

## Supplementary Material

### **Dispersion from C<sup>α</sup> or N<sup>H</sup>: 4D experiments for backbone resonance assignment of intrinsically disordered proteins**

**Helena Tossavainen<sup>1</sup>, Santeri Salovaara<sup>1</sup>, Maarit Hellman<sup>1</sup>, Riikka Ihalin<sup>2</sup>, Perttu Permi<sup>1,3\*</sup>**

<sup>1</sup>Department of Chemistry, Nanoscience Center, University of Jyväskylä, FI-40014 Jyväskylä, Finland.

<sup>2</sup>Department of Biochemistry, University of Turku, FI-20014 Turku, Finland.

<sup>3</sup>Department of Biological and Environmental Science, University of Jyväskylä, FI-40014 Jyväskylä, Finland

\*Corresponding author:

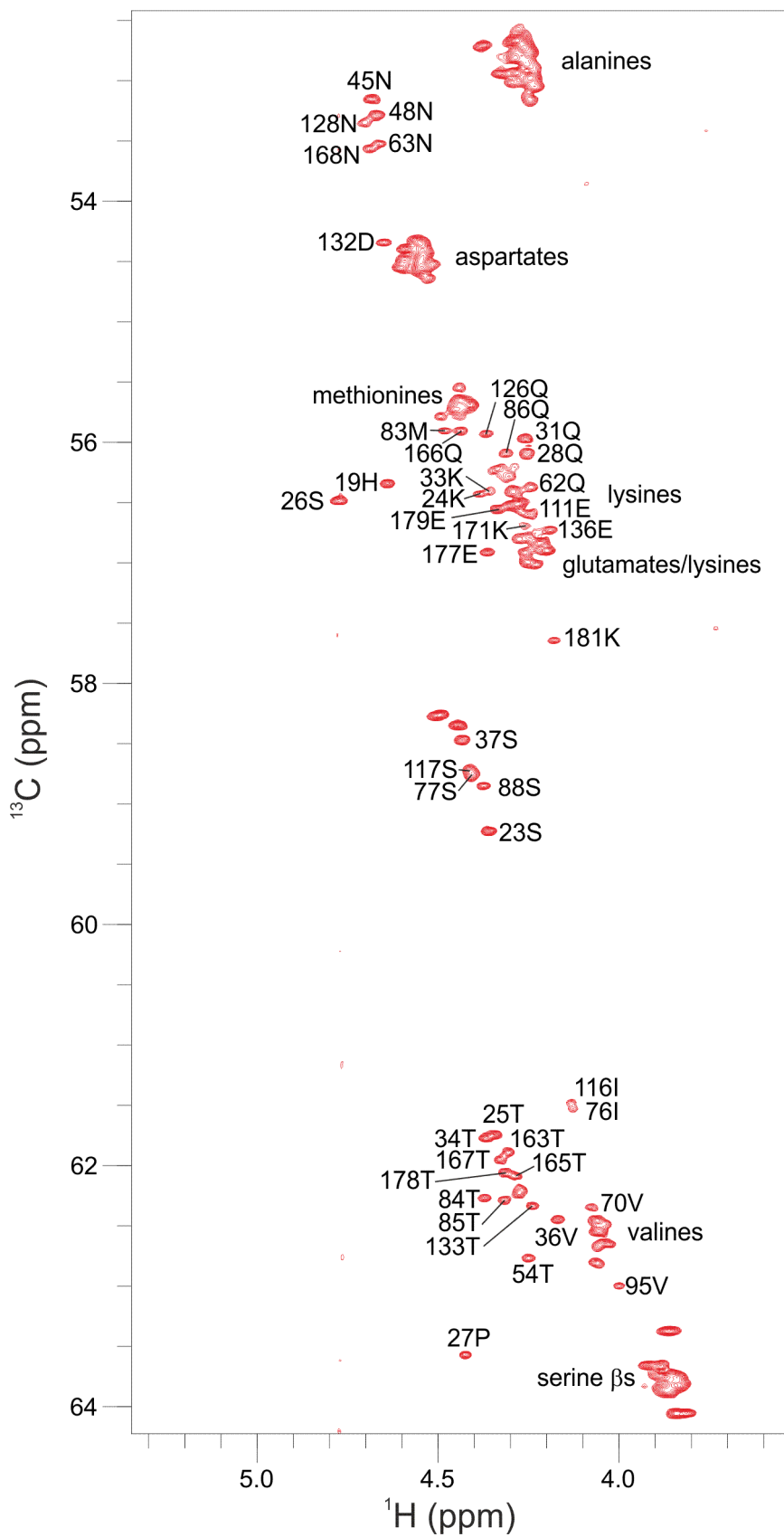
Perttu Permi, [perttu.permi@jyu.fi](mailto:perttu.permi@jyu.fi)

ORCID of the authors:

Helena Tossavainen, 0000-0002-1609-1651

Riikka Ihalin, 0000-0003-4118-0370

Perttu Permi, 0000-0002-6281-1138

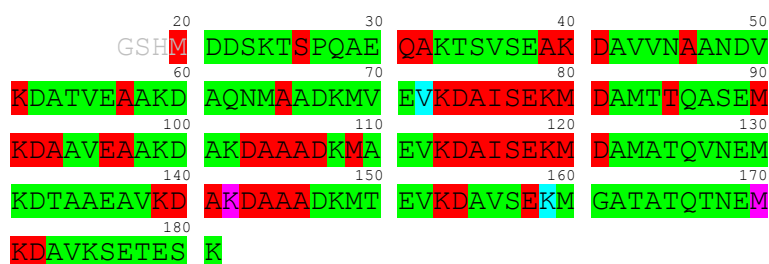


**Suppl. Fig. S1** The  $\text{C}^\alpha/\text{H}^\alpha$  region of the  $^1\text{H}$ ,  $^{13}\text{C}$  CT HSQC of BilRI measured at 800 MHz  $^1\text{H}$  frequency, 25 °C, 1 mM, pH 6.5. The resolved signals are labeled with sequence number and amino acid code.

## Assignment spectra

3D H<sup>N</sup> experiments, assigned C<sup>α/β</sup>, or C<sup>α</sup>, or C<sup>β</sup>

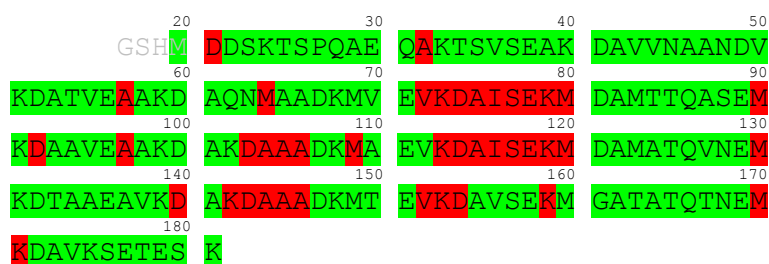
Assignment-%\*



C<sup>α</sup>: 66%

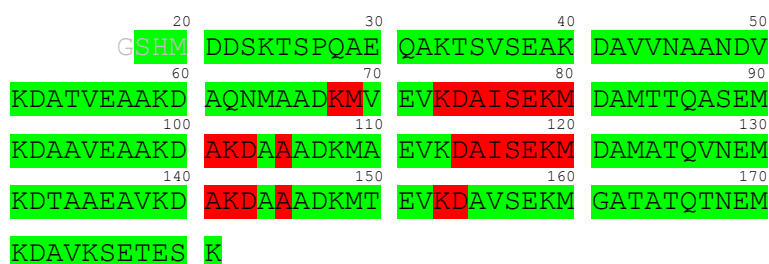
C<sup>β</sup>: 66%

3D H<sup>N</sup> experiments, assigned C<sup>γ</sup>



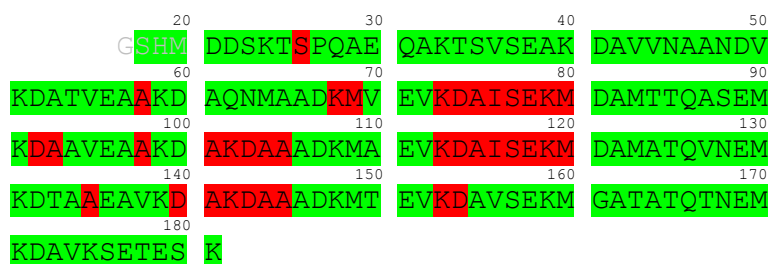
C<sup>γ</sup>: 74%

3D H<sup>α</sup> experiments, assigned H<sup>α</sup>



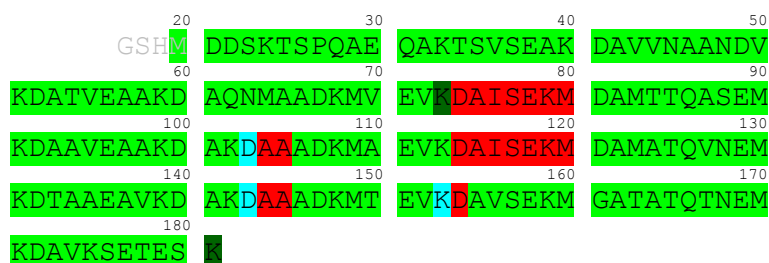
H<sup>α</sup>: 87%

4D H<sup>α</sup> experiments, assigned H<sup>α</sup>, C<sup>α</sup>



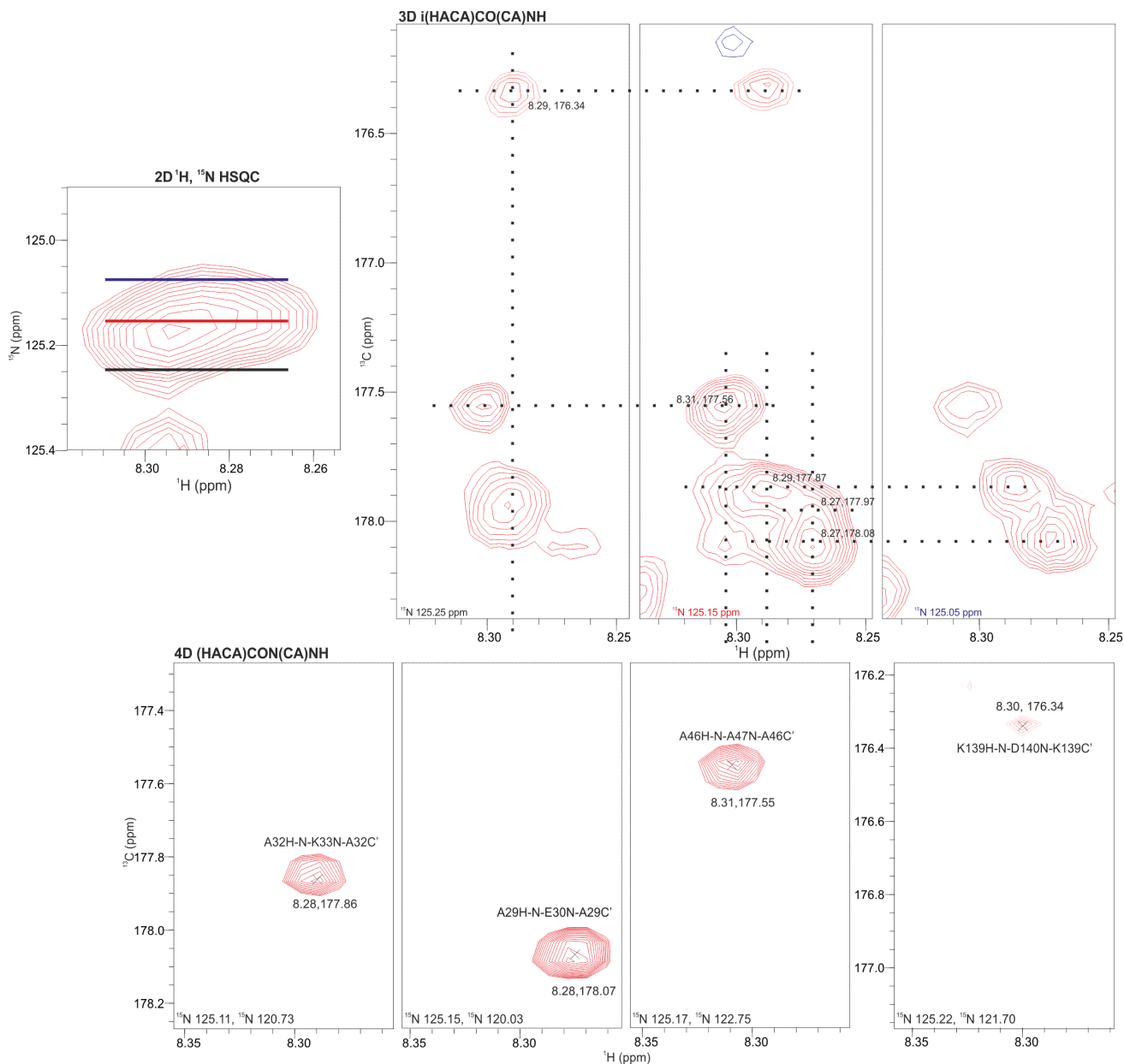
H<sup>α</sup>, C<sup>α</sup>: 76%

4D H<sup>N</sup> experiments, assigned H, N, C<sup>γ</sup> or H, N or N

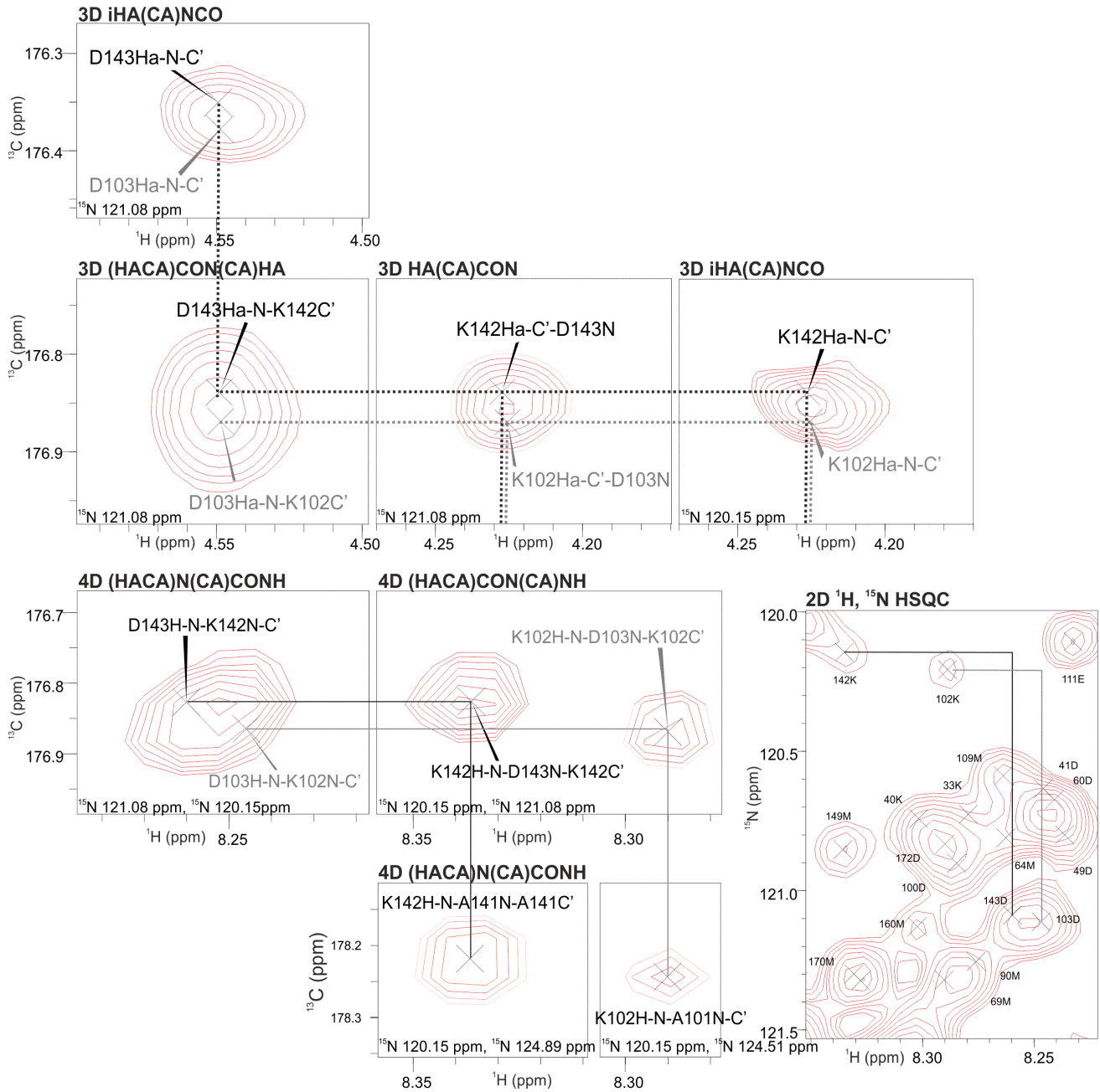


H<sup>N</sup>, N<sup>H</sup>: 86%

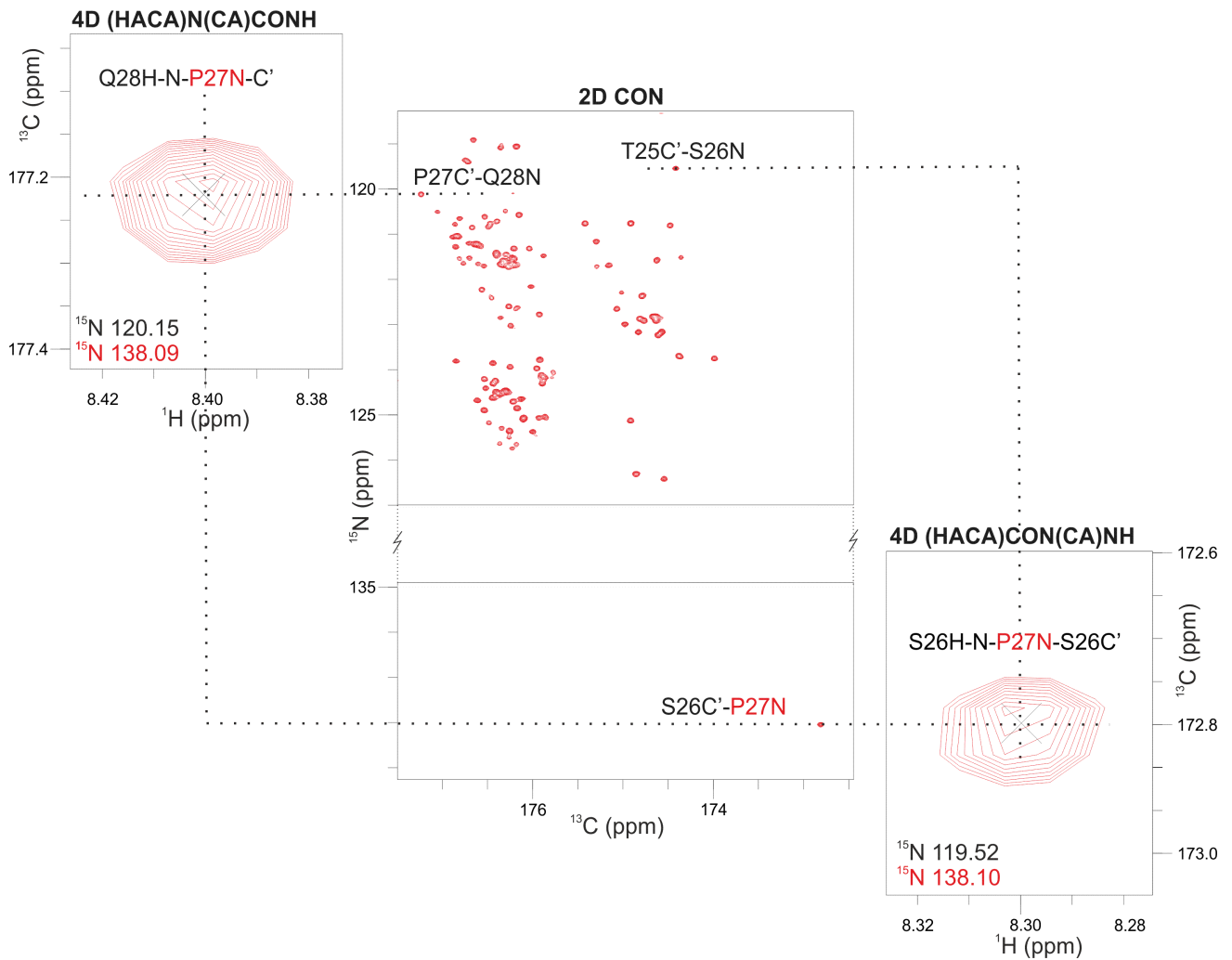
**Suppl. Fig. S2.** Assignment coverages and percentages from spectra used in the BilRI assignment. \*The N-terminal cloning artefact GSHM is excluded.



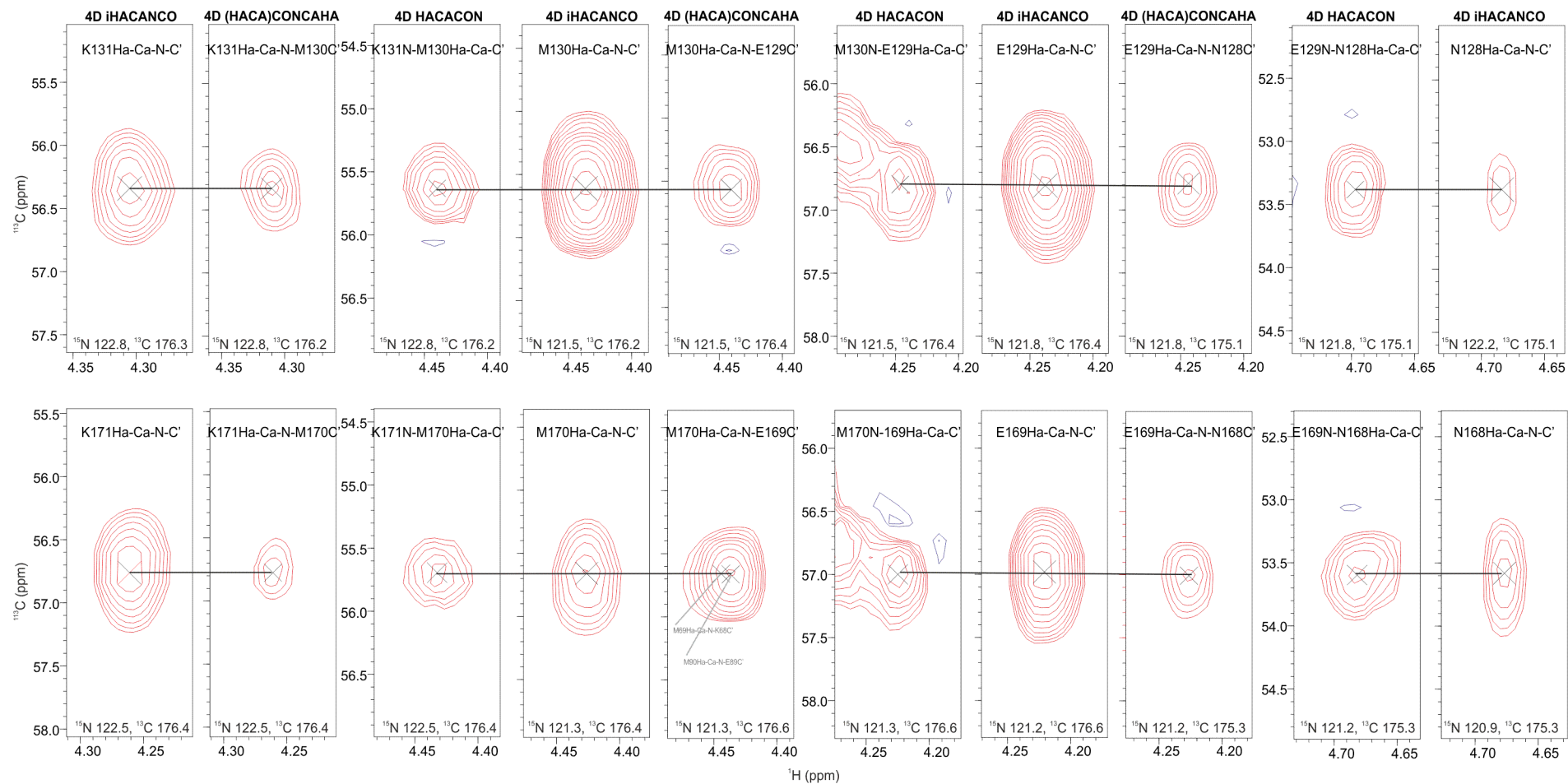
**Suppl. Fig. S3** Similar to Fig. 1 of the main text, the resolving power of the  $i+1$  amide nitrogen in the fourth dimension is presented. Compared are 3D  $i(\text{HACA})\text{CO}(\text{CA})\text{NH}$  and 4D  $(\text{HACA})\text{CON}(\text{CA})\text{NH}$  planes from the  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC region at upper left corner. From 2D  $\text{H}^{\text{N}}\text{-C}'$  planes of the 3D  $i(\text{HACA})\text{CO}(\text{CA})\text{NH}$  at  $^{15}\text{N}$  frequencies marked with lines of different colors it is difficult to determine the number and chemical shifts of peaks present in the overlapping peak area in the  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC. Tentative shifts are marked for five peaks. In contrast, in the 4D  $(\text{HACA})\text{CON}(\text{CA})\text{NH}$  spectrum four nicely resolved peaks are found.



**Suppl. Fig. S4** Representative planes from 4D (HACA)N(CA)CONH and (HACA)CON(CA)NH spectra, which demonstrate the resolving power of sequential  $N^H$  in resolving ambiguities present in 3D  $H^\alpha$ -detected experiments. Shown is the assignment of  $^{102}\text{Lys}$ - $^{103}\text{Asp}$ / $^{142}\text{Lys}$ - $^{143}\text{Asp}$  in the KDAKDAAADKM stretches (residues 99-106 and 139-146). The intraresidual and sequential  $H^\alpha$ ,  $N^H$ ,  $C'$  correlations of  $^{103}\text{Asp}$  and  $^{143}\text{Asp}$  appear as one peak in the three  $H^\alpha$ -detected experiments, and their assignments remain ambiguous. Although the  $^{143}\text{Asp}H, N$ - $^{142}\text{Lys} N, C$ / $^{103}\text{Asp}H, N$ - $^{102}\text{Lys} N, C'$  correlations are not resolved in the 4D (HACA)N(CA)CONH, thanks to distinct  $H^N$  shifts, the preceding residues display well separated peaks in the 4D (HACA)CON(CA)NH, which allows to proceed with the sequential walk. The sequential walk is presented in the  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectrum at the lower right corner. Grey labels mark correlations with peak maximum in adjacent  $^{15}\text{N}/^{15}\text{N}$ ,  $^{13}\text{C}$  planes.



**Suppl. Fig. S5** Assignment of proline-containing amino acid sequences using 4D (HACAN(CA)CONH and (HACA)CON(CA)NH. Shown is the assignment of  $^{28}\text{Gln}$ - $^{27}\text{Pro}$ - $^{26}\text{Ser}$  triplet in BiIRI. The (HACA)N(CA)CONH shows a correlation between  $^{28}\text{Gln}$   $\text{H}^{\text{N}}$ ,  $\text{N}^{\text{H}}$  and  $\text{N}^{\text{H}}$ ,  $\text{C}'$  of the preceding proline. At this proline  $\text{N}^{\text{H}}$  chemical shift in the (HACA)CON(CA)NH the  $\text{H}^{\text{N}}$ ,  $\text{N}^{\text{H}}$ ,  $\text{C}'$  of the preceding  $^{26}\text{Ser}$ . The corresponding peaks and the pathway between them are shown in the 2D CON spectrum.



**Suppl. Fig. S5** Assignment of identical stretches  $^{128}\text{Asn}$ - $^{129}\text{Glu}$ - $^{130}\text{Met}$ - $^{131}\text{Lys}$  and  $^{168}\text{Asn}$ - $^{169}\text{Glu}$ - $^{170}\text{Met}$ - $^{171}\text{Lys}$  using the described three 4D experiments iHACANCO, (HACA)CONCAHA and HACACON.