

## Supporting Information for:

# Different molecular recognition by three domains of the full-length GRB2 to SOS1 proline-rich motifs and EGFR phosphorylated sites

Keita Tateno<sup>a,†</sup>, Takami Ando<sup>a,†</sup>, Maako Tabata<sup>a</sup>, Haruka Sugasawa<sup>a</sup>, Toshifumi Hayashi<sup>a</sup>, Sanya Yu<sup>a</sup>, Sayeesh PM<sup>a</sup>, Kohsuke Inomata<sup>a</sup>, Tsutomu Mikawa<sup>b</sup>, Yuji Sugita<sup>b</sup>, Yutaka Ito<sup>a,\*</sup> and Teppei Ikeya<sup>a\*</sup>

<sup>a</sup>Department of Chemistry, Graduate School of Science, Tokyo Metropolitan University, 1-1 Minamiosawa, Hachioji, Tokyo 192-0397, Japan; <sup>b</sup>RIKEN Center for Biosystems Dynamics Research, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan; <sup>c</sup>Theoretical Molecular Science Laboratory, RIKEN Cluster for Pioneering Research, Wako, Saitama 351-0198, Japan; <sup>d</sup>Laboratory for Biomolecular Function Simulation, RIKEN Center for Biosystems Dynamics Research, Kobe, Hyogo 650-0047, Japan; <sup>e</sup>Computational Biophysics Research Team, RIKEN Center for Computational Science, Kobe, Hyogo 650-0047, Japan

<sup>†</sup> These authors contributed equally to this work

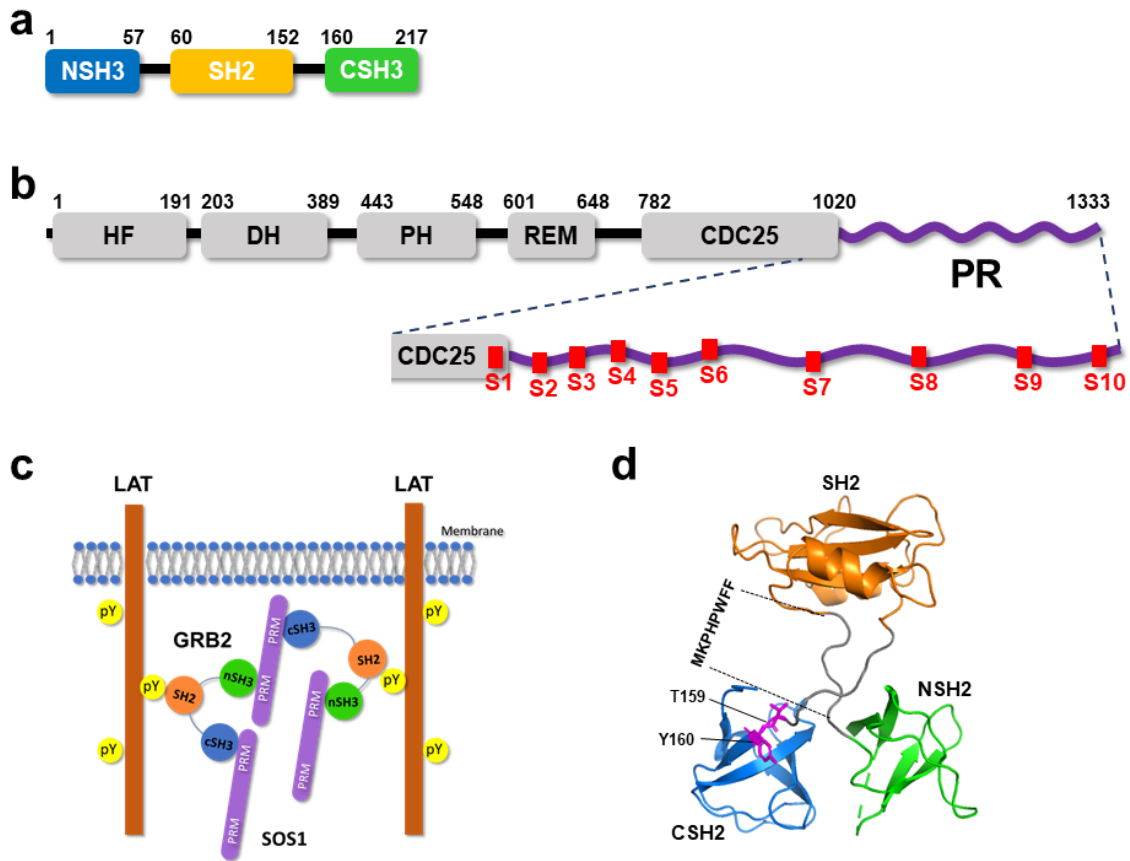
\* To whom correspondence should be addressed. E-mail: T.I and Y.I (tikeya@tmu.ac.jp and ito-yutaka@tmu.ac.jp)

## Contents

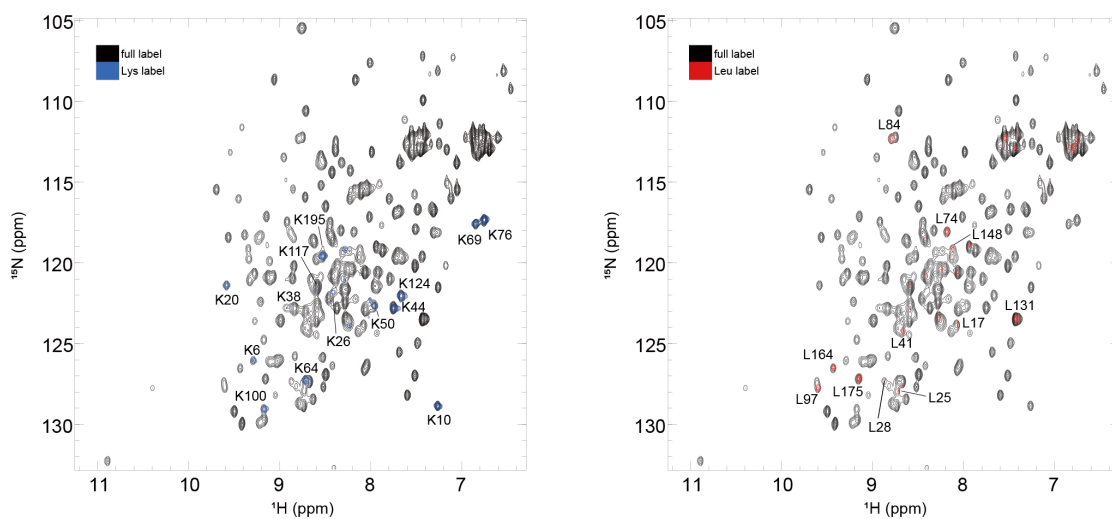
- 1. Supplementary Figure S1.** Domain and three dimensional (3D) structures of GRB2 and SOS1, and schematic illustration of LLPS.
- 2. Supplementary Figure S2.** Overlays of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$ -uniformly and amino-acid-selectively labelled samples.
- 3. Supplementary Figure S3.** Backbone resonance assignments of GRB2.
- 4. Supplementary Figure S4.** Table summarising the backbone and side-chain resonance assignments of GRB2.
- 5. Supplementary Figure S5.** Overlays of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of the full-length GRB2 and each isolated domain.
- 6. Supplementary Figure S6.** Overlays of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra from multipoint titrations of  $^{15}\text{N}$ -labelled GRB2 with SOS-S5 PRM.
- 7. Supplementary Figure S7.** Overlays of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra from multipoint titrations of  $^{15}\text{N}$ -labelled GRB2 with SOS-S9 PRM.
- 8. Supplementary Figure S8.** Overlays of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra from multipoint titrations of  $^{15}\text{N}$ -labelled GRB2 with SOS-S10 PRM.
- 9. Supplementary Figure S9.** Plots of relative change in chemical shift perturbation (CSP) of backbone  $^1\text{H}^{\text{N}}$  and  $^{15}\text{N}$  nuclei of GRB2 upon the titration with S4, S5, S9, and S10 PRMs.
- 10. Supplementary Figure S10.** Residual errors after the bootstrap calculations.
- 11. Supplementary Figure S11.** Chemical shift perturbation profile mapped on each residue, against the concentration of ligand with SOS S4 PRM.
- 12. Supplementary Figure S12.** Chemical shift perturbation profile mapped on each residue, against the concentration of ligand with SOS S5 PRM.
- 13. Supplementary Figure S13.** Chemical shift perturbation profile mapped on each residue, against the concentration of ligand with SOS S9 PRM.
- 14. Supplementary Figure S14.** Chemical shift perturbation profile mapped on each residue, against the concentration of ligand with SOS S10 PRM.
- 15. Supplementary Figure S15.** ITC analysis for the binding of the full length GRB2 to S2,

S4, S5, S6, S8 S9, and S10 PRMs..

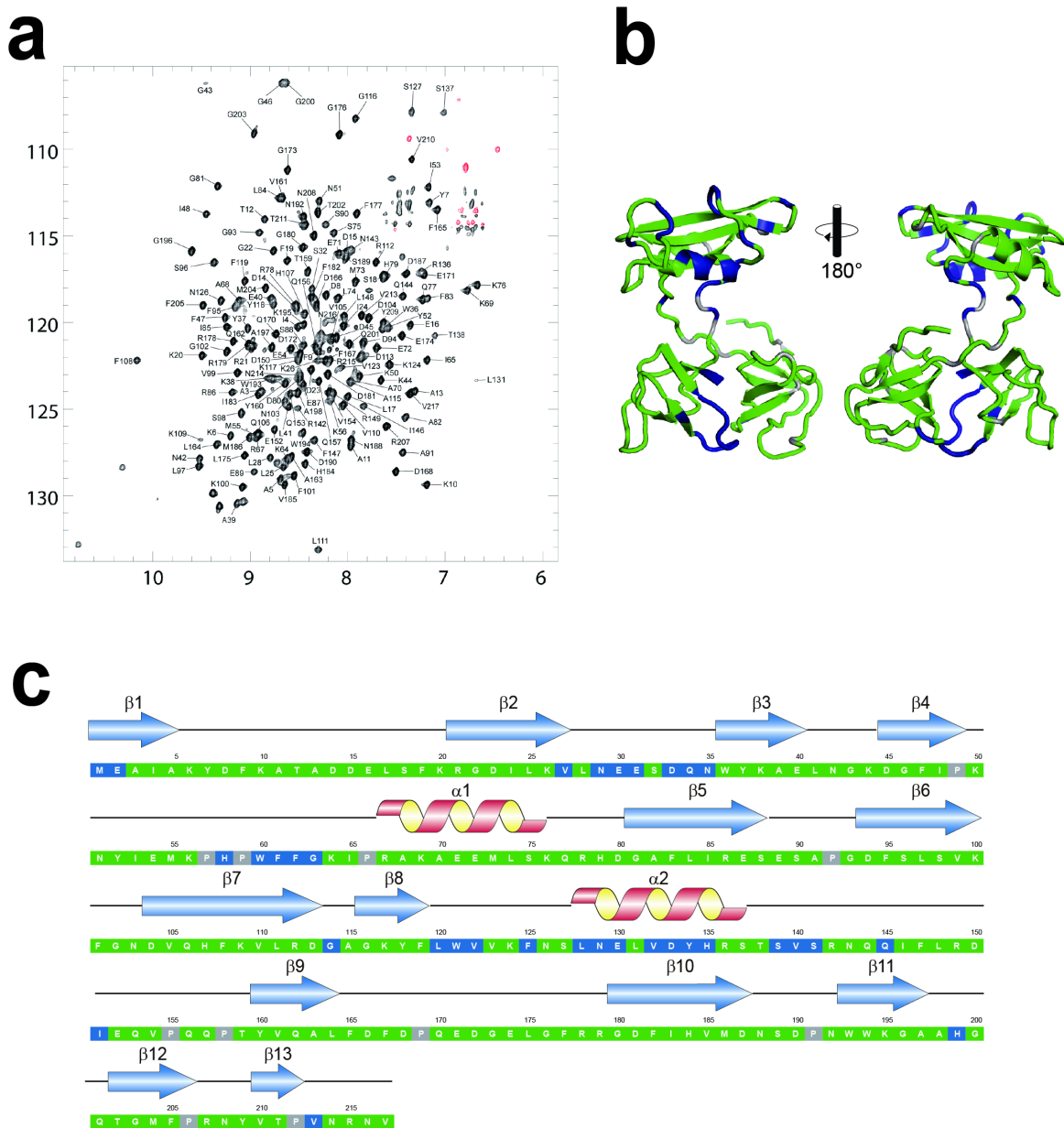
- 16. Supplementary Figure S16.** O ;verlays of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra from multipoint titrations of  $^{15}\text{N}$ -labelled GRB2 exclusively with the EGFR phosphorylated peptide.
- 17. Supplementary Figure S17.** Overlays of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra from multipoint titrations of  $^{15}\text{N}$ -labelled GRB2 with the EGFR phosphorylated peptide, in the presence of S4 PRM.
- 18. Supplementary Table S1** Dissociation constants derived from ITC analysis for the full length GRB2 and isolated SH3 domains against SOS1 PRMs.



**Figure S1. Domain and three dimensional (3D) structures of GRB2 and SOS1, and schematic illustration of LLPS.** (a) the domain structure of GRB2. (b) the domain structure of SOS1 with the positions of proline-rich motifs, SOS1 S1-S10, indicated by red boxes. (c) Schematic illustration of the considered multivalent interactions among GRB2, SOS1 and LAT in liquid-liquid phase separation (LLPS), proposed in the previous reports. (d) The crystal structure of GRB2.



**Figure S2. Overlays of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of  $^{15}\text{N}$ -uniformly and amino-acid-selectively labelled samples. The spectra of lysine- (left) and leucine- (right) selectively labelled GRB2 were overlaid onto those of the  $^{15}\text{N}$ -uniformly labelled samples.**

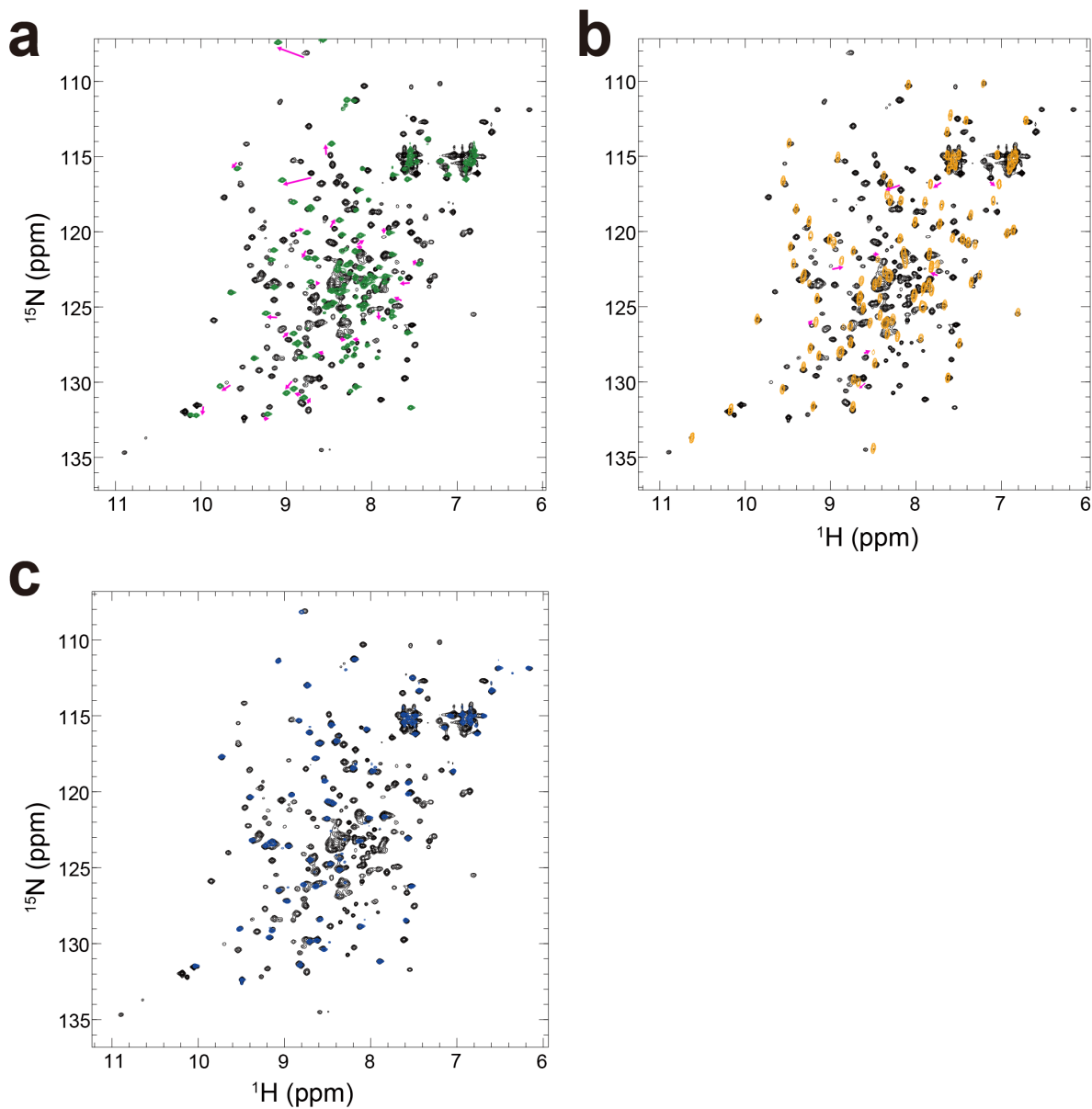


**Figure S3. Backbone resonance assignments of GRB2.** (a) 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum and backbone resonance assignments of GRB2. Positive and negative signals (aliased on the  $^{15}\text{N}$  axis) are color-coded in black and red, respectively. (b) The backbone resonance assignments are mapped onto the crystal structure of GRB2. Assigned and unassigned residues are colour-coded in green and blue, respectively. (c) The backbone resonance assignments aligned with

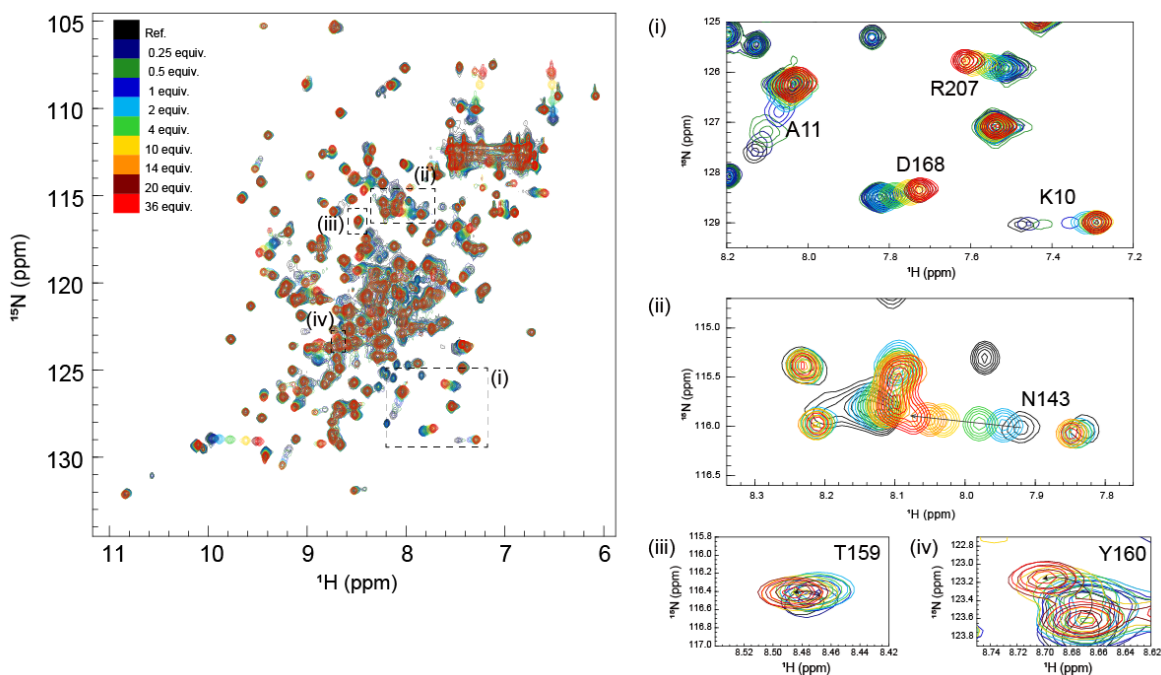
secondary structures. Assigned, unassigned, and proline residues are colour-coded in green, blue, and grey, respectively.



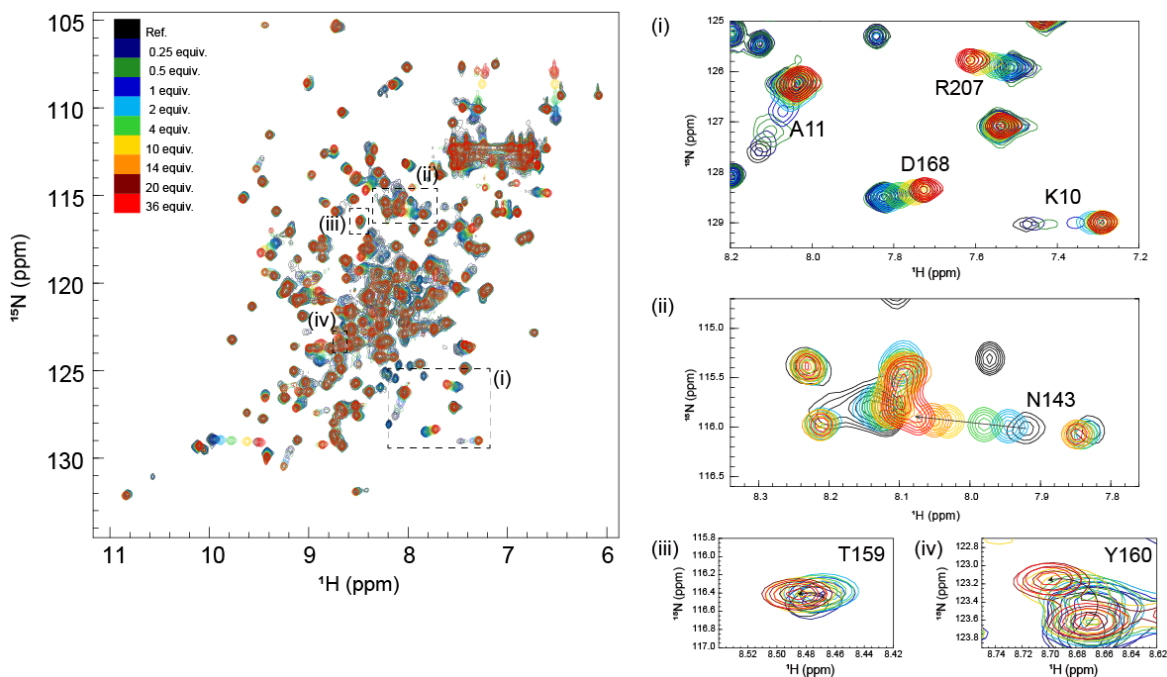




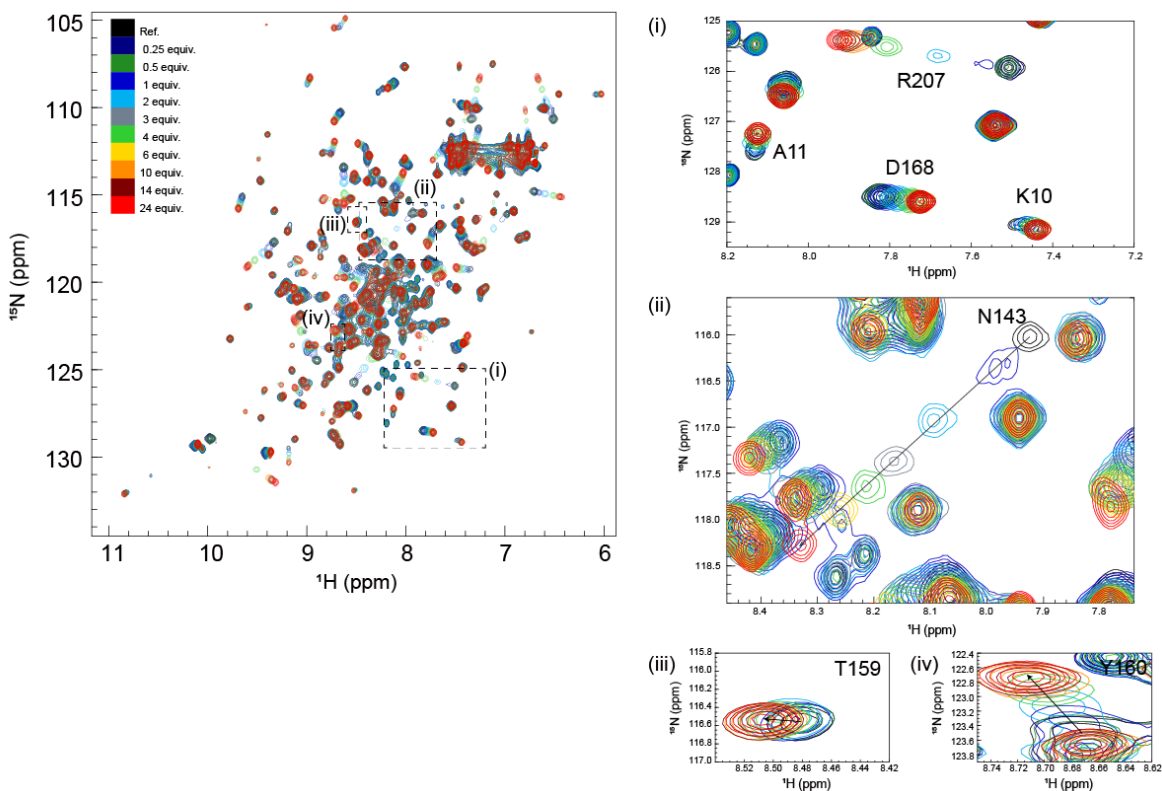
**Figure S5. Overlays of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of the full-length GRB2 and each isolated domain.** The black spectra indicate the full-length GRB2, and green, orange, and blue represent isolated NSH3, SH2, CSH3, respectively. The red arrows indicate representative peaks that significant chemical shift changes were observed.



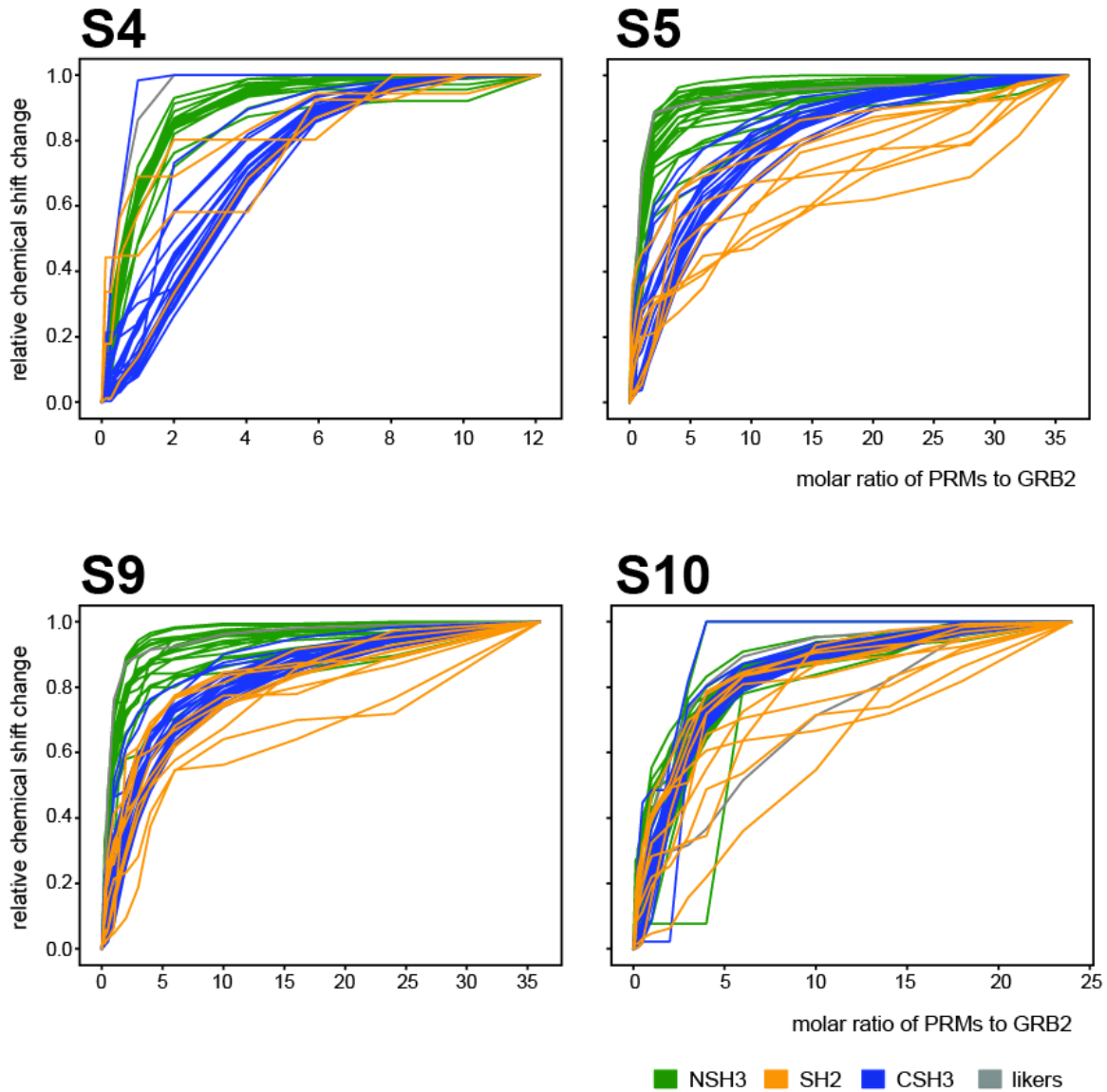
**Figure S6. Overlays of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra from multipoint titrations of  $^{15}\text{N}$ -labelled GRB2 with SOS-S5 PRM (DSPPAOPPRQPT).** The peptide concentration was increased stepwise (the protein: peptide molar ratio of 1:0.25, 1:0.5, 1:1, 1:2, 1:4, 1:10, 1:14, 1:20, and 1:36). In this figure, the colour codes of  $^1\text{H}$ - $^{15}\text{N}$  correlation cross-peaks at each titration point, showing the molar ratio of GRB2: SOS-S5, are as follows: black (1:0); dark blue (1:0.25); dark green (1:0.5); blue (1:1); cyan (1:2); green (1:4); yellow (1:10); orange (1:14); dark red (1:20), red (1:36). Cross peaks that showed large chemical shift changes were annotated.



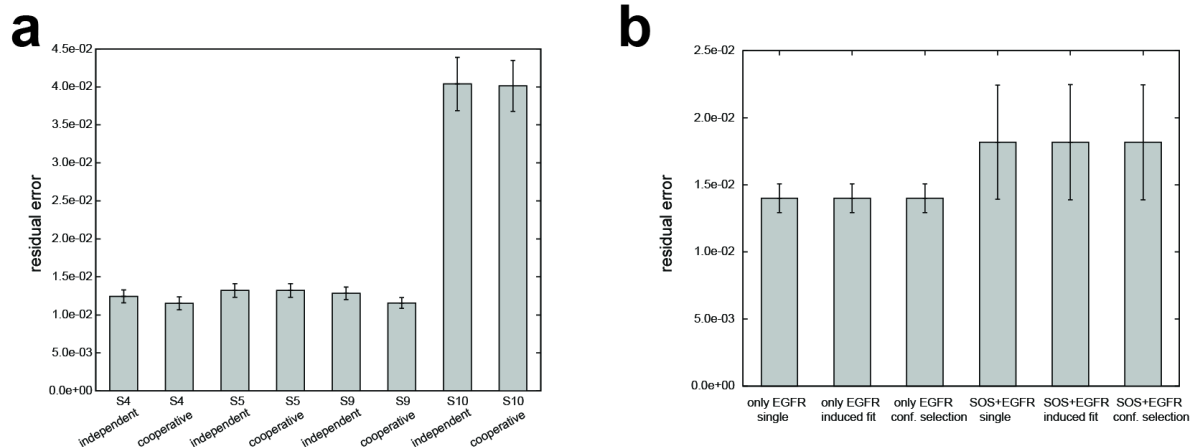
**Figure S7. Overlays of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra from multipoint titrations of  $^{15}\text{N}$ -labelled GRB2 with SOS-S9 PRM (IAGPPVPPRQST).** The peptide concentration was increased stepwise (the protein: peptide molar ratio of 1:0.25, 1:0.5, 1:1, 1:2, 1:4, 1:10, 1:14, 1:20, and 1:36). In this figure, the colour codes of  $^1\text{H}$ - $^{15}\text{N}$  correlation cross-peaks at each titration point, showing the molar ratio of GRB2: SOS-S9, are as follows: black (1:0); dark blue (1:0.25); dark green (1:0.5); blue (1:1); cyan (1:2); green (1:4); yellow (1:10); orange (1:14); dark red (1:20), red (1:36). Cross peaks that showed large chemical shift changes were annotated.



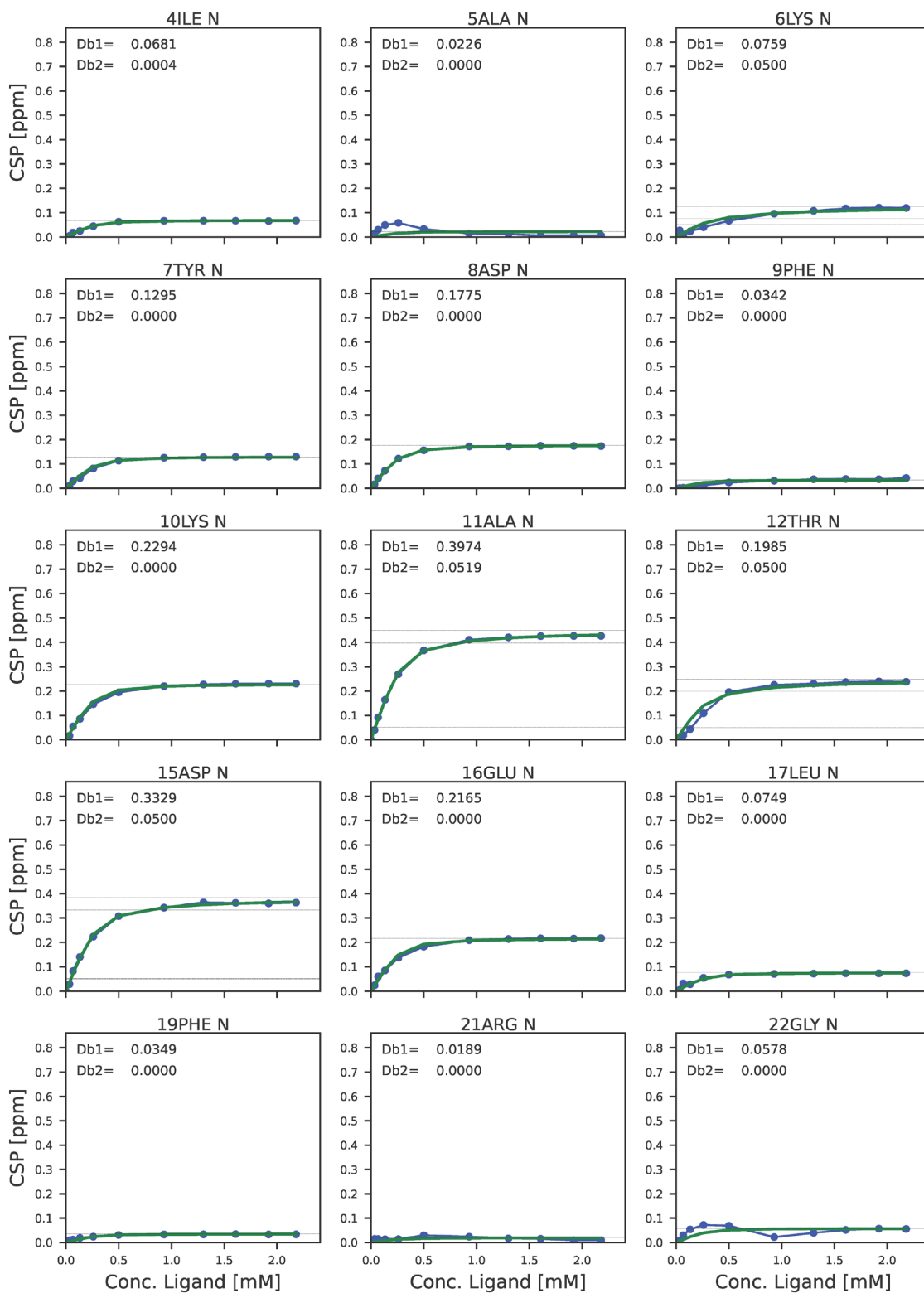
**Figure S8. Overlays of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra from multipoint titrations of  $^{15}\text{N}$ -labelled GRB2 with SOS-S10 PRM (PKLPPKTYKREH).** The peptide concentration was increased stepwise (the protein: peptide molar ratio of 1:0.25, 1:0.5, 1:1, 1:2, 1:3, 1:4, 1:6, 1:10, 1:14, and 1:24). In this figure, the colour codes of  $^1\text{H}$ - $^{15}\text{N}$  correlation cross-peaks at each titration point, showing the molar ratio of GRB2: SOS-S10, are as follows: black (1:0); dark blue (1:0.25); dark green (1:0.5); blue (1:1); cyan (1:2); grey (1:3), green (1:4); yellow (1:6); orange (1:10); dark red (1:14), red (1:24). Cross peaks that showed large chemical shift changes were annotated.

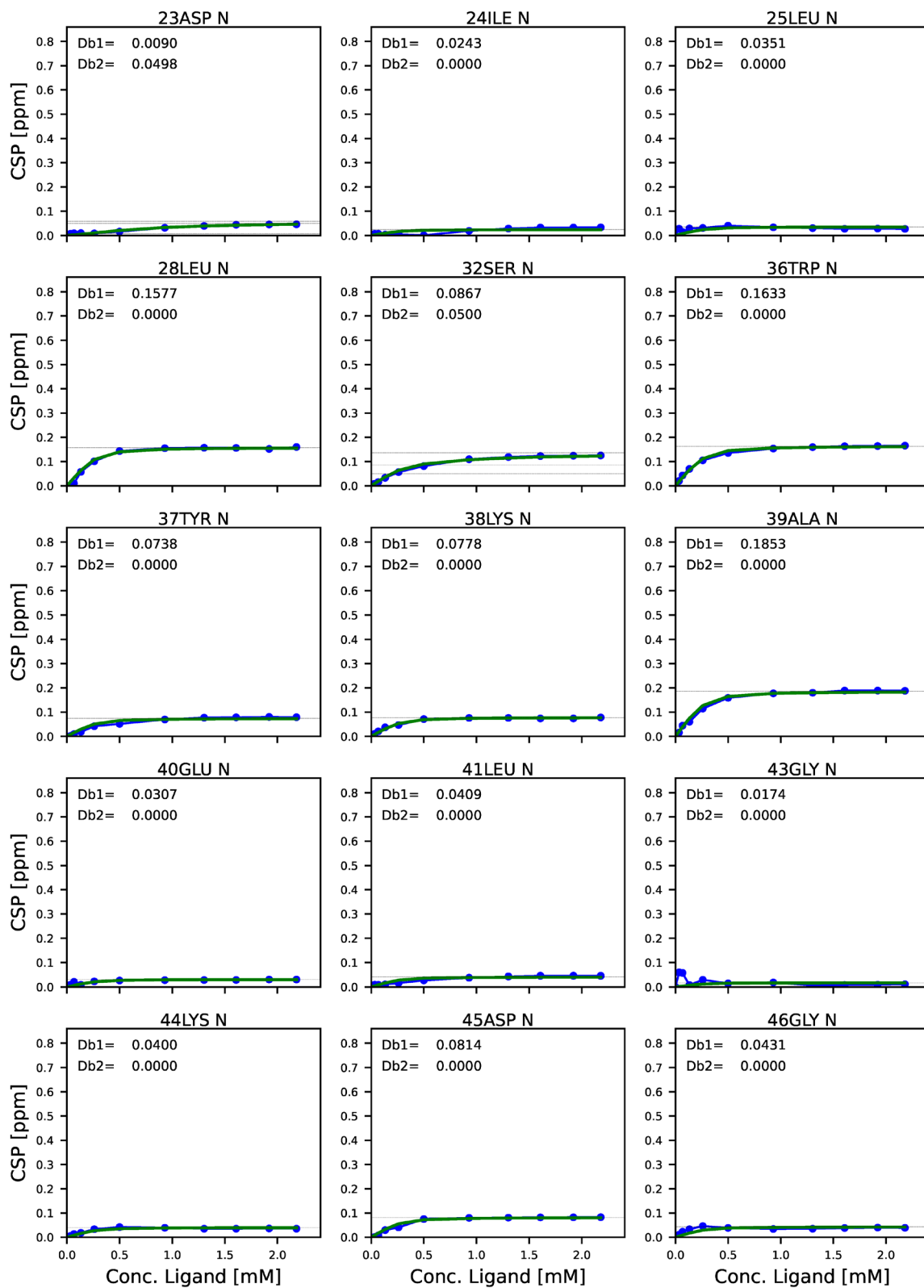


**Figure S9.** Plots of relative change in chemical shift perturbation (CSP) of backbone  $^1\text{H}^{\text{N}}$  and  $^{15}\text{N}$  nuclei of GRB2 upon the titration with S4, S5, S9, and S10 PRMs. A line represents the relative CSP beyond 0.06 ppm for a residue. Lines are colour-coded for each domain and linker.

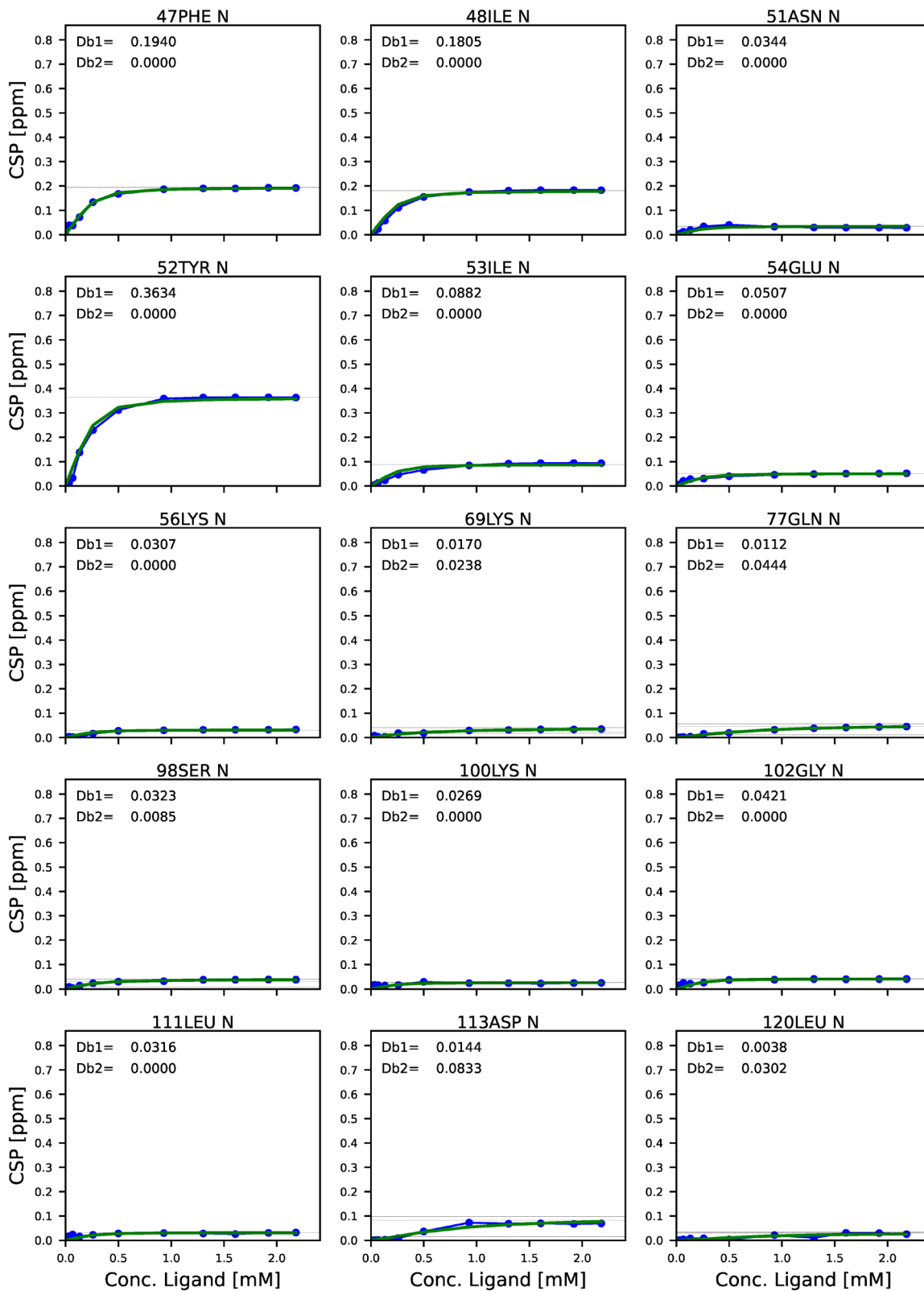


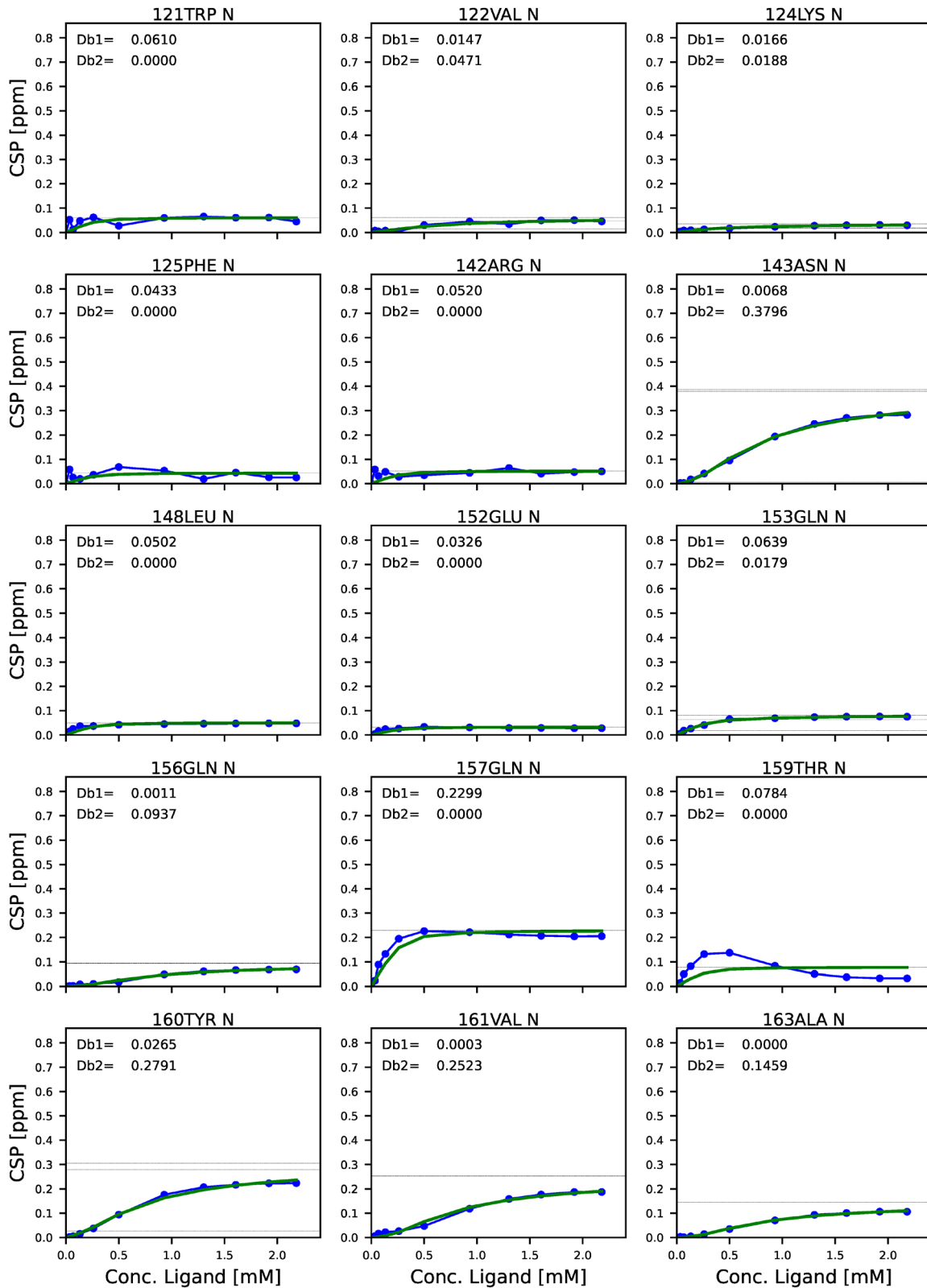
**Figure S10. Residual errors after the bootstrap calculations.** Residual errors after the bootstrap calculations with the CSPs of the S4, S5, S9, and S10 PRMs titration experiments for two-independent and two-cooperative binding models. residual errors after the bootstrap calculations for two-independent and two-cooperative binding models (right)

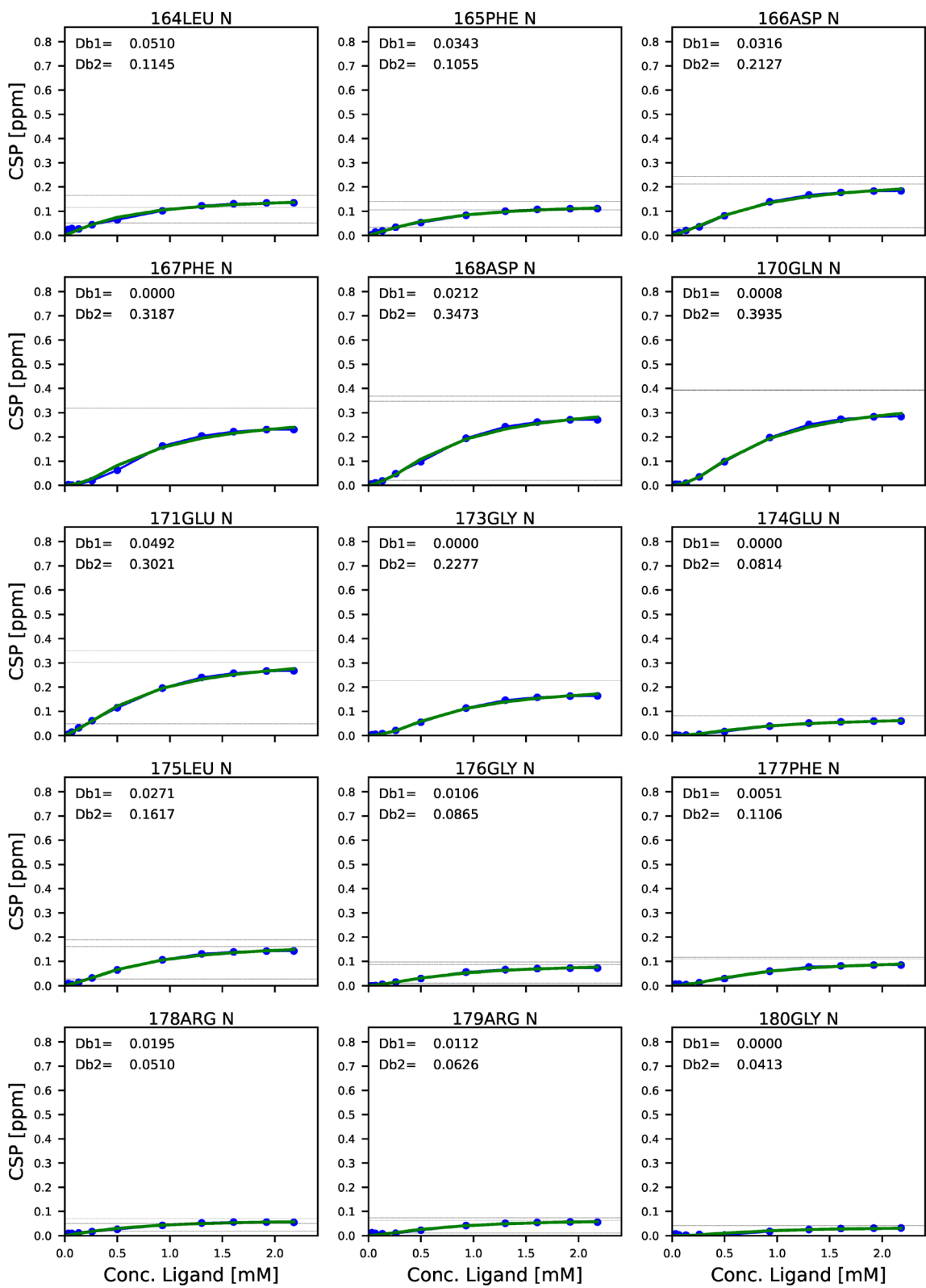


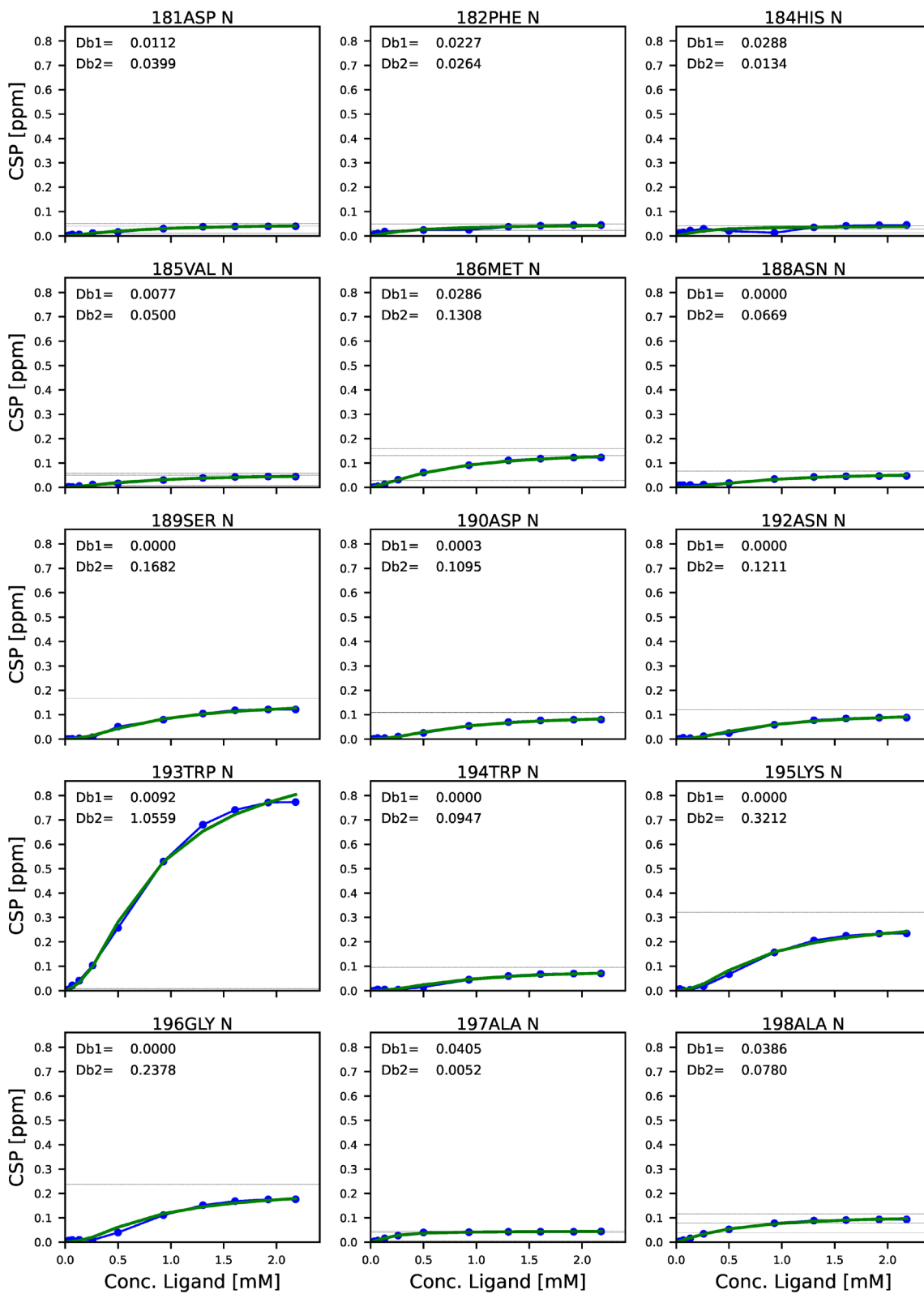


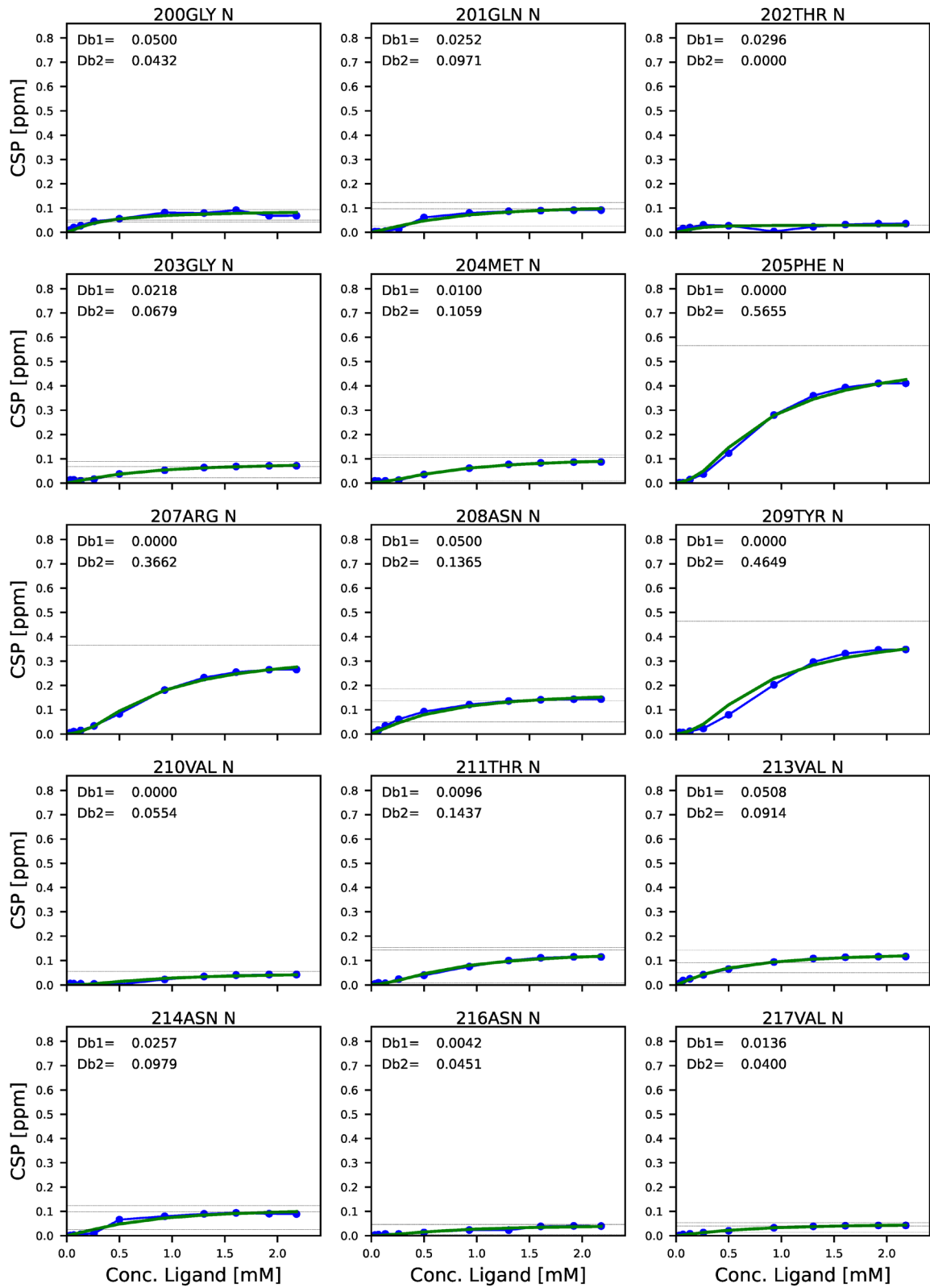




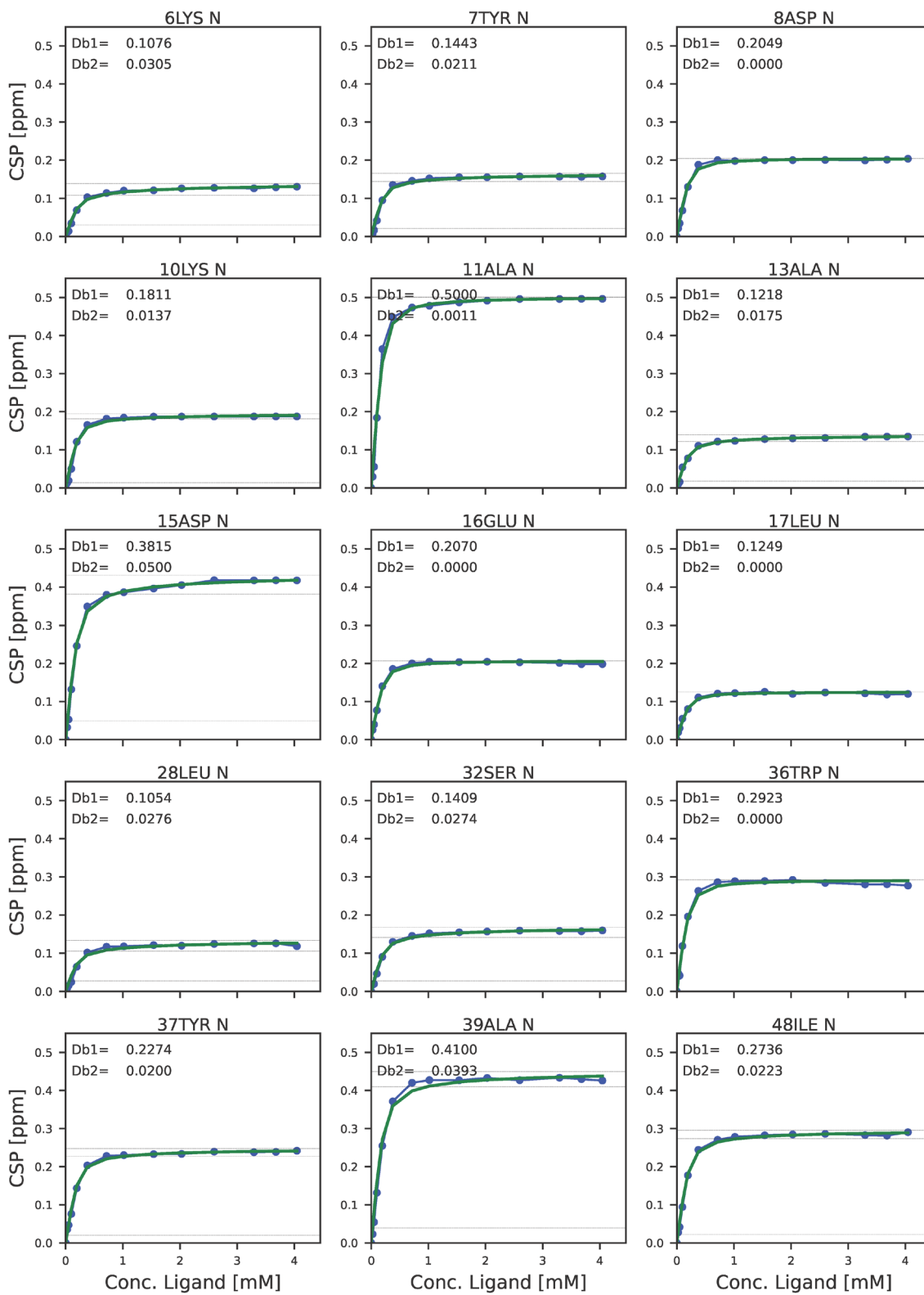


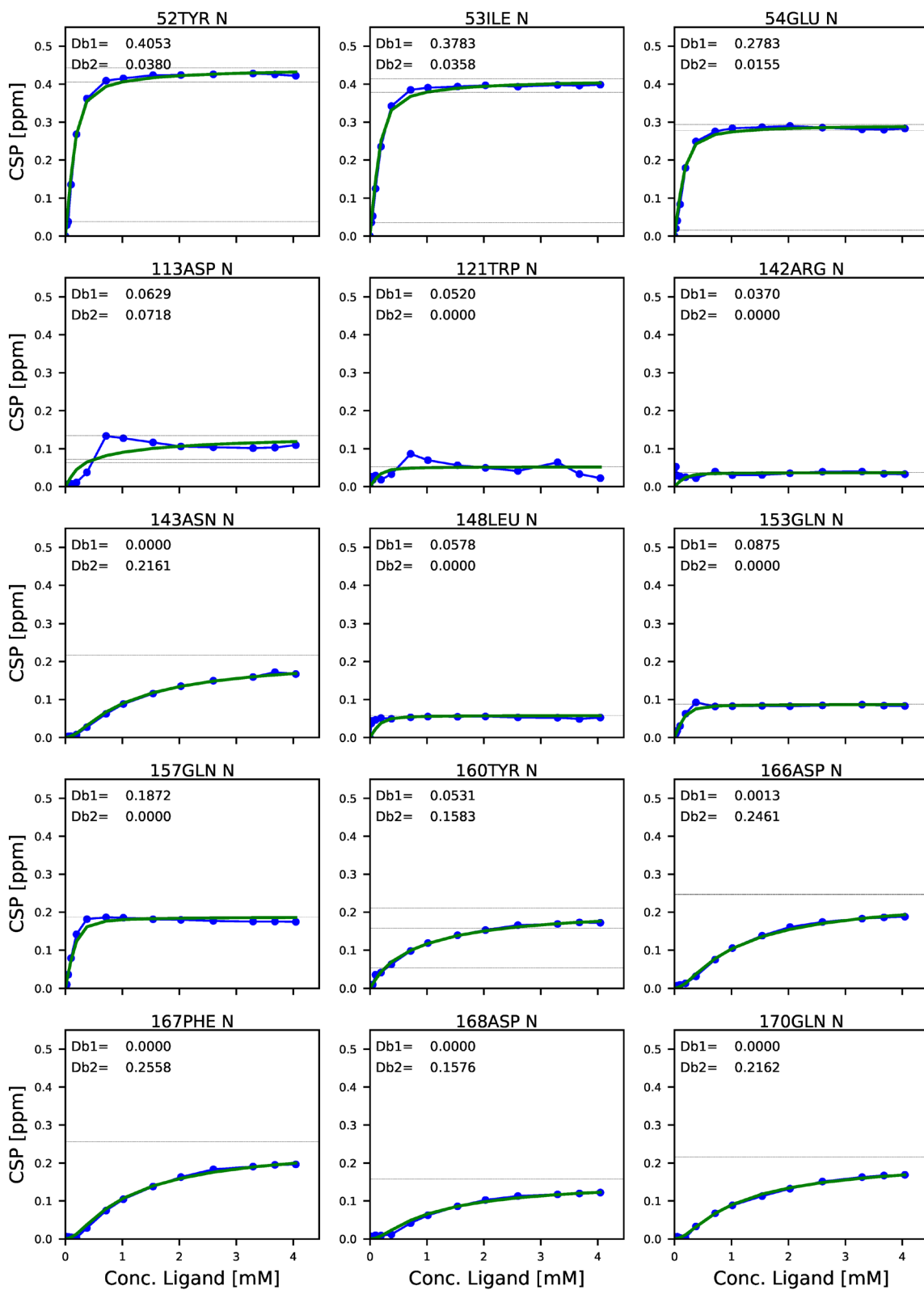




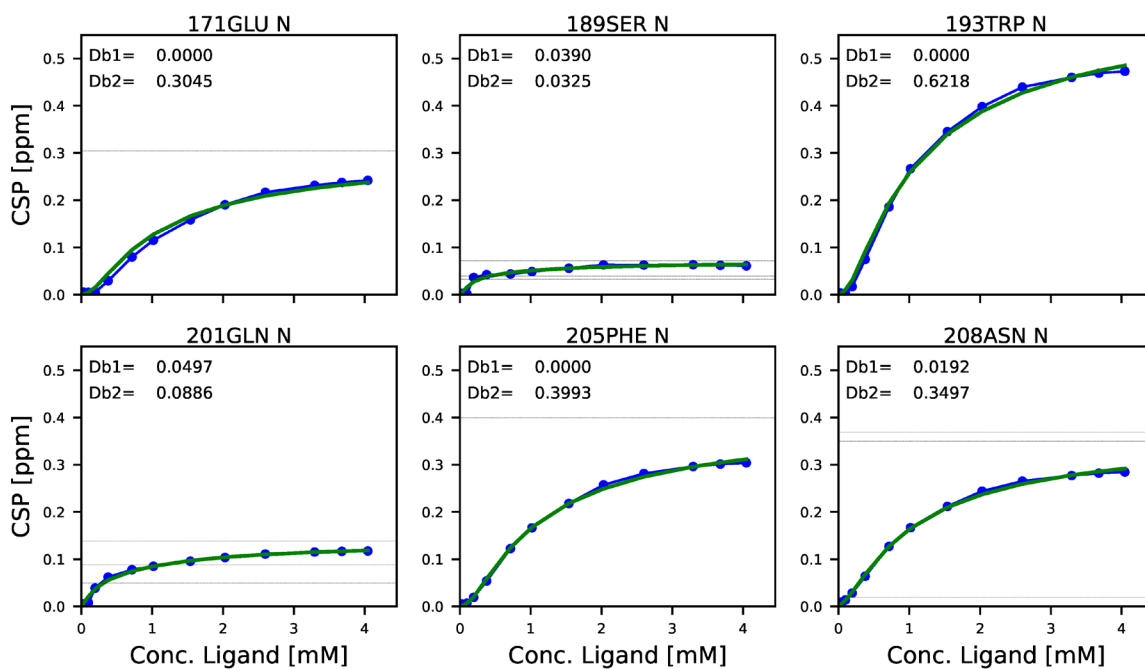


**Figure S11. Chemical shift perturbation profile mapped on each residue, against the concentration of ligand with SOS S4 PRM.** The blue dots and lines indicate CSP values at each titration point, while the green lines represent the theoretical curves calculated based on the two-site independent binding model.

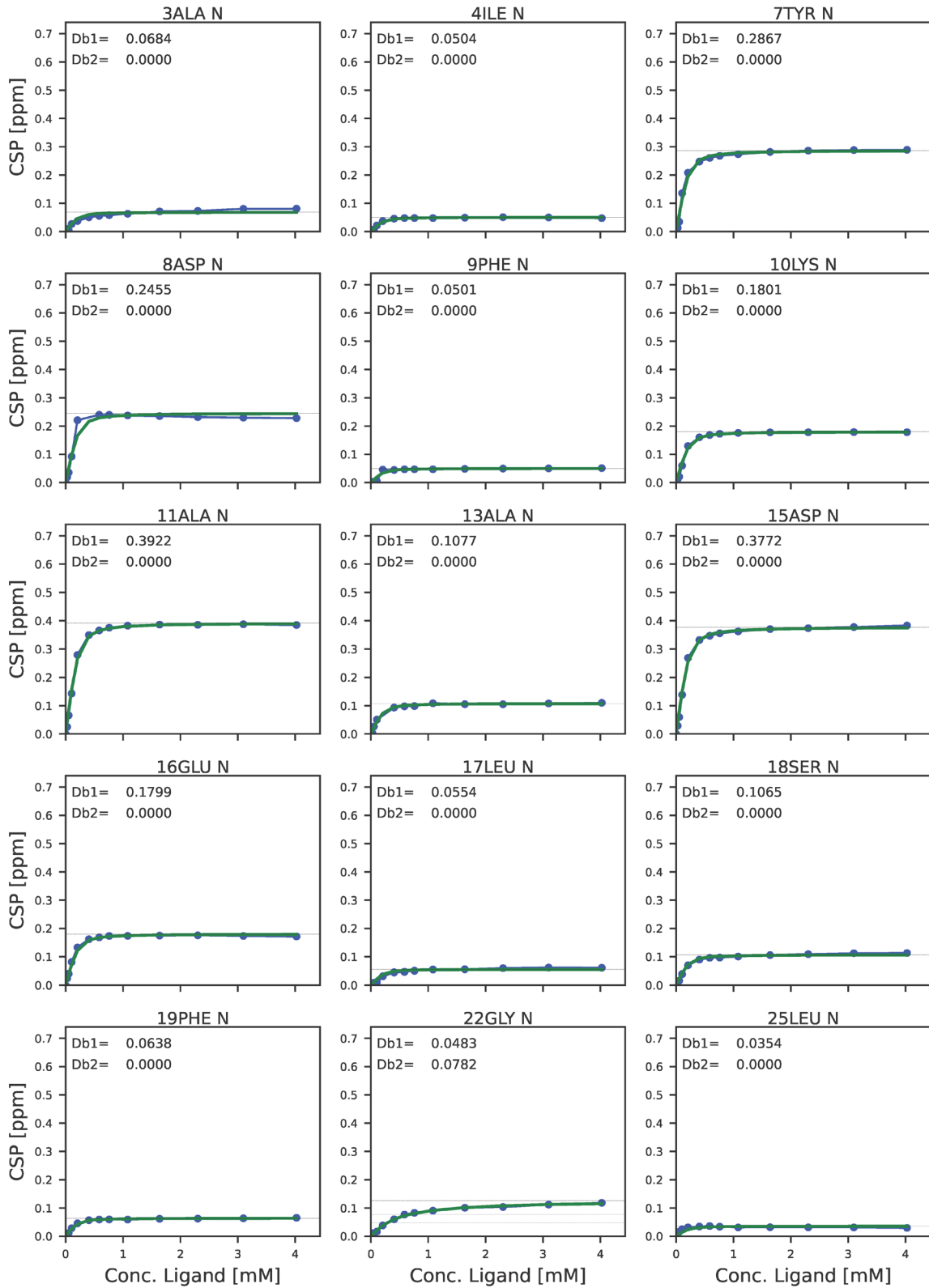


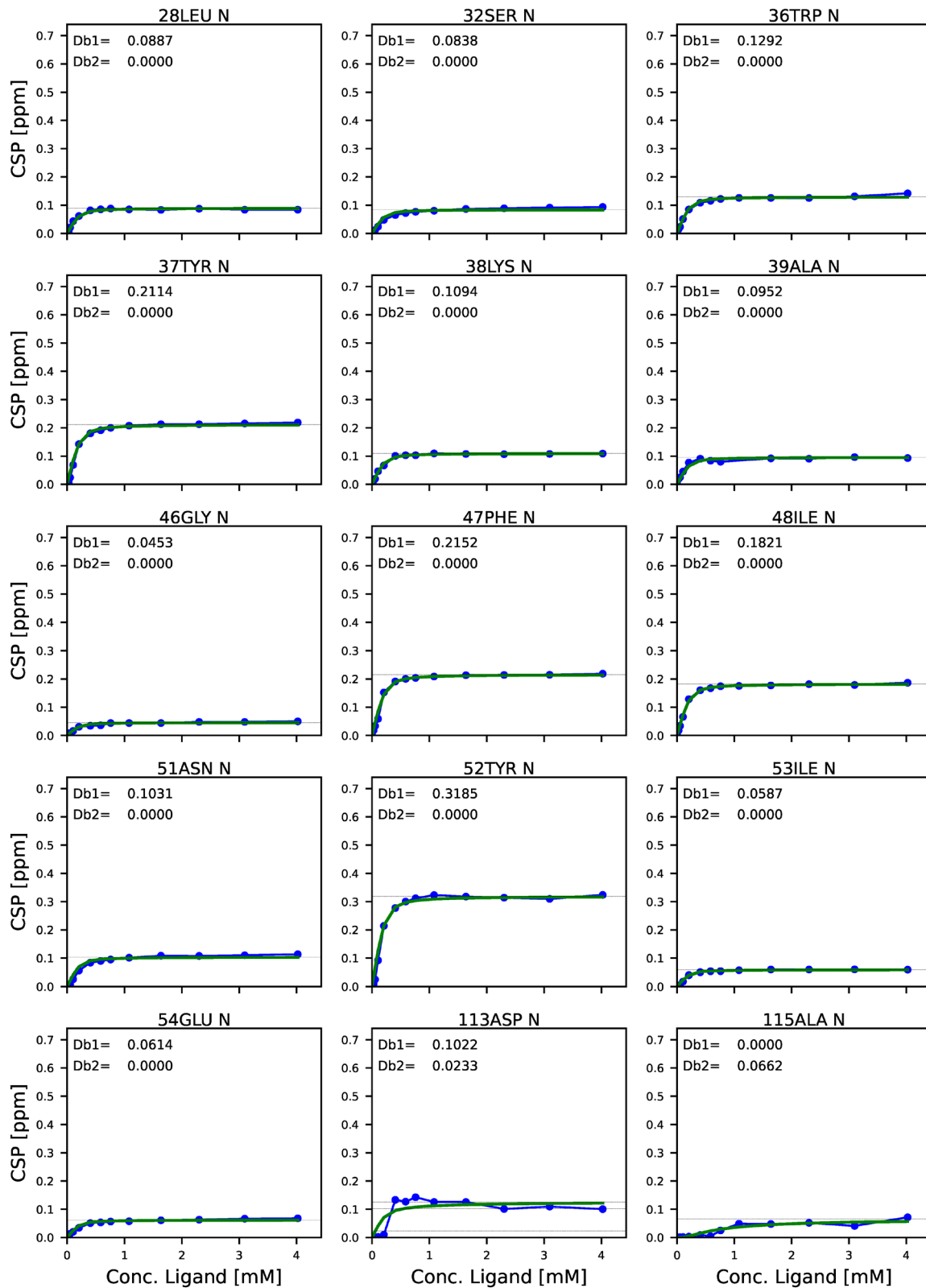


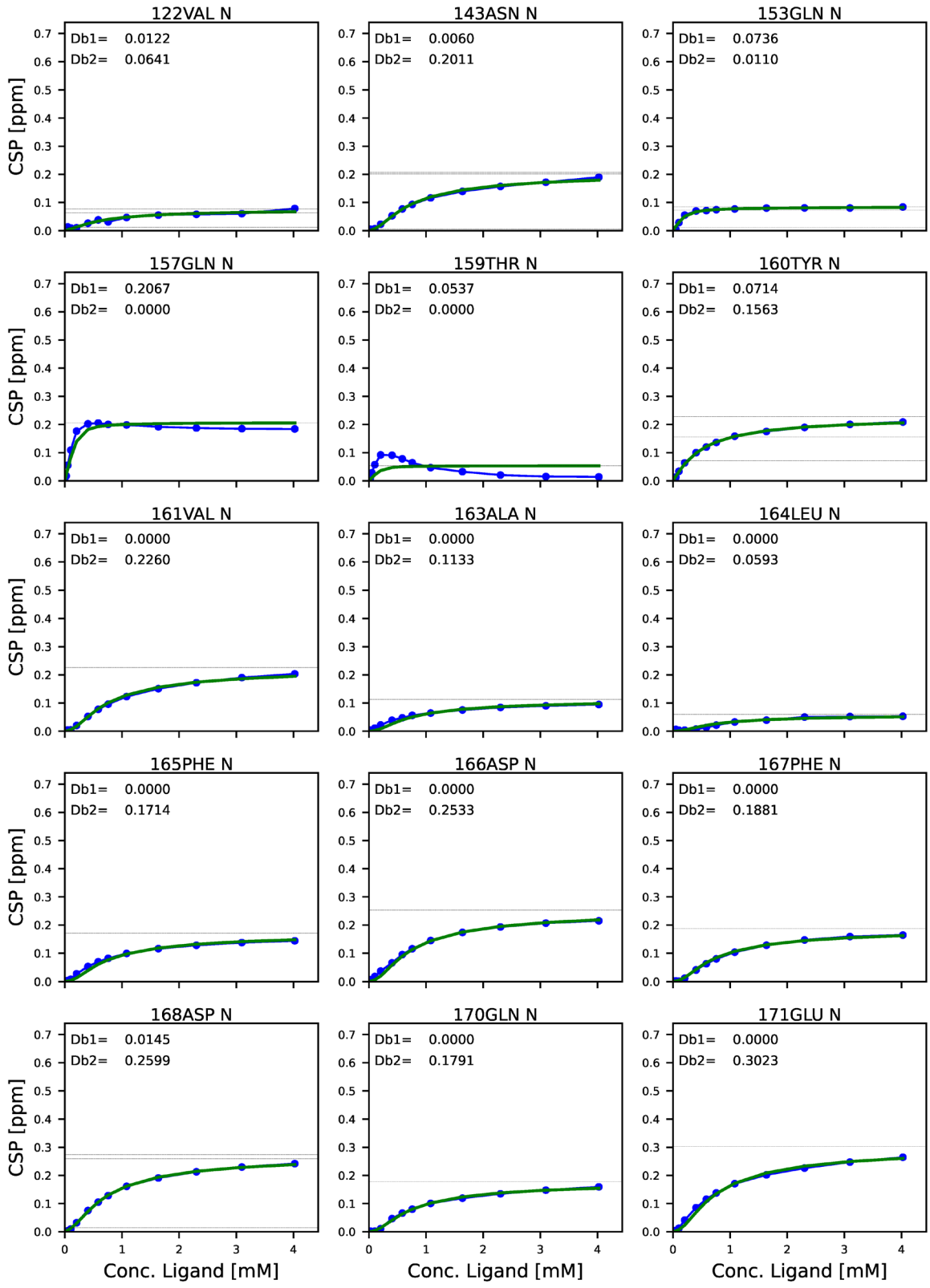


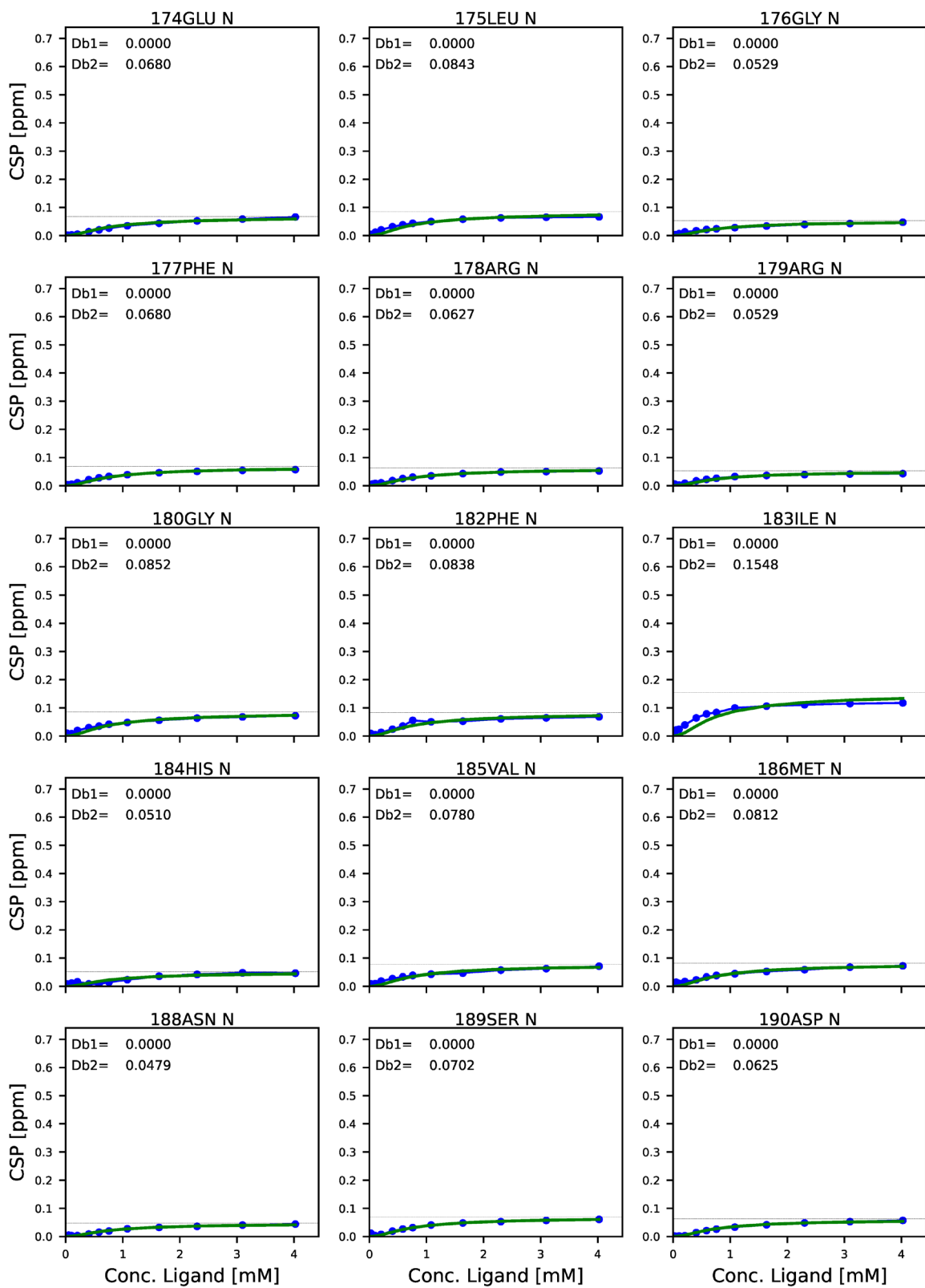


**Figure S12. Chemical shift perturbation profile mapped on each residue, against the concentration of ligand with SOS S5 PRM.** The blue dots and lines indicate CSP values at each titration point, while the green lines represent the theoretical curves calculated based on the two-site independent binding model.

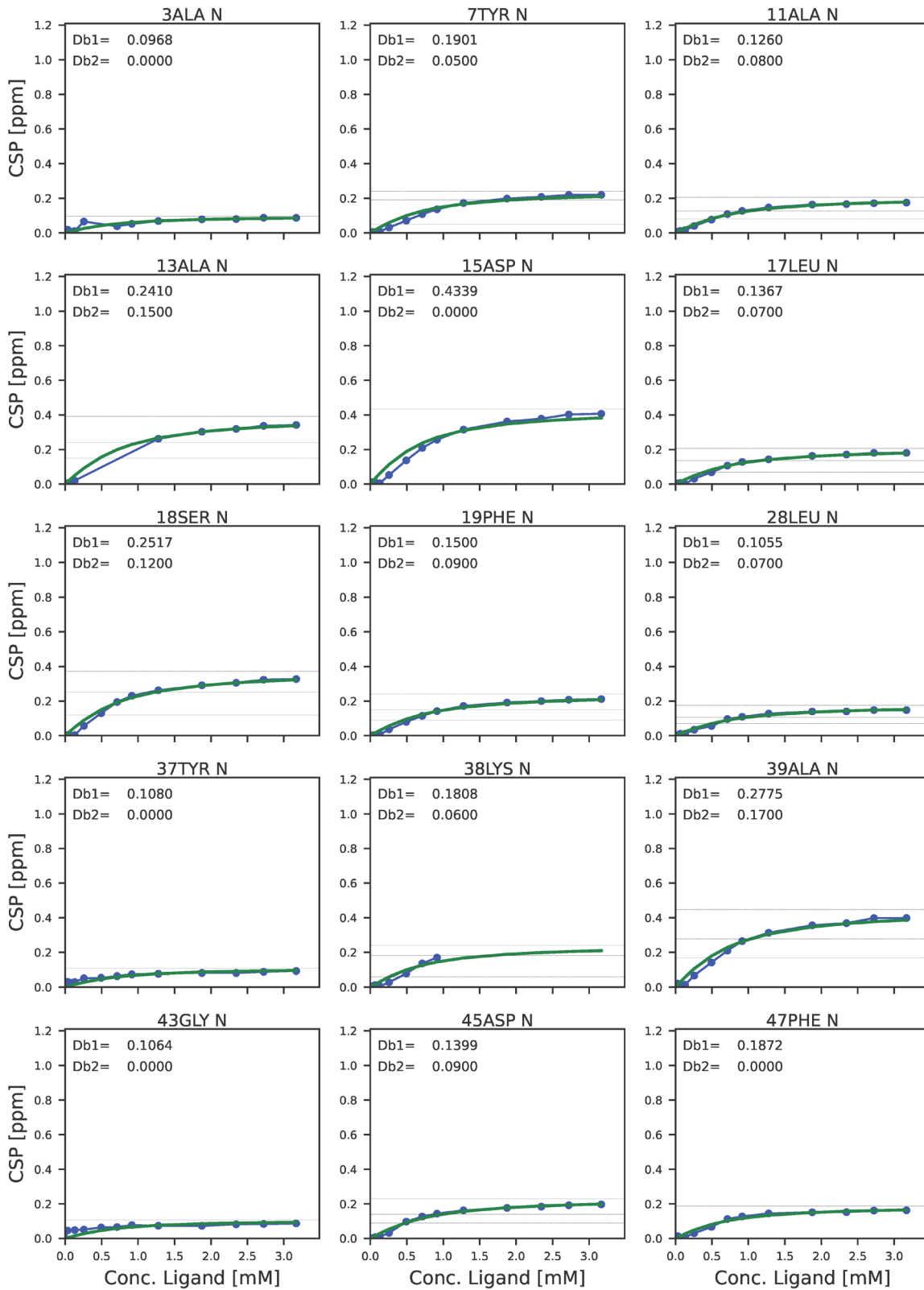


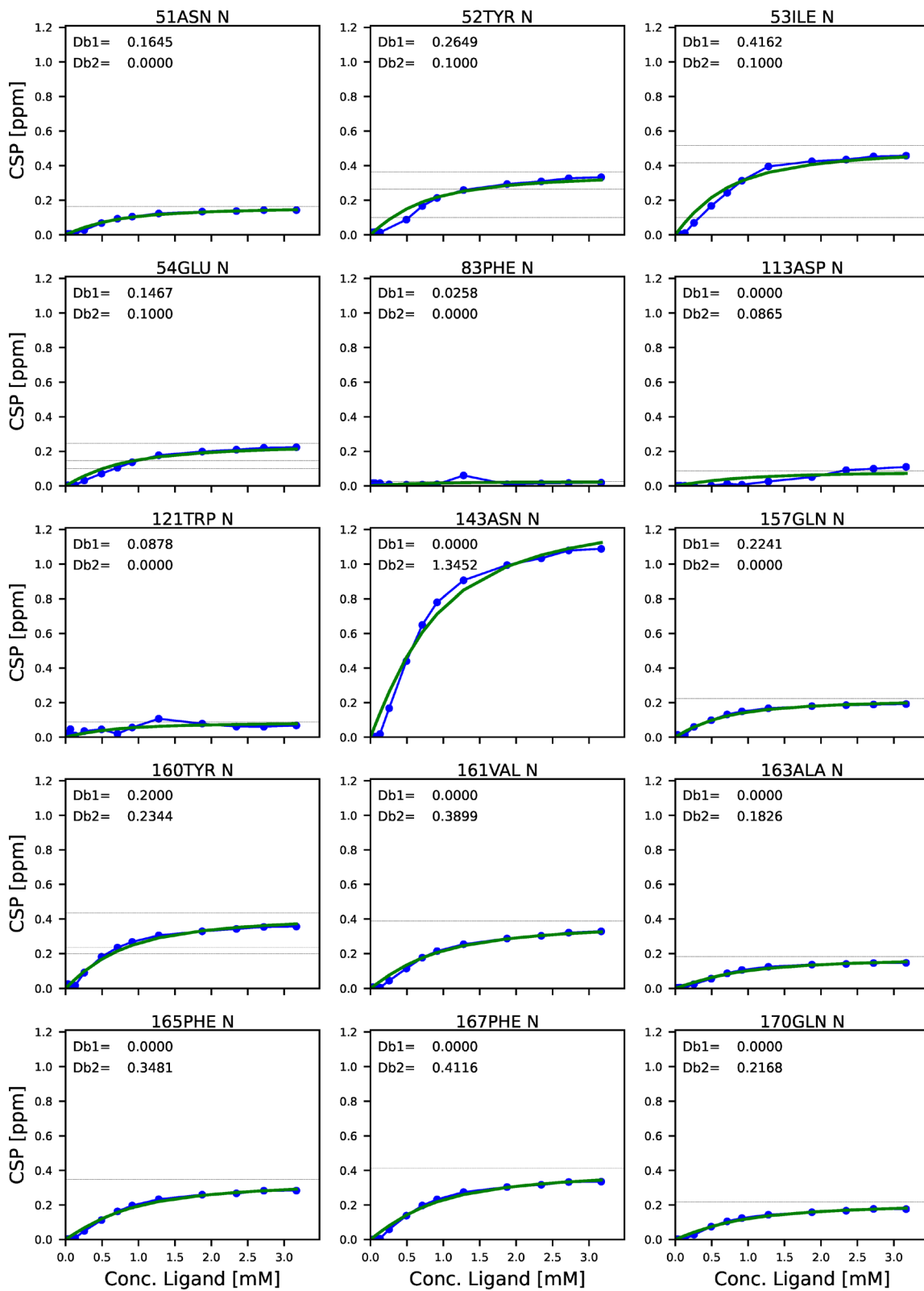




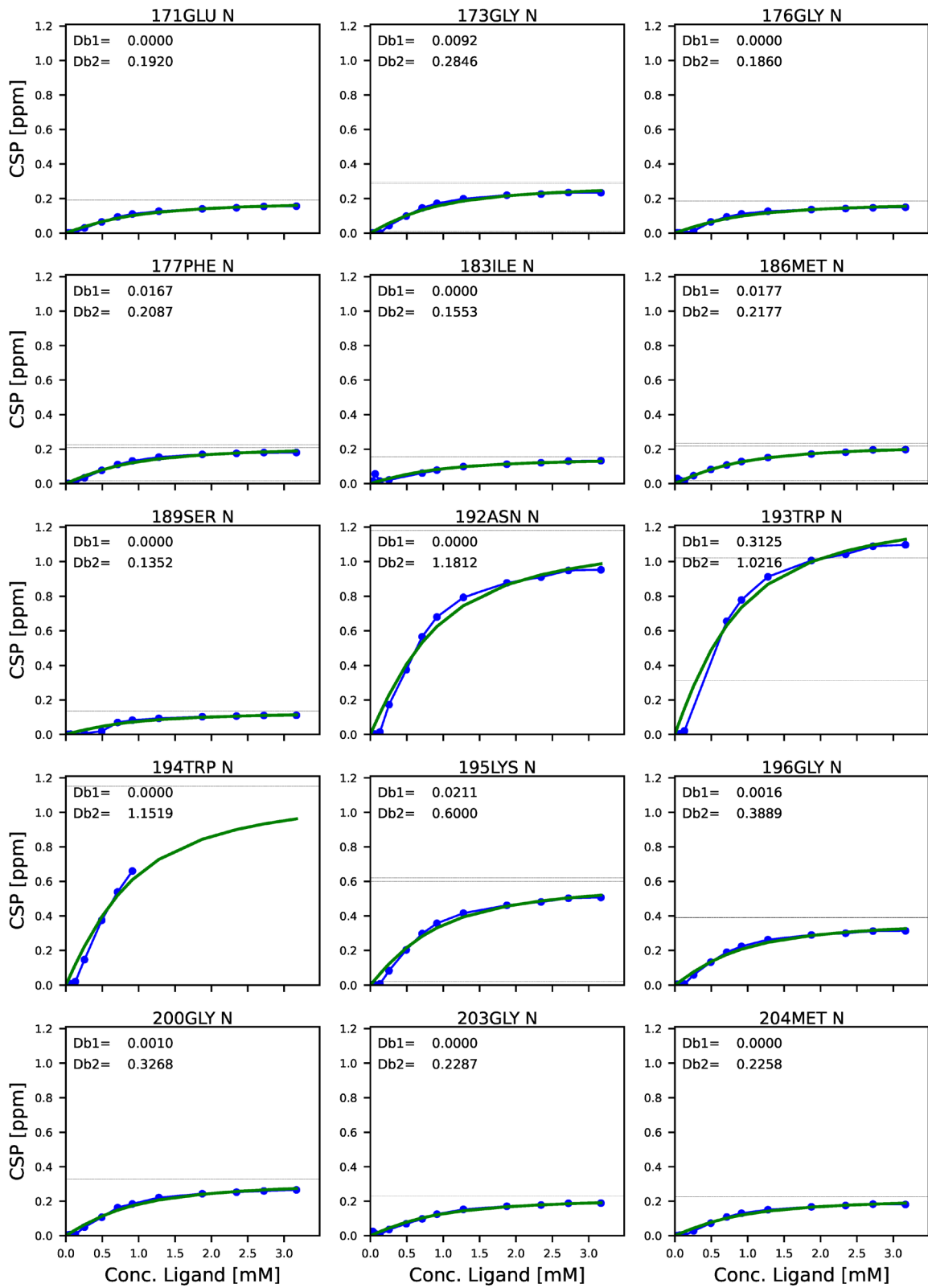


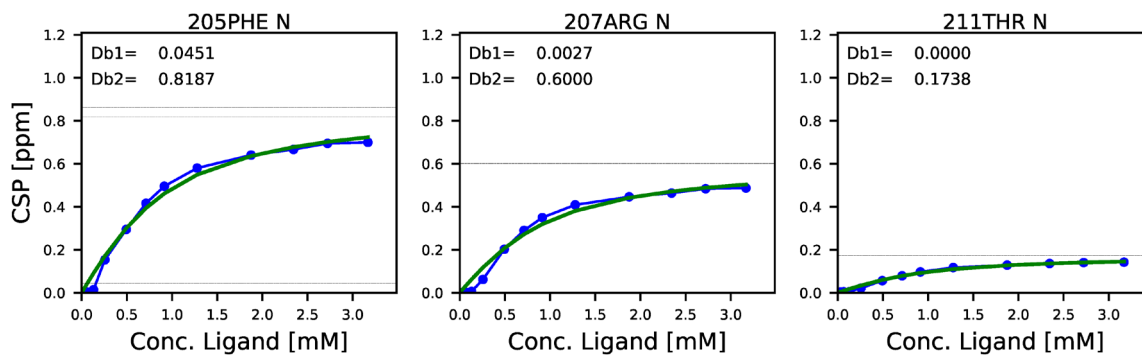
**Figure S13. Chemical shift perturbation profile mapped on each residue, against the concentration of ligand with SOS S9 PRM.** The blue dots and lines indicate CSP values at each titration point, while the green lines represent the theoretical curves calculated based on the two-site independent binding model.



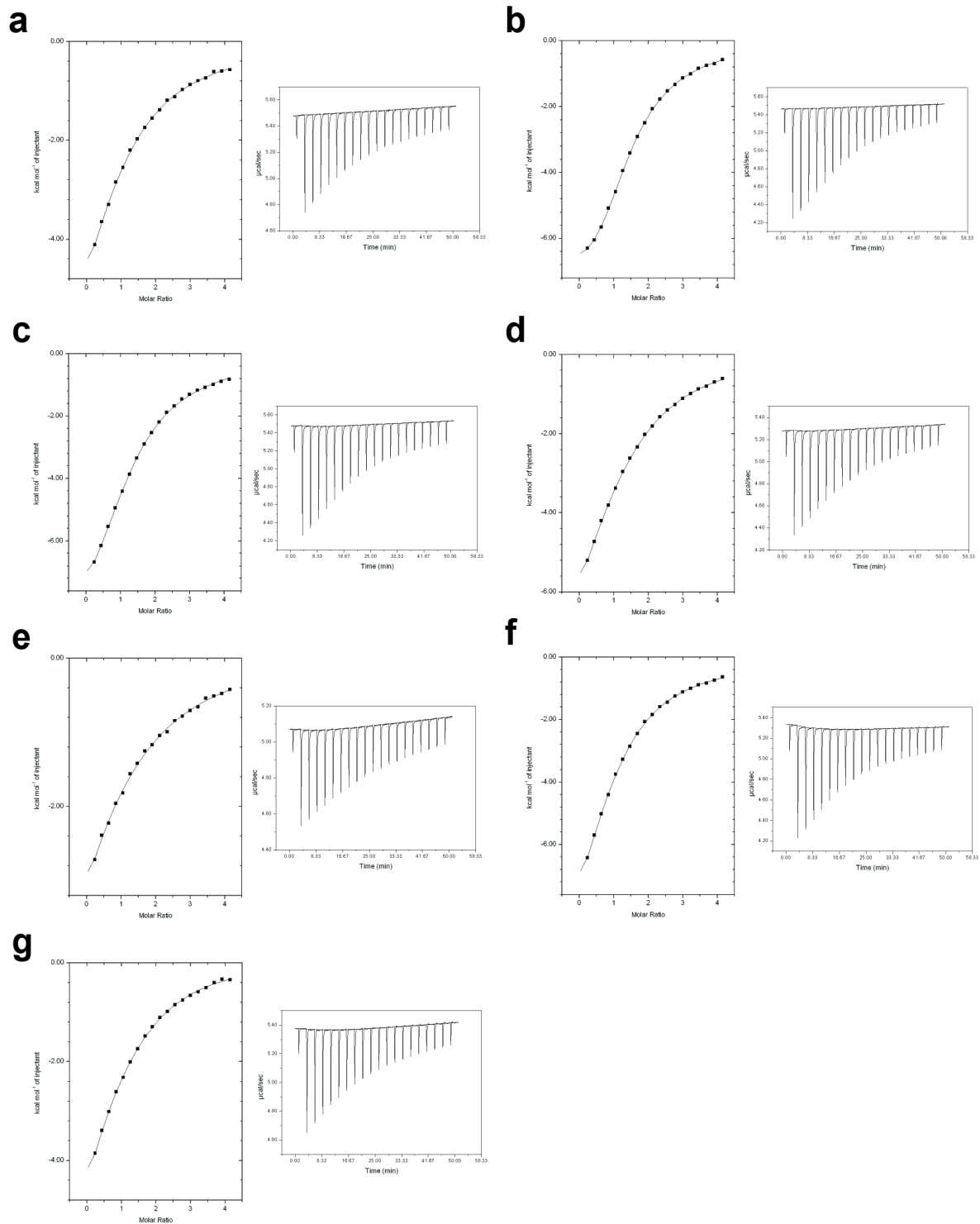






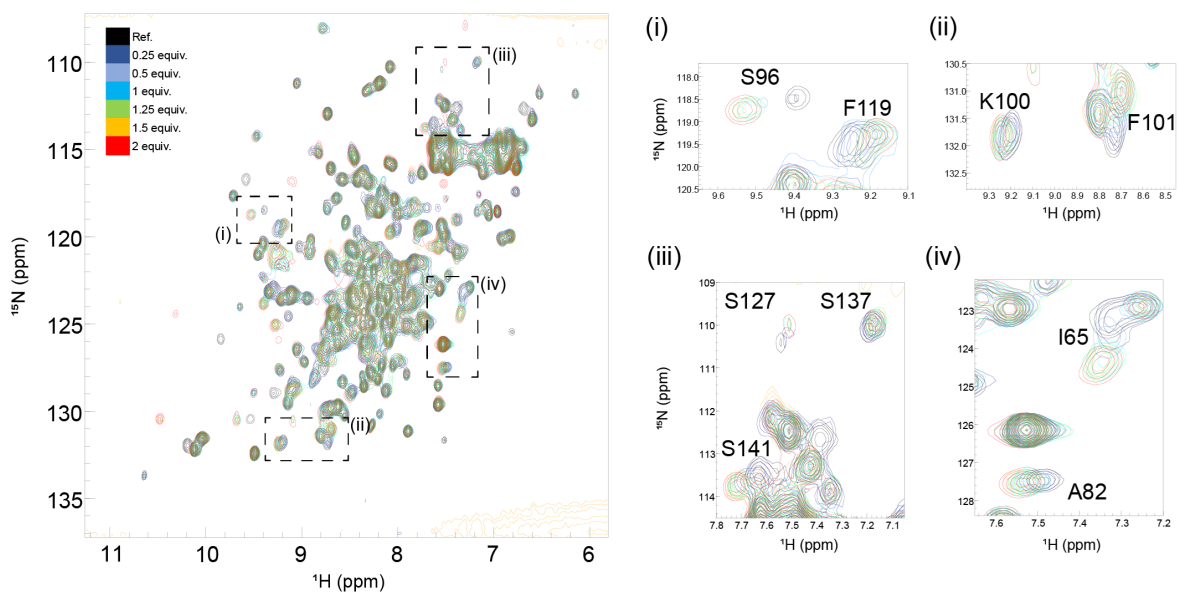


**Figure S14. Chemical shift perturbation profile mapped on each residue, against the concentration of ligand with SOS S10 PRM.** The blue dots and lines indicate CSP values at each titration point, while the green lines represent the theoretical curves calculated based on the two-site independent binding model.

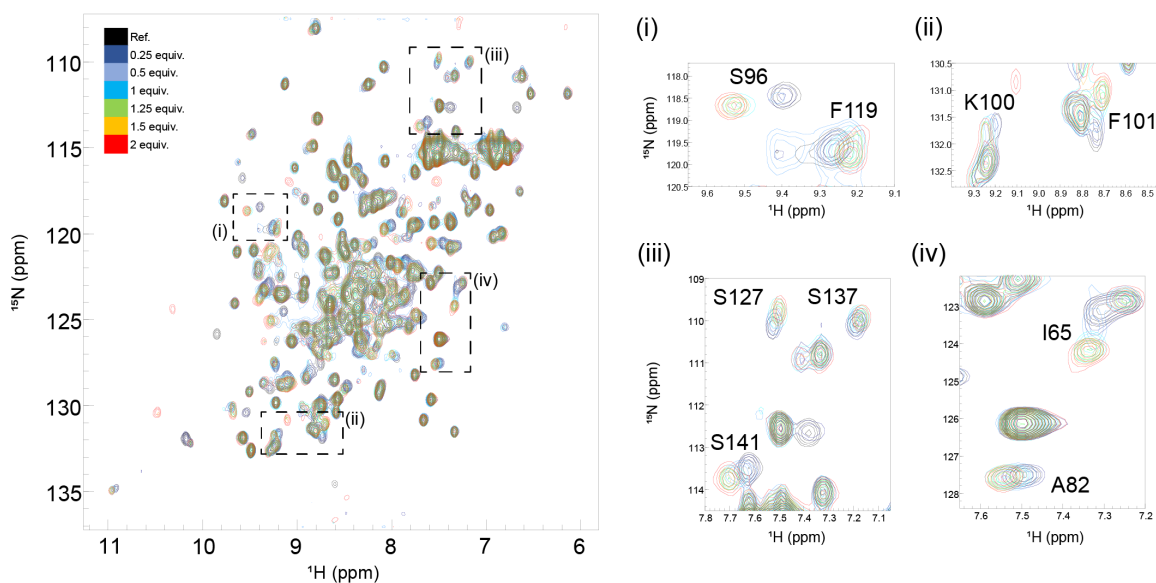


**Figure S15. ITC analysis for the binding of the full-length GRB2 to PRMs.** Figures a-g indicate ITC analysis for S2 (a), S4 (b), S5 (c), S6 (d), S8 (e) S9 (f), and S10 (g) PRMs,

respectively. Data with S1, S3, and S7 are not shown here because of no clear thermogram curves. The solid lines show the fit of the data to a "Two Set of Sites" model based on the binding of a ligand to a macromolecule, as incorporated in the Origin software.



**Figure S16. Overlays of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra from multipoint titrations of  $^{15}\text{N}$ -labelled GRB2 exclusively with the EGFR phosphorylated peptide (EpYINSQV).** The peptide concentration was increased stepwise (the protein: peptide molar ratio of 1:0.25, 1:0.5, 1:1, 1:1.25, 1:1.5, and 1:2). In this figure, the colour codes of  $^1\text{H}$ - $^{15}\text{N}$  correlation cross-peaks at each titration point, showing the molar ratio of GRB2: EGFR, are as follows: black (1:0); green (1:0.25); blue (1:0.5); yellow (1:0.75); magenta (1:1); cyan (1:1.25); brown (1:1.5); coral (1:2); red (1:4). Cross peaks that showed large chemical shift changes were annotated.



**Figure S17. Overlays of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra from multipoint titrations of  $^{15}\text{N}$ -labelled GRB2 with the EGFR phosphorylated peptide (EpYINSQV), in the presence of S4 PRM. The peptide concentration was increased stepwise (the protein: peptide molar ratio of 1:0.25, 1:0.5, 1:1, 1:1.25, 1:1.5, and 1:2). In this figure, the colour codes of  $^1\text{H}$ - $^{15}\text{N}$  correlation cross-peaks at each titration point, showing the molar ratio of GRB2: EGFR, are as follows: black (1:0); green (1:0.25); blue (1:0.5); yellow (1:0.75); magenta (1:1); cyan (1:1.25); brown (1:1.5); coral (1:2); red (1:4). Cross peaks that showed large chemical shift changes were annotated.**

**Table S1. Dissociation constants derived from ITC analysis for the full length GRB3 and isolated SH3 domains against SOS1 PRMs**

PRM	Sequence	ITC analysis			
		$K_D$ ( $\mu$ M) from the full length GRB2 <sup>a</sup>		$K_D$ ( $\mu$ M) from isolated SH3 <sup>b</sup>	
		$K_{D1}$	$K_{D2}$	$K_{D1}$	$K_{D2}$
S1	VPPPVPPIRRR (1148-1159)	NA	NA	—	—
S2	VPPPVPPIRRR (1148-1159)	65.1 $\pm$ 23.7	NA	—	—
S3	VPPPVPPIRRR (1148-1159)	NA	NA	—	—
S4	VPPPVPPIRRR (1148-1159)	5.3 $\pm$ 0.4	51.3 $\pm$ 1.7	39 $\pm$ 1	125 $\pm$ 13
S5	DSPPAOPPIRQPT (1177-1188)	15.9 $\pm$ 1.5	109.7 $\pm$ 6.5	56 $\pm$ 5	1396 $\pm$ 87
S6	DSPPAOPPIRQPT (1177-1188)	40.1 $\pm$ 6.5	97.3 $\pm$ 10.5	117 $\pm$ 2	1718 $\pm$ 33
S7	DSPPAOPPIRQPT (1177-1188)	NA	NA	—	—
S8	DSPPAOPPIRQPT (1177-1188)	122.9 $\pm$ 18.6	NA	—	—
S9	IAGPPVPPIRQST (1287-1298)	35.6 $\pm$ 4.2	640.0 $\pm$ 291.1	82 $\pm$ 1	1318 $\pm$ 44
S10	PKLPPKTYKREH (1303-1314)	71.5 $\pm$ 24.2	130.8 $\pm$ 54.4	—	—

<sup>a</sup>The dissociation constants were obtained by globally fitting ITC data to the “Two Set of Sites” model using MicroCal ITC-ORIGIN Analysis Software (Malvern Panalytical). <sup>b</sup>The dissociation constants are published data from ITC in the previous work using isolated SH3 domains {McDonald, 2009 #67}.