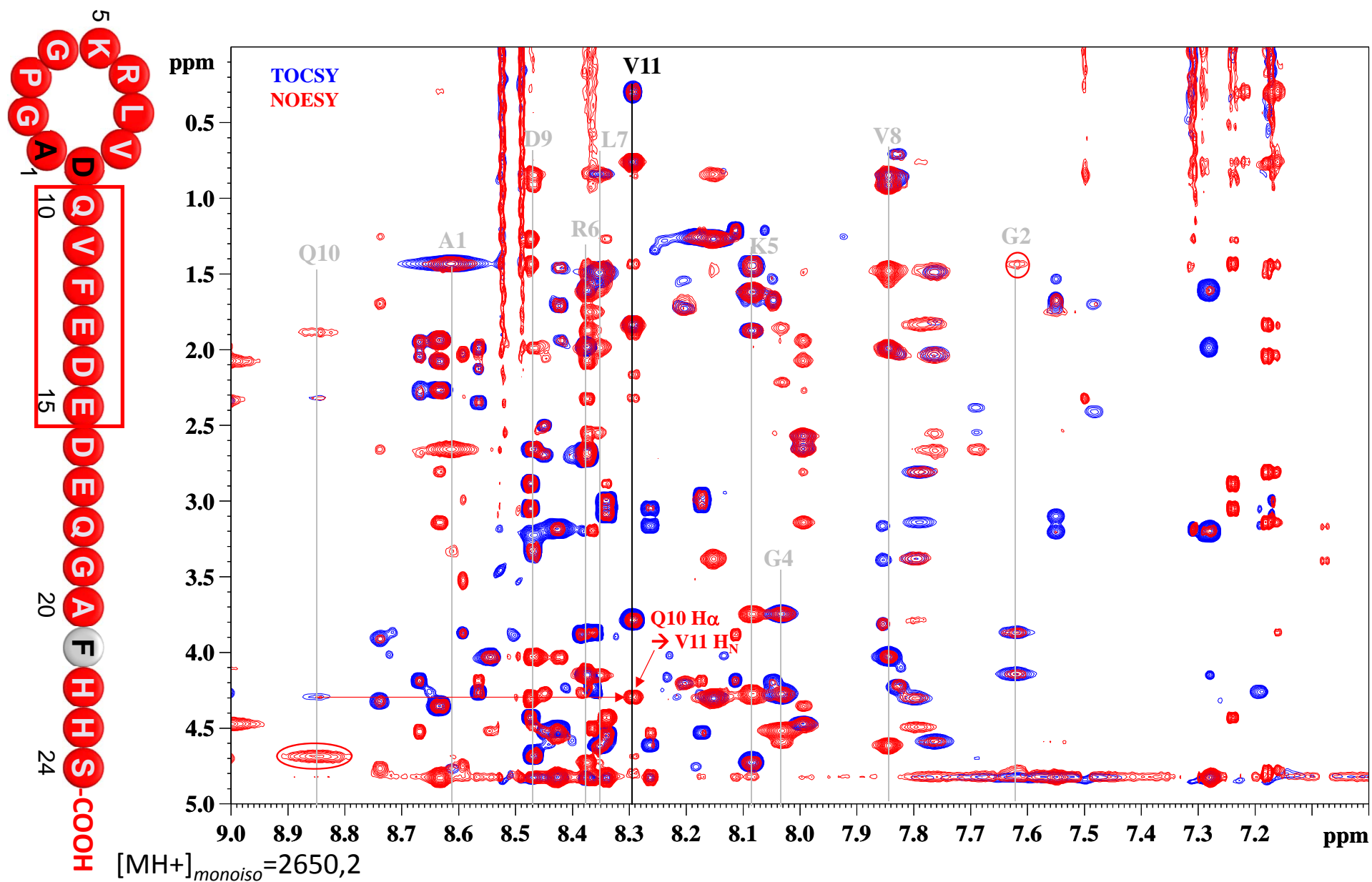


Figure S14  
 L21F - 11

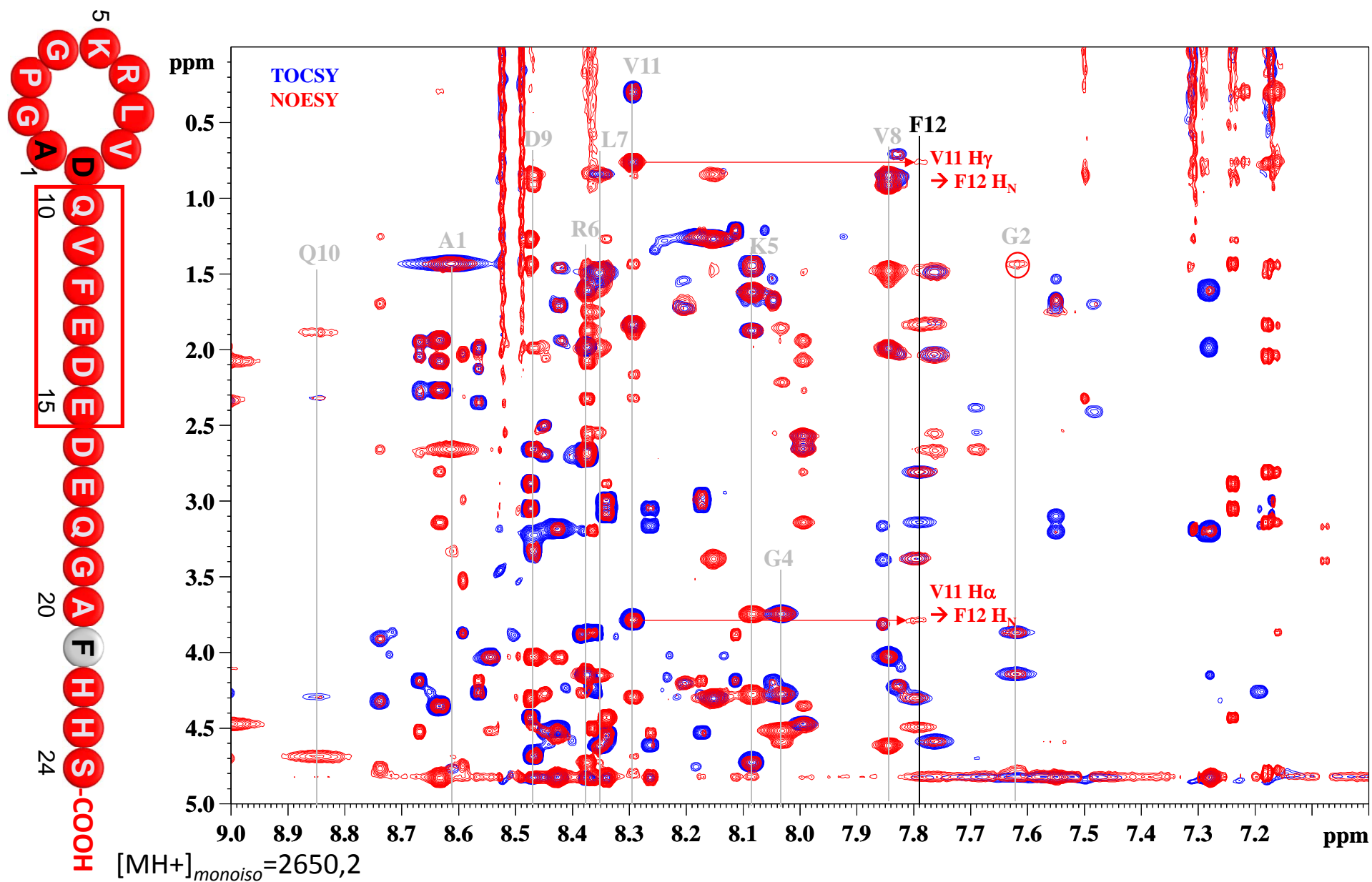


Figure S14  
 L21F - 12



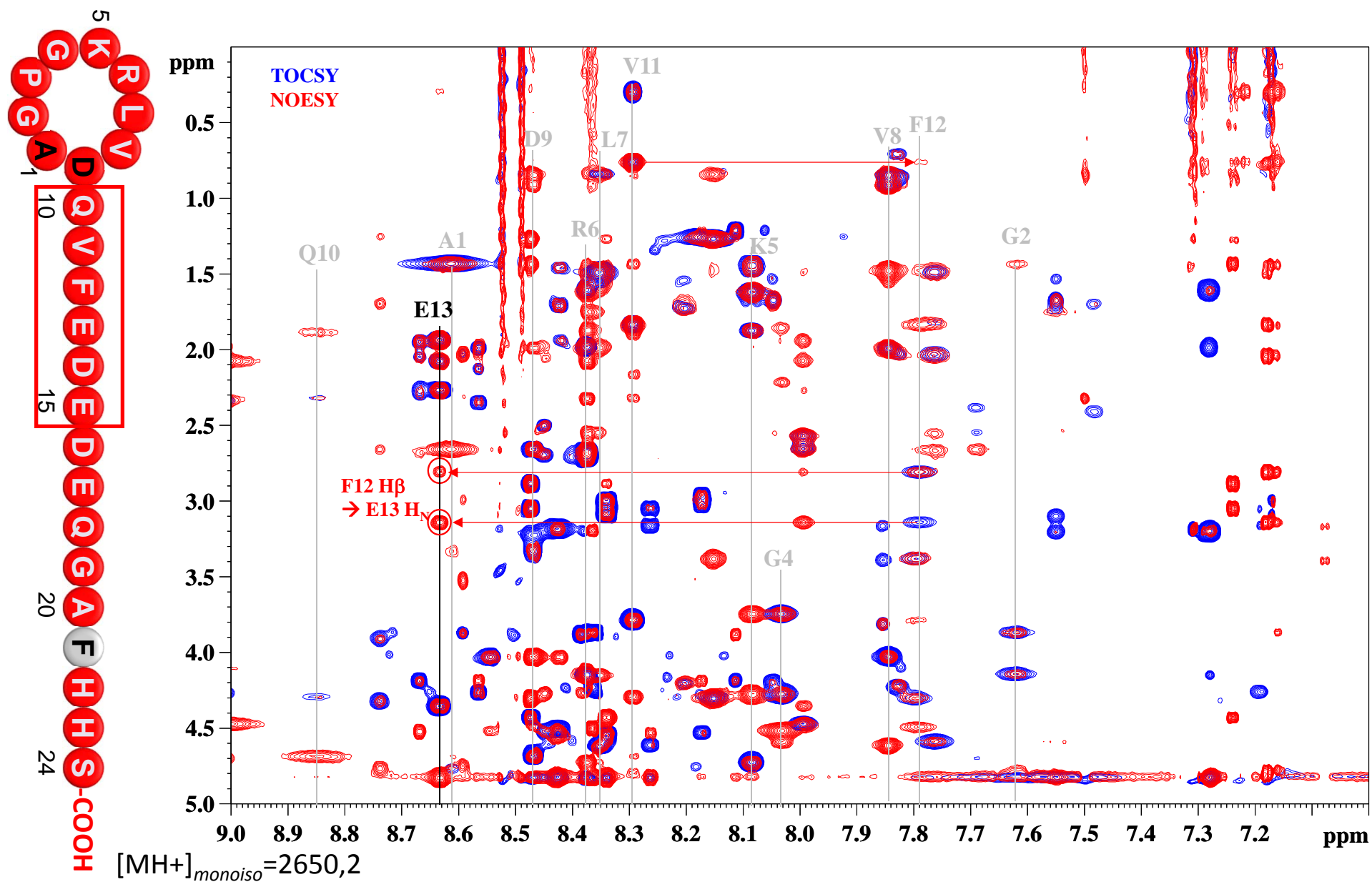


Figure S14  
 L21F - 13

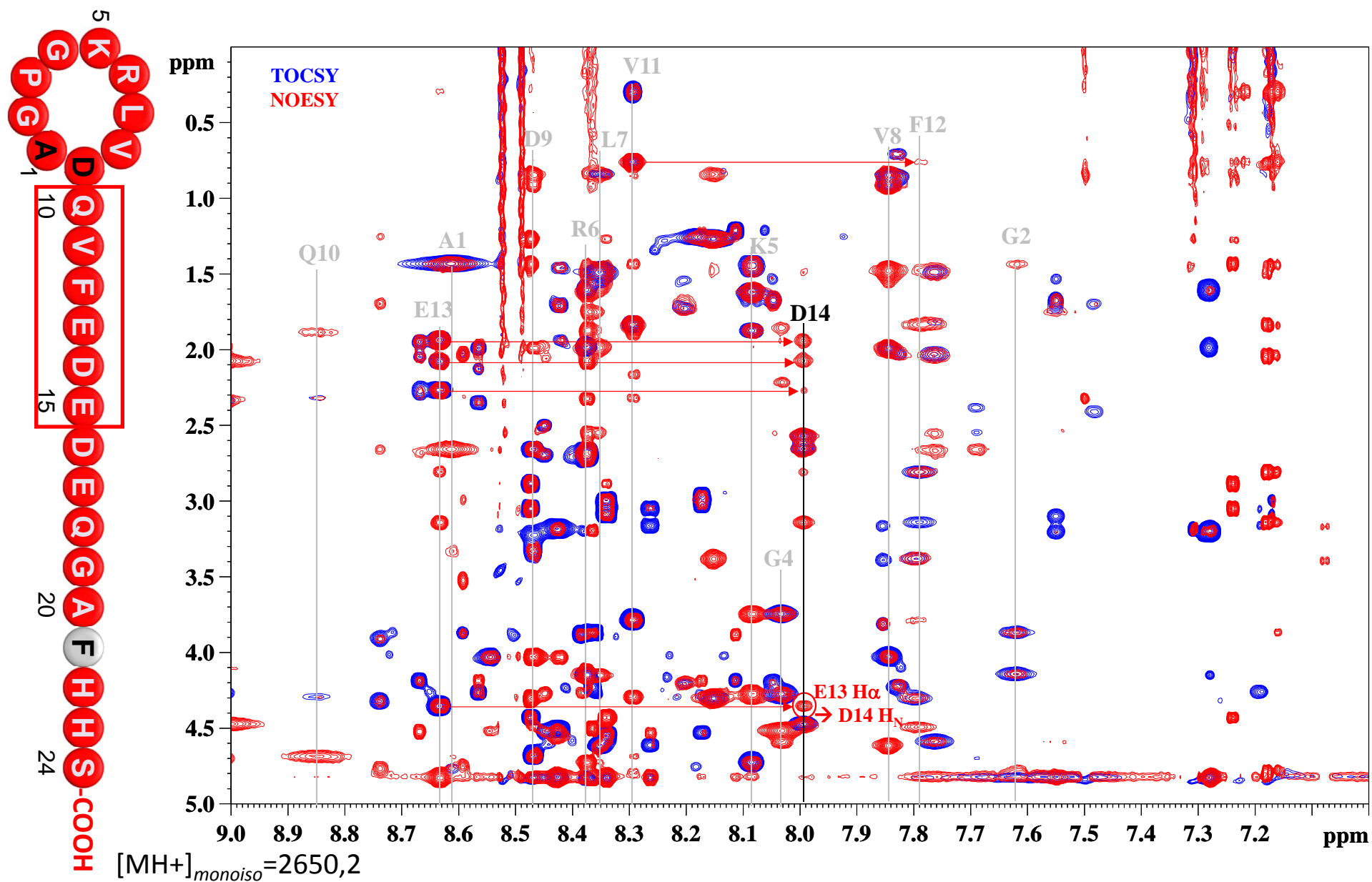


Figure S14  
 L21F - 14



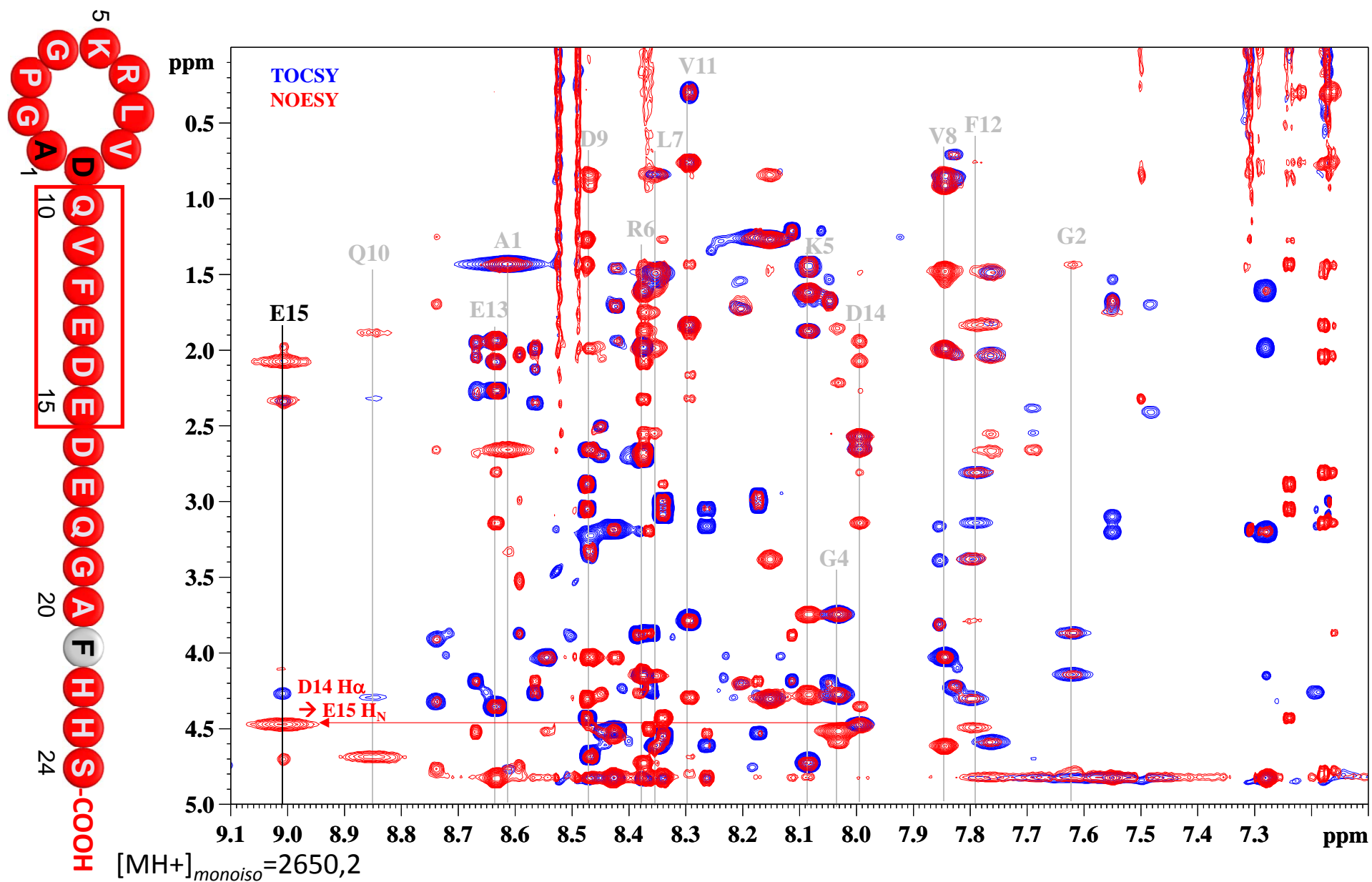


Figure S14  
 L21F - 15

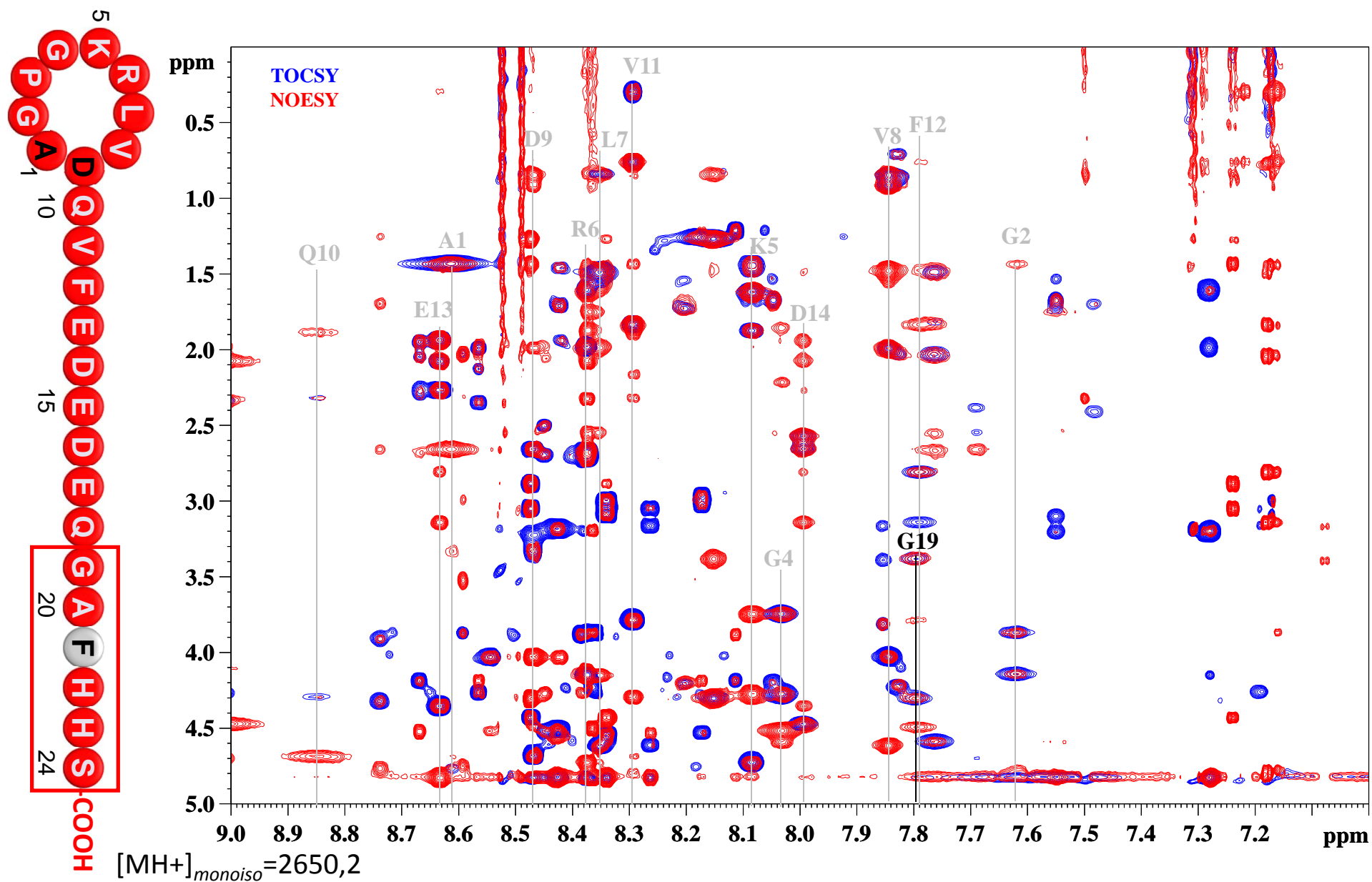


Figure S14  
L21F - 16

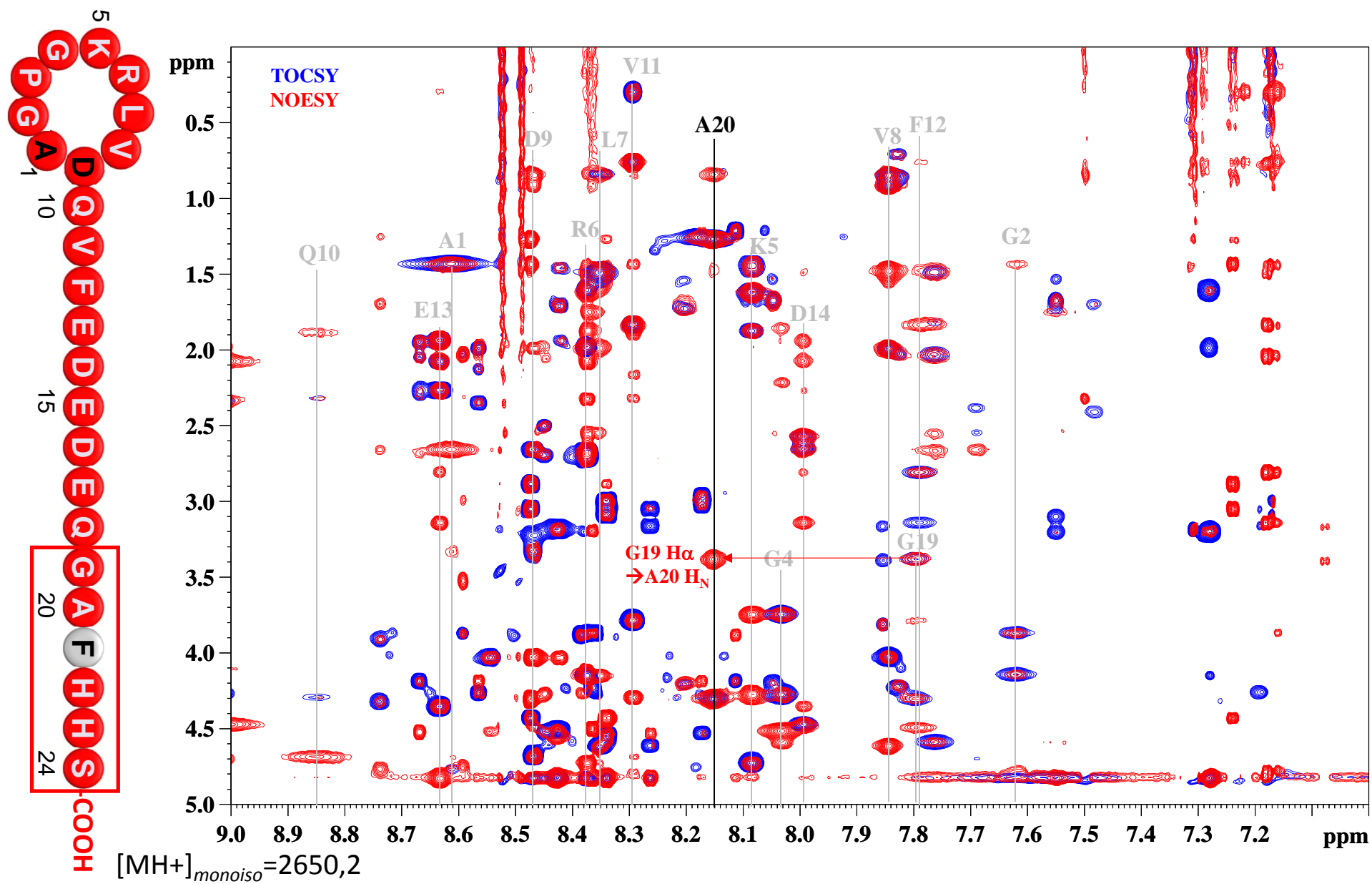


Figure S14  
L21F - 17

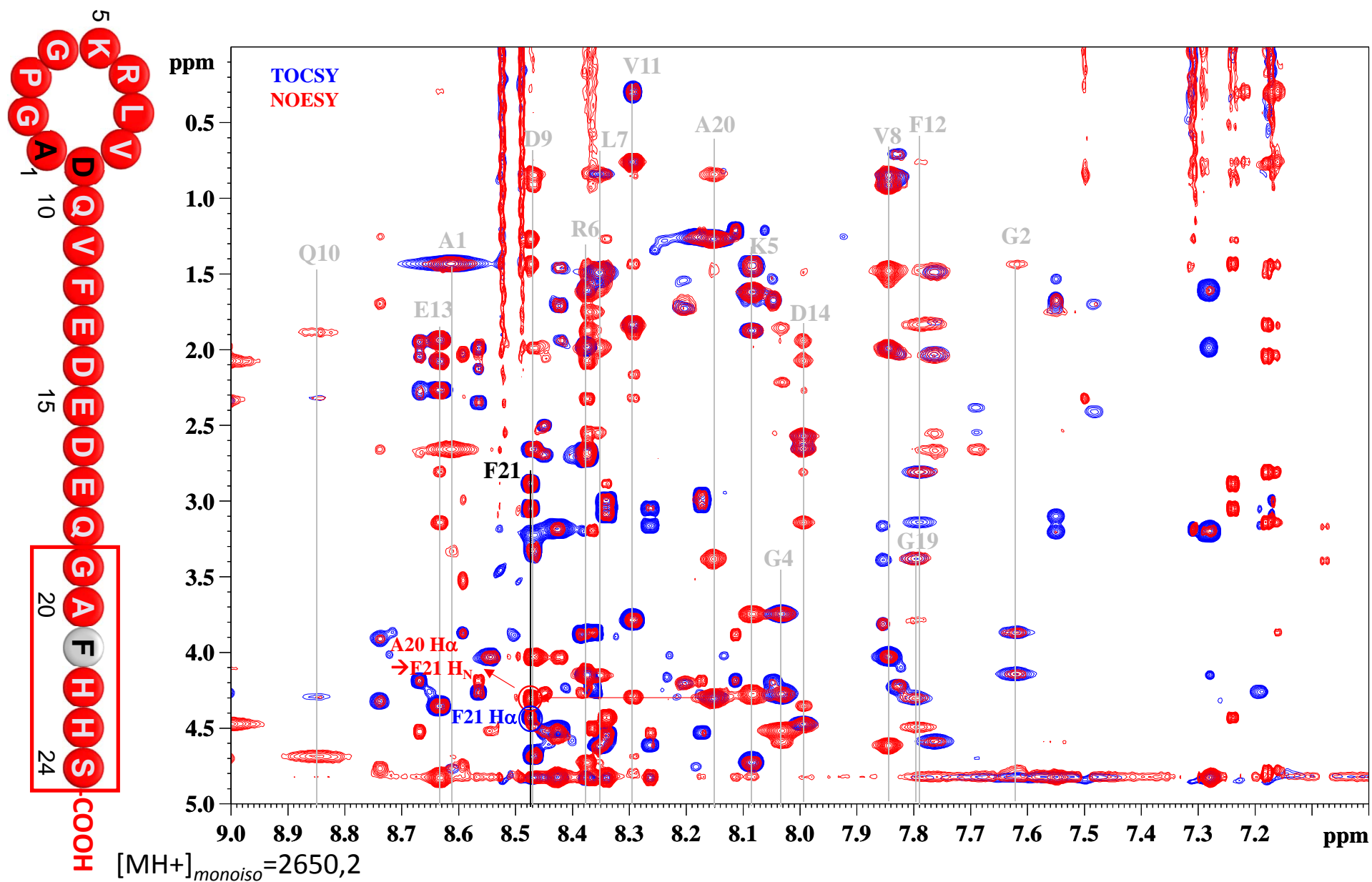


Figure S14  
L21F - 18

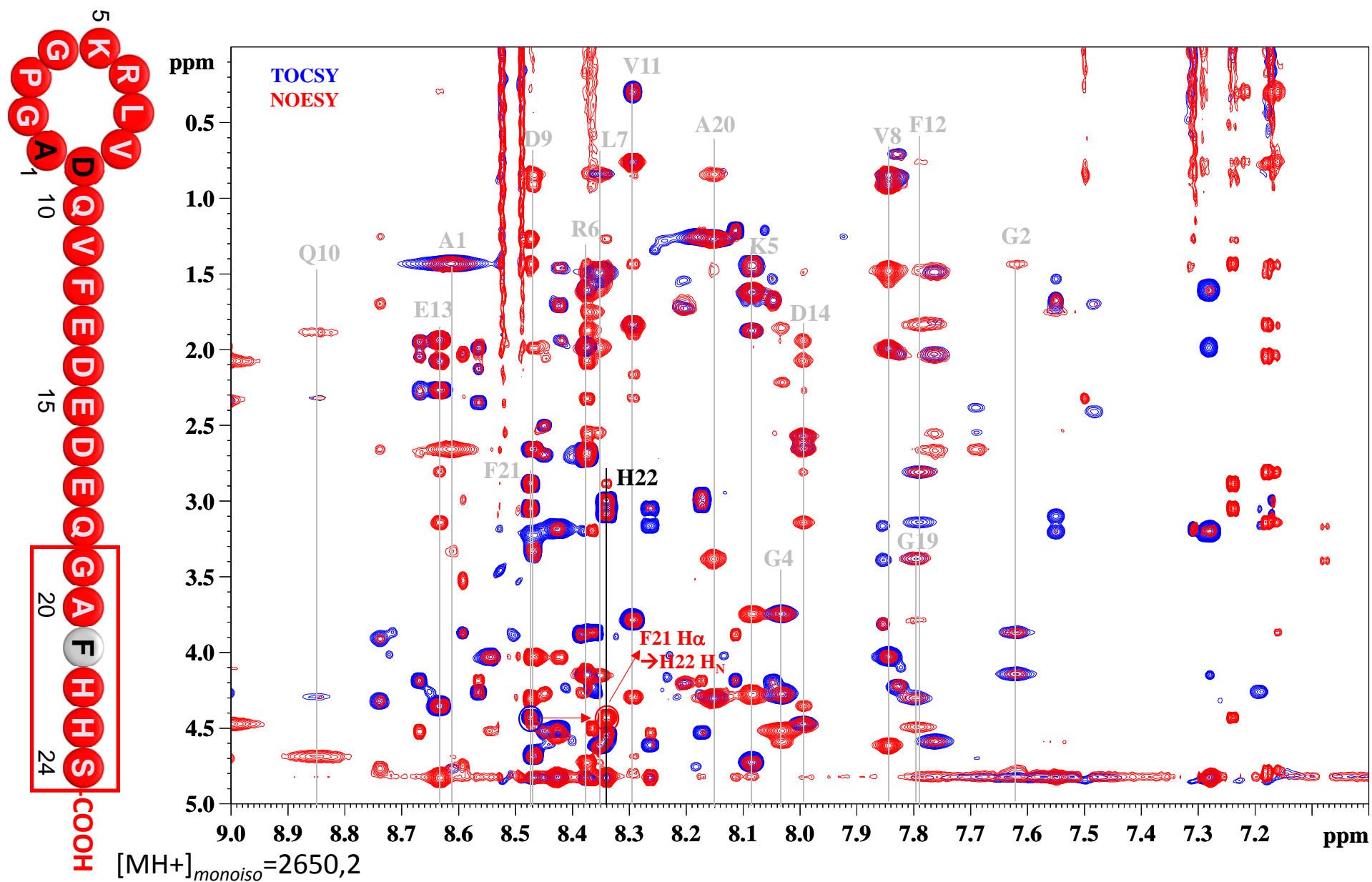


Figure S14  
L21F - 19

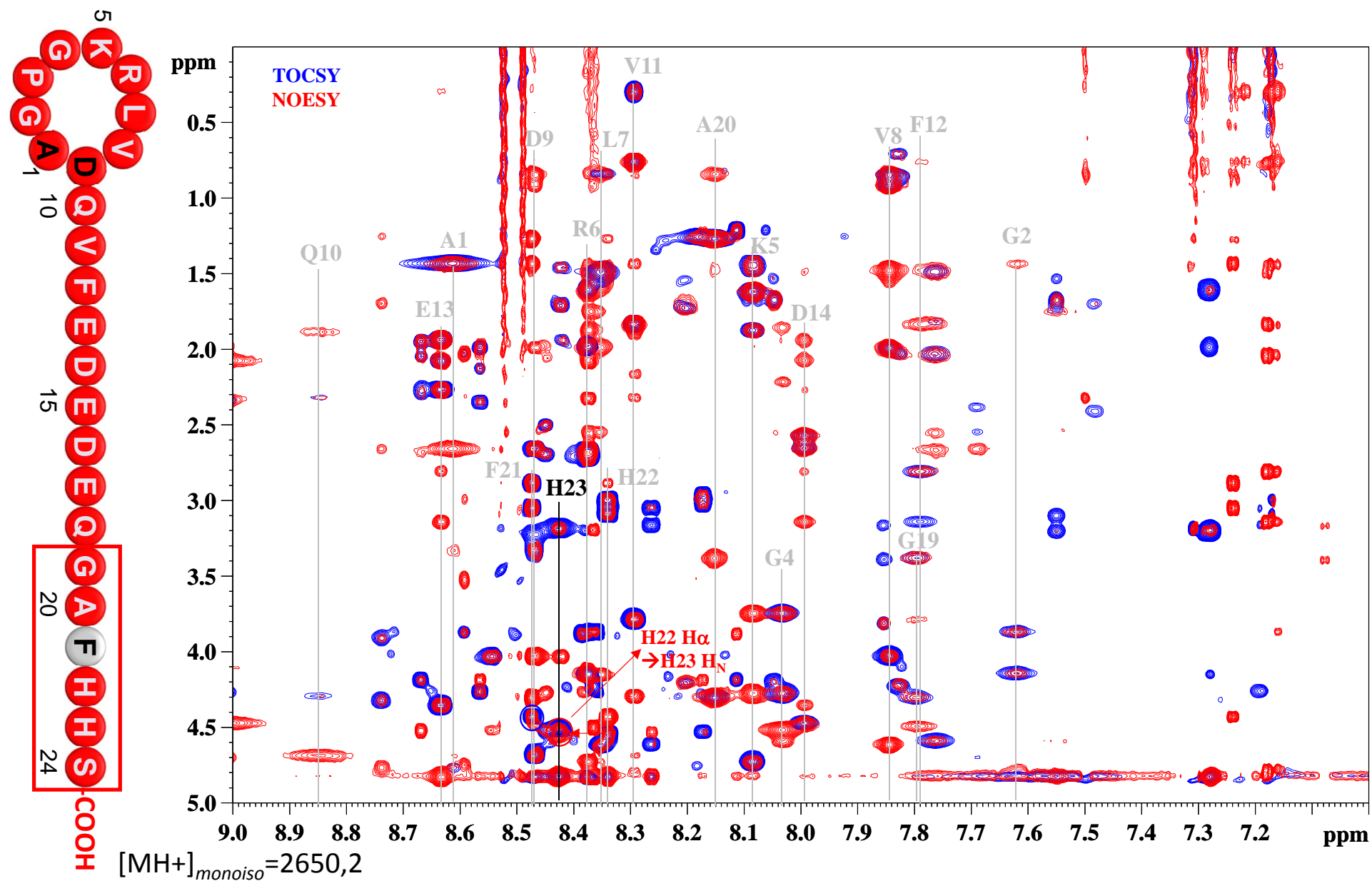


Figure S14  
 L21F - 20