

Supporting Information

NMR Spectroscopy of the Main Protease of SARS-CoV-2 and Fragment-Based Screening Identify Three Protein Hotspots and an Antiviral Fragment

*François-Xavier Cantrelle⁺, Emmanuelle Boll⁺, Lucile Brier⁺, Danai Moschidi,
Sandrine Belouzard, Valérie Landry, Florence Leroux, Frédérique Dewitte, Isabelle Landrieu,
Jean Dubuisson, Benoit Deprez,* Julie Charton, and Xavier Hanouille**

anie_202109965_sm_miscellaneous_information.pdf
anie_202109965_sm_SI.xlsx

Author Contributions

F.C. Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing—review & editing: Supporting

E.B. Investigation: Equal; Methodology: Equal; Writing—review & editing: Equal

L.B. Formal analysis: Equal; Investigation: Equal; Writing—review & editing: Supporting

D.M. Formal analysis: Equal; Investigation: Equal; Writing—review & editing: Supporting

S.B. Conceptualization: Equal; Formal analysis: Equal; Investigation: Equal; Validation: Equal; Writing—review & editing: Equal

V.L. Formal analysis: Equal; Investigation: Equal; Writing—review & editing: Supporting

F.L. Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Writing—review & editing: Supporting

F.D. Investigation: Supporting; Writing—review & editing: Supporting

I.L. Resources: Supporting; Writing—review & editing: Equal

J.D. Funding acquisition: Equal; Writing—review & editing: Supporting

B.D. Funding acquisition: Equal; Project administration: Supporting; Writing—review & editing: Equal

J.C. Formal analysis: Equal; Investigation: Equal; Project administration: Equal; Writing—review & editing: Equal

X.H. Conceptualization: Lead; Funding acquisition: Lead; Investigation: Lead; Project administration: Lead; Supervision: Lead; Writing—original draft: Lead; Writing—review & editing: Lead.

SUPPORTING INFORMATION

Table of Contents**Experimental Procedures**

| | |
|--|-----|
| Expression and purification of 3CLp | 3 |
| NMR spectroscopy | 3 |
| NMR Assignment of 3CLp | 3 |
| Fragment library | 4 |
| NMR ligand-based primary screening | 4 |
| NMR protein-based secondary screening | 4 |
| NMR titration | 5 |
| In vitro inhibition of 3CLp | 5 |
| Jump dilution (reversibility) assay | 5 |
| Cellular antiviral and cytotoxicity assays | 5 |
| Ligand efficiency | 6 |
| X-Ray crystallography | 6 |
| Supporting Data | 6 |
| NMR spectroscopy of SARS-CoV-2 3CLp dimer | 6 |
| Analysis of the 3CLp binding hotspots | 7 |
| Supporting Figures | |
| Figure S1 | 8 |
| Figure S2 | 9 |
| Figure S3 | 10 |
| Figure S4 | 11 |
| Figure S5 | 12 |
| Figure S6 | 13 |
| Figure S7 | 14 |
| Figure S8 | 15 |
| Figure S9 | 16 |
| Figure S10 | 17 |
| Figure S11 | 18 |
| Figure S12 | 19 |
| Figure S13 | 19 |
| Adapted version of Figure 4 and Figure 5 for color-blind persons | 20 |
| Table S1 | 21 |
| Scheme S1 | 26 |
| Scheme S2 | 63 |
| Table S3 | 102 |
| References | 103 |

SUPPORTING INFORMATION

Experimental Procedures

Expression and purification of 3CLp

The DNA sequence coding for the SARS-CoV-2 3CLp (GenBank MN908947.3) was synthesized with codon optimization for expression in *Escherichia coli* (Genecust). The synthetic gene was inserted in our home-made pHIS-SUMO plasmid, which is based on the pET-24a (Merck) vector and contains a codon-optimized sequence coding for a N-terminal 6xHis-SUMO tag. The pHIS-SUMO-3CLp vector allows the expression, in *E. coli* BL21(DE3) (Novagen), of a 6xHis-SUMO-3CLp fusion protein. The bacteria were grown at 37°C in Luria-Bertani (LB) medium supplemented with Kanamycin. When A_{600nm} reached ~0.9, the temperature was lowered to 21°C and induction was carried out with 0.3 mM isopropyl- β -D-galactopyranoside for 12 hours. Cells were harvested and then lysed using homogenizer (Avestin Emulsiflex C-3) in lysis buffer (50mM Tris.Cl pH8.0, 300mM NaCl) supplemented with both DNaseI and RNaseA. In a first step, the 6xHis-SUMO-3CLp was purified using a HisTrap HP column (Cytivia) and eluted with an elution buffer containing 400mM imidazole. The fractions containing 6xHis-SUMO-3CLp were selected, pooled and dialyzed (cut-off 12-14 kDa) 2 hours at 4°C against 2 L of cleavage buffer (40 mM Tris-Cl pH7.5, 100 mM NaCl, 5 mM β -mercaptoethanol). The 6xHis-SUMO tag was cleaved SENP2 protease (His-tagged), added into the dialysis bag, and by dialyzing the sample over-night at 4°C against 2 L of fresh cleavage buffer. The 6xHis-SUMO tag and SENP2 protease were both removed when the sample was passed through a HisTrap HP column (Cytiva). The flow-through containing the native 3CLp was dialyzed (cut-off 6-8kDa) twice at 4°C against 3 L of NMR buffer (100 mM NaPi pH 6.8, 50 mM NaCl, 1mM EDTA). The native 3CLp was concentrated using a stirred cell with a 10kDa membrane (Amicon) and then supplemented with 2mM DTT-d₁₀. The purified 3CLp (320 μ M - 10.8 mg/mL) was flash frozen in liquid nitrogen and then stored at -80°C until used (aliquots of 500 μ L). The final yield was about 100mg per L of culture. The native sequence of the purified 3CLp (306 residues) was checked by MALDI-TOF analysis (Axima Assurance, Shimadzu).

The expression and purification of ²H,¹⁵N or ²H,¹⁵N,¹³C-labeled 3CLp is similar to the one for unlabeled 3CLp (see above) but with the following modifications. Bacteria were grown in a M9-based semi-rich medium (M9 medium supplemented with either ¹⁵NH₄Cl (1 g/L), D-glucose-d₇ (3g/L), Isogro-¹⁵N,D powder growth medium (0.5 g/L; Sigma-Aldrich) or ¹⁵NH₄Cl (1 g/L), D-glucose-¹³C₆-d₇ (3g/L), Isogro-¹³C,¹⁵N,D powder growth medium (0.5 g/L; Sigma-Aldrich), and kanamycin (25 μ g/mL). After purification the ²H,¹⁵N and ²H,¹⁵N,¹³C-labeled 3CLp were stored in NMR buffer (50mM NaPi pH 6.8, 40mM NaCl, 0.1 mM EDTA, 3mM THP (Tris(hydroxypropyl)phosphine), 5% D₂O). The ²H,¹⁵N- and ²H,¹⁵N,¹³C-labeled 3CLp were concentrated up to 247 μ M (8.37 mg/mL) and 530 μ M (17.9 mg/mL), respectively, flash frozen in liquid nitrogen and then stored at -80°C until used. The final yields for the doubly and triply labeled 3CLp were about 70 mg per L of culture.

The plasmid coding for the monomeric R298A mutant of SARS-CoV-2 3CLp was obtained by site-directed mutagenesis (Genecust) from the pHIS-SUMO-3CLp vector. The expression and purification of this mutant were performed as for the wild-type.

NMR spectroscopy

NMR experiments were performed either using Bruker Avance Neo 900 MHz or Bruker AvanceIII HD 600MHz NMR spectrometers equipped with a 5mm cryogenic triple resonance probe, CPTCI (¹H, ¹⁵N, ¹³C) and CPQCI (¹H, ¹⁵N, ¹³C, ¹⁹F), respectively. The 900MHz and 600MHz spectrometers are equipped with SampleJet and SampleCase sample changer, respectively. The proton chemical shifts were referenced using the methyl signal of TMSP (sodium 3-trimethylsilyl-[2,2,3,3-d₄]-propionate) at 0 ppm. Spectra were processed with the Bruker TopSpin software package 4.0.7 and 3.6.2 or with NMRpipe^[1]. Data analyses were done with CcpNmr Analysis^[2] and NMRFAM Sparky^[3] softwares.

NMR Assignment of 3CLp

NMR backbone assignments of SARS-CoV-2 3CLp were performed using several sets of TROSY-based 3D ¹H,¹⁵N,¹³C spectra acquired, with non-uniform sampling^[4], on an Avance Neo 900 MHz (Bruker) equipped with a cryogenic triple resonance probe, CPTCI (¹H, ¹⁵N, ¹³C). All NMR experiments were performed at 305 K in 5 mm Shigemitsu tubes. 2D ¹H,¹⁵N TROSY-HSQC and 3D ¹H,¹⁵N,¹³C TROSY-HNCACB, -HN(CO)CACB, -HNCO, -HN(CA)CO, -HN(CO)CA spectra were acquired on a dimeric ²H,¹⁵N,¹³C-3CLp sample at 530 μ M in NMR buffer (50 mM NaPi pH 6.8, 40 mM NaCl, 0.1 mM EDTA, 3 mM THP (Tris(hydroxypropyl)phosphine), 5% D₂O). 2D ¹H,¹⁵N TROSY-HSQC and 3D ¹H,¹⁵N,¹³C TROSY-HNCACB, -HN(CO)CACB spectra were acquired on a dimeric ²H,¹⁵N,¹³C-3CLp sample (530 μ M) bound to boceprevir (2 mM) in NMR buffer (50 mM NaPi pH 6.8, 40 mM NaCl, 0.1 mM EDTA, 3 mM THP (Tris(hydroxypropyl)phosphine), 3% DMSO-d₆ 5% D₂O). 2D ¹H,¹⁵N TROSY-HSQC and 3D ¹H,¹⁵N,¹³C TROSY-HNCACB, -HN(CO)CACB, -HNCO spectra were acquired on a monomeric ²H,¹⁵N,¹³C-3CLp R298A mutant sample at 440 μ M in NMR buffer (50 mM NaPi pH 6.8, 40 mM NaCl, 0.1 mM EDTA, 3 mM THP (Tris(hydroxypropyl)phosphine), 5% D₂O). To perform the NMR backbone assignments, we used a combined and integrated strategy including: classical sequential assignment using the experimental NMR data listed above, analyses of both the experimental NMR chemical shift perturbations upon boceprevir binding and the crystallographic structure of SARS-CoV-2 3CLp bound to boceprevir^[5] (PDB: 7c6s), chemical shift predictions performed with SHIFTX2^[6] on a

SUPPORTING INFORMATION

crystallographic structure of dimeric 3CLp (PDB: 7k3t, 10.2210/pdb7K3T/pdb), and previous NMR assignments for SARS-CoV 3CLp isolated N-terminal and C-terminal domains (Entries 17251 and 17911, respectively)^[7].

Backbone assignments of SARS-CoV-2 3CLp have been deposited in the Biological Magnetic Resonance Data Bank (Entry 50780).

Fragment library

The library used in this work contains 960 commercially available fragments. 640 and 320 fragments were purchased from Life Chemicals and Maybridge (ThermoFisher Scientific), respectively. Among the 960 fragments, 320 correspond to a highly hydrophilic subset, 320 are from a fluorine-containing (¹⁹F) subset, and 320 correspond to an unbiased subset. In total, our fragment library contains 427 fluorine fragments (427/960 = 44.5 %). The stock solutions of the fragments were at 100 mM concentration in DMSO-d₆.

NMR ligand-based primary screening

In order to optimize the acquisition time, the fragments from the library (960) were split into 192 cocktails of 5 fragments. Each NMR sample (530 μL in 5mm tubes) contains a cocktail of 5 fragments (377 μM each) in NMR buffer (100mM NaPi pH 6.8, 50mM NaCl, 0.2mM DTT-d₁₀, 2.5% DMSO-d₆, 7% D₂O). For the primary screening, these 192 fragment cocktails were analyzed by ¹H Water-LOGSY spectra^[8], at 293 K, both in the absence and in the presence of 25 μM of unlabeled 3CLp (1:15 ratio). 128 and 64 fragment cocktails were analyzed on the 900 MHz and 600 MHz NMR spectrometers equipped with CPTCI or CPQCI cryogenic probes, respectively. A control experiment was performed on 3CLp alone (without fragments) in the NMR buffer. The spectra were recorded with 32768 complex points, the mixing time was set to 1.7 sec, and the number of scans was 512 or 320 on the 900 MHz and 600 MHz spectrometers, respectively. Additionally, some of the fragment cocktails were also analyzed by ¹⁹F spectroscopy on the 600 MHz spectrometer equipped with a ¹⁹F-cryogenic probe (CPQCI). This includes the 64 cocktails prepared with the 320 fragments from the fluorine-containing (¹⁹F) subset (5 ¹⁹F fragments per tube), as well as 27 others cocktails that contain, at least, one ¹⁹F fragment from the others subsets. For 1D ¹⁹F spectra, the carrier frequency was set at -135 ppm, the spectral width was 192 ppm, the relaxation delay was 2 sec and 16 scans were recorded per experiment. The excitation pulse was a 600 μs 90° BURBOP pulse with 20 kHz B₁ field strength that was designed to cover a large excitation bandwidth (120 kHz)^[9], allowing detection of the full range ¹⁹F chemical shifts in a single scan. Broadening of the ¹⁹F signals can be observed for fragments that bind to the protein. ¹H decoupling using a waltz16 scheme was applied during acquisition. The spectra were processed and analyzed with Topspin 4.0.6 or 3.6.2 (Bruker). The identification of the hits was done using their ¹H and/or ¹⁹F NMR spectra provided by the suppliers.

NMR protein-based secondary screening

¹H,¹⁵N TROSY-HSQC spectra were acquired on 200 μL samples in 3 mm tubes containing 100 μM of ²H,¹⁵N-doubly labeled 3CLp sample in NMR buffer (50 mM NaPi pH6.8; 40 mM NaCl; 3 mM Tris(hydroxypropyl)phosphine; 3% DMSO-d₆; 5% D₂O) and 2 mM of the fragments (1 fragment per tube). A control experiment was performed on 3CLp alone (without fragments). 2D ¹H,¹⁵N-TROSY-HSQC spectra were acquired, at 305 K on a 900 MHz spectrometer equipped with a cryogenic probe, with 64 scans and 3072 and 128 complex points in the ¹H and ¹⁵N dimensions, respectively.

Upon binding of the fragments, both the ¹H and ¹⁵N combined chemical shift perturbations (CSP) and the signal intensity changes were analyzed. The CSP were calculated using the Equation 1 (Eq. 1), whereby (ΔδH) and (ΔδN) are the chemical shift perturbations in the ¹H and ¹⁵N dimensions, respectively^[10].

$$\text{CSP} = \sqrt{1/2 \cdot [(\Delta\delta\text{H})^2 + 0.2 \cdot (\Delta\delta\text{N})^2]} \quad (\text{Eq. 1})$$

The changes in signal intensity were assessed by calculating the relative signal intensity (rel. Intensity = I/I₀) along the 3CLp sequence where I and I₀ correspond to the intensity of the ¹H,¹⁵N TROSY-HSQC peak in the presence and in the absence of the fragment, respectively.

NMR titration

The affinity of the interaction between fragment **F01** and 3CLp was determined using NMR spectroscopy. A series of ¹H,¹⁵N TROSY-HSQC spectra were acquired on 100μM ²H,¹⁵N 3CLp with increasing amounts of fragment **F01** (0, 30, 100, 250, 600 and 1000 μM). The 2.5 % DMSO-d₆ concentration was kept constant in all experiments. The ¹H and ¹⁵N combined chemical shift perturbations (CSP) upon **F01** binding were calculated using equation 1 (see above). For determination of the dissociation constant (K_D), the CSP (ppm) were plotted as a function of the molar ratio [F01]/[3CLp] and the data were fitted to a curve (Equation 2, Eq. 2), corresponding to a 1:1

SUPPORTING INFORMATION

interaction, where $\Delta\delta_{\max}$ is the maximum value for CSP, X corresponds to the [F01]/[3CLp] ratio, and [3CLp] is the ^2H , ^{15}N 3CLp concentration.

$$\text{CSP} = \frac{\Delta\delta_{\max}}{2} \times \left(1 + X + \frac{K_D}{[3\text{CLp}]} - \sqrt{\left(1 + X + \frac{K_D}{[3\text{CLp}]}\right)^2 - 4 \times X}\right) \quad (\text{Eq. 2})$$

The K_D value presented corresponds to the mean (\pm S.D.) calculated over 18 different 3CLp NMR resonances (see **Figure S11**).

In vitro inhibition of 3CLp

The *in vitro* inhibition of the 3CLp enzymatic activity by fragment **F01** was assayed using a quenched FRET (Fluorescence Resonance Energy Transfer) assay with the {Dabcyl}-KTSAVLQSGFRKM-{Glu(Edans)} peptide as substrate. The assay was performed in buffer 50 mM HEPES, 0.1 mg/mL BSA, 0.01% TRITON and 2 mM GSH at pH 7.5. The 3CLp (15 nM) was incubated with the fragment **F01** (0.98 to 980 μM) or without (positive control), for 60 minutes at room temperature prior to addition of the peptide substrate (10 μM). Final concentration of DMSO did not exceed 1%. Then the reaction progress was monitored, every 90 seconds, for 30 minutes by measuring fluorescence intensity, with a Victor 3V instrument (Perkin-Elmer), using excitation and emission wavelengths of 340(25) nm and 535(25) nm, respectively. The initial velocities were used in calculations to determine the IC_{50} of **F01**. IC_{50} value was calculated from concentration-response curves by a nonlinear regression analysis at four parameters (using XL fitTM 5.2.0.0. from IDBS (Guilford, United Kingdom) or GraphPad Prism 7 (San Diego, USA)).

Jump dilution (reversibility) assay

The following assay was used to test if the **F01** binding to 3CLp is reversible or not. Fragment **F01** was pre-incubated for 60 minutes at 10 times its IC_{50} with 100x 3CLp (1500 nM). Then, the incubate was quickly diluted 100-fold with substrate solution (10 μM) before measuring the fluorescence kinetics, as explained above. Final concentrations after dilution are $\text{IC}_{50}/10$ for **F01**, 15 nM for 3CLp, and 10 μM for the substrate peptide. In the same time, control experiments were performed with 60 minutes incubation of 3CLp (30 nM) with fragment **F01** at either 0.2x or 20x its IC_{50} , and with a 2-fold dilution with substrate peptide at 20 μM . Final concentrations, after dilution, are $\text{IC}_{50}/10$ or 10x IC_{50} for **F01**, 15 nM for 3CLp, and 10 μM for the substrate peptide. All the experiments were performed in triplicate.

Cellular antiviral and cytotoxicity assays

Vero-81 cells (ATCC, CCL-81) and Vero-E6 were grown at 37°C with 5% CO_2 in Dulbecco's modified eagle medium (DMEM, Gibco) supplemented with 10% heat-inactivated fetal bovine serum (FBS, Eurobio). SARS-CoV-2 virus (strain SARS-CoV-2/human/FRA/Lille_Vero-TMPRSS2/2020) was amplified in Vero 81 cells expressing TMPRSS2 as described previously^[11].

The antiviral activity of fragment **F01** was assessed using dose response experiments. Vero-81 cells were seeded in 48-well plates and infected 24h later at a MOI of 0.1 in presence of increasing concentration of fragment **F01** (from 0 to 1000 μM). 16h later, cell supernatants were collected for virus titration and cells were lysed in non-reducing Laemmli buffer. Proteins were separated onto a 10% SDS-polyacrylamide gel electrophoresis and transferred on nitrocellulose membrane (Amersham). Membrane-bound N proteins were detected with a rabbit polyclonal antibody (Novus) and a horseradish peroxidase-conjugated secondary antibody. Detection was carried out by chemoluminescence (Pierce) and signals were quantified by using the gel quantification function of ImageJ. The experiment was repeated 3 times in duplicates. Viral titers were also measured by the TCID₅₀/ml method using Vero-E6 cells. The N protein quantification was used in calculations to determine the EC_{50} of **F01**. The EC_{50} value of **F01** is based on the N protein quantification and was calculated from concentration-response curve by a nonlinear regression analysis at four parameters using GraphPad Prism 7 (San Diego, USA).

The cytotoxicity of fragment **F01** was assayed on Vero-81 cells using live imaging following both Hoechst 33342 and NucView 488 Caspase-3 staining. Briefly, Vero-81 cells were seeded in 384-well plate, and 24h later, fragment **F01** was added (from 0 to 500 μM) to the culture medium, as well as Hoechst 33342 and NucView 488 Caspase-3 substrate. Right after compounds addition, live imaging was performed for 20 hours using an In Cell Analyzer 6000 (GE Healthcare). The cytotoxicity was defined based on the ratio of the apoptotic cell population (NucView staining) and the total population (Hoechst staining). Cytotoxicity data were fitted to an exponential growth model using GraphPad Prism 7 (San Diego, USA).

Ligand efficiency

The ligand efficiency (LE) of fragment **F01** is 0.29-0.30 kcal.mol⁻¹.heavy atom⁻¹ (LE = $-2.303(\text{RT}/\text{N}) \times \log\text{KD}$ ^[12], or LE = $1.4 \times \text{pIC}_{50}/\text{N}$ ^[13], where R is the ideal gas constant, T is the temperature, N is the number of heavy atoms and IC_{50} is the half-maximal *in vitro* inhibitory concentration).

SUPPORTING INFORMATION

X-Ray crystallography

A 3CLp sample at 5mg/mL in storage buffer (50 mM Tris-Cl pH7.5, 20 mM NaCl, 1 mM EDTA, 1 mM DTT) was used for crystallogensis. Initial conditions were obtained in 96-well plate (sitting drops) with a Cybi-Disk robot and commercial screening kits. The crystallization conditions were further manually optimized in 24-well plates (hanging drops). Crystals with flower-shape were obtained in 0.2 M sodium formate, 20% PEG 3350 at room temperature. These crystals were crushed with a micro-tool to make a seed stock and new crystals were grown in the same condition using the microseeding technique with a cat whisker. The crystals were soaked in a cryoprotectant solution (0.2 M sodium formate, 20% PEG 3350, 10% glycerol, 5% DMSO) before freezing in liquid nitrogen. For the 3CLp in complex with **F01**, the 3CLp crystals were soaked for 4 hours in a solution containing 10 mM fragment F01, 5% DMSO, 0.2 M sodium formate, 20% PEG 3350 and then briefly soaked into 0.2 M sodium formate, 20% PEG 3350, 15% glycerol, 2 mM F01 and 2% DMSO before freezing.

X-Ray data were collected on the Proxima1 and Proxima2 beamlines^[14] of the SOLEIL synchrotron facility (Paris, France). The data collection was done remotely using the MXCuBE2^[15] software and the crystals were handled by a Staubli sample changer. The data were collected at 100K using an Eiger-X 16M or Eiger-X 9M (Dectris) detector. The data were processed with XDS^[16] (xdsme scripts from the synchrotron facility, <https://github.com/legrandp/xdsme>). The molecular replacement (using the PDB entry 7K3T, DOI: 10.2210/pdb7K3T/pdb) and the refinement steps were done using the CCP4i2 interface^[17] of the CCP4 program suite^[18]. The statistics for data collection and refinement are summarized in the **Table S2**.

The final models and the structure factors corresponding to the 3CLp with its C145 oxidized, the 3CLp bound to fragment **F01** (N-(5-chloropyridin-2-yl)-3-oxo-2,3-dihydro-1H-indene-1-carboxamide) have been deposited in the Protein Data Bank as entries 7NTS and 7P51, respectively.

Supporting data

NMR spectroscopy of SARS-CoV-2 3CLp dimer

Starting from an efficient expression system of SARS-CoV-2 3CLp, in which the protease is expressed as a fusion protein including a N-terminal 6xHis-SUMO tag, we produced and purified samples with different isotopic labeling schemes ($u\text{-}^2\text{H}, ^{15}\text{N}$; $u\text{-}^2\text{H}, ^{15}\text{N}, ^{13}\text{C}$) to study by liquid-state NMR spectroscopy. During the purification, the fusion tag is cleaved and SARS-CoV-2 3CLp (306 residues) with both native N- and C-terminal ends is released (SI, Figure S1). This has been shown to be crucial for both the enzymatic activity and the proper dimerization of the protease. According to its dimeric high molecular weight ($2 \times 33.8 = 67.6$ kDa), we used per-deuterated protein samples (~ 500 μM), TROSY-based pulse sequences, an acquisition temperature of 305 K, and finally a 900 MHz NMR spectrometer equipped with a cryogenic probe. We obtained good quality $^1\text{H}, ^{15}\text{N}$ -TROSY HSQC spectrum, with ~ 280 resonances, on a $^2\text{H}, ^{15}\text{N}$ -3CLp sample (Figure 1) and thus moved on a $^2\text{H}, ^{15}\text{N}, ^{13}\text{C}$ -3CLp sample, at 530 μM , and recorded both 2D $^1\text{H}, ^{15}\text{N}$ TROSY-HSQC and 3D $^1\text{H}, ^{15}\text{N}, ^{13}\text{C}$ TROSY- HNCACB, -HN(CO)CACB, -HNCO, -HN(CA)CO, -HN(CO)CA spectra. Due to unfavorable magnetic relaxation properties, related to the protease dynamics and its molecular weight, some NMR signals were not observed and we thus had to record supplementary 3D NMR data on other triply labeled samples, including $^2\text{H}, ^{15}\text{N}, ^{13}\text{C}$ -3CLp (530 μM) bound to boceprevir (2 mM), and a monomeric $^2\text{H}, ^{15}\text{N}, ^{13}\text{C}$ -3CLp R298A mutant (440 μM), in order to reduce the protein dynamics and the molecular weight, respectively. To perform the NMR backbone assignments of SARS-CoV-2 3CLp, we used a combined and integrated strategy including: classical sequential assignment using the experimental NMR data listed above; analyses of both the experimental NMR chemical shift perturbations upon boceprevir binding and the crystallographic structure of SARS-CoV-2 3CLp bound to boceprevir^[5] (PDB: 7c6s); chemical shift predictions performed with SHIFTX2^[6] on a crystallographic structure of dimeric 3CLp (PDB: 7k3t); and previous NMR assignments for SARS-CoV 3CLp isolated N-terminal and C-terminal domains (BMRB entries: 17251 and 17911, respectively)^[7]. We assigned 183 proton amide correlations (183/293 non-proline residues, 63%), 2 NH from Tryptophan side-chains and further obtained 239/306 $C\alpha$, 207/280 $C\beta$ (non-glycine residues) and 234/306 C' chemical shifts, respectively (Figure 1). Most of the unassigned proton amides lie in the first two β -barrel domains or at the dimerization interface (SI, Figure S2). The backbone assignments of SARS-CoV-2 3CLp have been deposited in the Biological Magnetic Resonance Data Bank (BMRB entry 50780). Previous attempts to record multidimensional NMR data on SARS-CoV^[19] and more recently on SARS-CoV-2^[20] 3CLp have failed. These new NMR data open the field to a large range of future studies of the dimeric 3CLp in solution and at temperature (305K) close to physiological, an important parameter when considering dynamics. To assess the potential of our experimental system, we analyzed the 3CLp spectral perturbations, chemical shift perturbations (CSP) and peak intensity variations, upon binding of either boceprevir or GC376 (SI, Figures S3-S4). In both cases, the NMR perturbations induced are highest in the active site of the protease but also propagate further in its two catalytic domains, and even toward its C-terminal end with GC376. This last molecule induces spectral perturbations both at the active site and at the dimerization interface of 3CLp, the two regions of the protease that are targeted to develop inhibitors^[21-24]. Indeed, NMR perturbations may arise from ligand binding but also from the subsequent conformational changes in the protein. Moreover, in the presence of GC376, a few 3CLp NMR resonances split into two new resonances (SI, Figure S5). This might be due to the two conformations of the P3 moiety of the inhibitor that have been observed by Fu *et al.* in a crystallographic structure of the 3CLp:GC376 complex (PDB: 7d1m)^[5]. The split resonances include peaks from Val42, Asn142, Gln192 and Gly2. The later one showing that we can detect the conformational consequences, upon binding of an inhibitor in the active site of one protomer, in the N-terminal end of the

SUPPORTING INFORMATION

other protomer (SI, Figure S5). Interestingly, when using a SARS-CoV-2 R298A 3CLp monomeric mutant, we observed ~395 resonances in the 2D ^1H , ^{15}N TROSY-HSQC spectrum, which is ~115 more than in the wild-type dimer (SI, Figure S6) but also ~100 more than expected. This could be due to the two orientations of the C-terminal domain III that have been described for SARS-CoV 3CLp R298A^[25]. This needs to be further investigated but still highlights the potential for in-solution studies of the 3CLp where its conformational flexibility is not constrained by crystal packing. Thus, based on the assignment data we are able to not only detect ligand binding and map the binding site(s), but also to analyze the conformational rearrangement(s) upon binding throughout the whole dimer, providing essential molecular detail for medicinal chemistry inhibitor development.

Analysis of the 3CLp binding hotspots

The NMR CSPs pattern in Class I is similar to the ones observed in 3CLp following binding of either boceprevir or GC376 (SI, Figure S8), two potent inhibitors. The NMR CSPs we have observed upon binding of fragment **F01** (see Figure 4), a Class I hit, indeed nicely match with the active site residues that are involved in the binding of GC376 in the crystallographic structure of the 3CLp:GC376 complex (SI, Figure S9a). Even being smaller than GC376 (507.5 Da), **F01** (286.7 Da) induced high CSPs for NMR resonances corresponding to residues distributed all along the 3CLp active site cleft: Asn142, Gly143 and Glu166 in the pocket S1; Val42, Ile43, Asp48 and Leu50 in the pocket S2; Arg188 and Glu166 in pocket S3; and Val186, Thr190, Gln192 and Val171 in the pocket S4. Moreover, we can observe that upon binding of **F01**, the CSPs also propagate toward the dimerization interface of 3CLp (including Ser301, Gly302, Thr304, Phe305, Gly124 and Phe3), as we previously observed with GC376 (SI, Figure S4). This observation points out that these molecules not only act at the catalytic site but also induce conformational perturbations at distance.

These NMR data are fully supported by the analysis of the crystal structure of fragment **F01**-bound 3CLp that we solved (Figure 6 and SI, Figure S10 and Table S3; PDB: 7p51). **F01** binds in the active site of 3CLp. Its 3-oxo-2,3-dihydro-indene ring and its 5-chloro-2-pyridyl group occupy the S1 and S2 pockets of the 3CLp, respectively. Three hydrogen bonds (H-bond) are formed between **F01** and 3CLp. The first one involves the O atom of the ketone in the indene ring of **F01** that is electrophilic and could covalently react with the catalytic Cys145. This group, located in a key position of the active site, rather behaves as a H-bond acceptor and interacts with the side chain of His163. This later residue has indeed been shown to make H-bond with pyridine ring or other heterocycles from several ligands^[24]. The second H-bond is between the O atom of the carboxamide function in **F01** and the backbone proton amide of Glu166. Finally, the last H-bond, which is water-mediated, involves the H atom of the carboxamide function and the side chain of Asn142, located in the oxyanion loop around the S1 pocket. The binding of **F01** induces conformational changes in all the active site of 3CLp (Figure 6 and SI, Figure S10b). The 5-chloro-2-pyridyl group pushes on the Met49 side chain and induces the displacement of the α -helix (Ser46-Leu50) around the S2 pocket. It also triggers the re-orientation of the Met165 side chain toward the loop L3 that consequently undergoes conformational changes (residues Arg188 to Ala194). Around the S1 pocket, both Asn142 and Glu166 are slightly displaced. This last movement propagates to the 3CLp dimeric interface as the side chain of Glu166 in protomer A pushes the Ser1 located in protomer B. The analyses show that the NMR CSPs observed in the 3CLp spectrum upon **F01** binding indeed both match with the fragment binding site (S1 and S2 pockets) and the induced conformational changes (S3 pocket, loop L3 and dimerization interface) (SI, Figure S10c).

Thus, our data show that conformational plasticity^[26,27] and allosteric regulations^[21,23,28] within 3CLp can be studied using NMR spectroscopy, especially the tight interplay between substrate binding, active site conformation and dimerization. Indeed, it has been shown that following 3CLp dimerization, the Ser1 from the second protomer interacts with Glu166 of the first one and helps to stabilize the active conformation of the S1 pocket^[29,30]. We can thus monitor whether upon binding, the compounds affect 3CLp locally, only around the binding site, or have conformational impact through allosteric pathways.

The hits from Class II, such as **F30**, induced NMR CSPs that would correspond to their binding into the S2 and S3 pockets located in the domain I-side of the 3CLp substrate binding site, as SEN1269 that has been identified by Günther *et al.*^[23] from a repurposing molecule library (SI, Figure S9b). This molecule binds to S2 and induced the displacement of the short α -helix (Ser46-Leu50), for which we observed the highest CSPs upon binding of Class II hits (SI, Figure S8).

The NMR CSPs induced upon binding of the Class III hits, which includes **F15**, could be predicted to resemble the binding of x1086 and x1187, two fragments that have been shown to bind at the dimer interface of the SARS-CoV-2 protease^[21,24] in a hydrophobic pocket made by residues both in the N-ter (Met6, Phe8) and C-ter (Arg298, Gln299, Val303) ends (SI, Figure S9c). With **F15**, we also observed a high CSP for the resonance corresponding to Gln127, which is at the dimeric interface, and that has been shown to make a hydrogen bond with x1086.

SUPPORTING INFORMATION

Supporting Figures

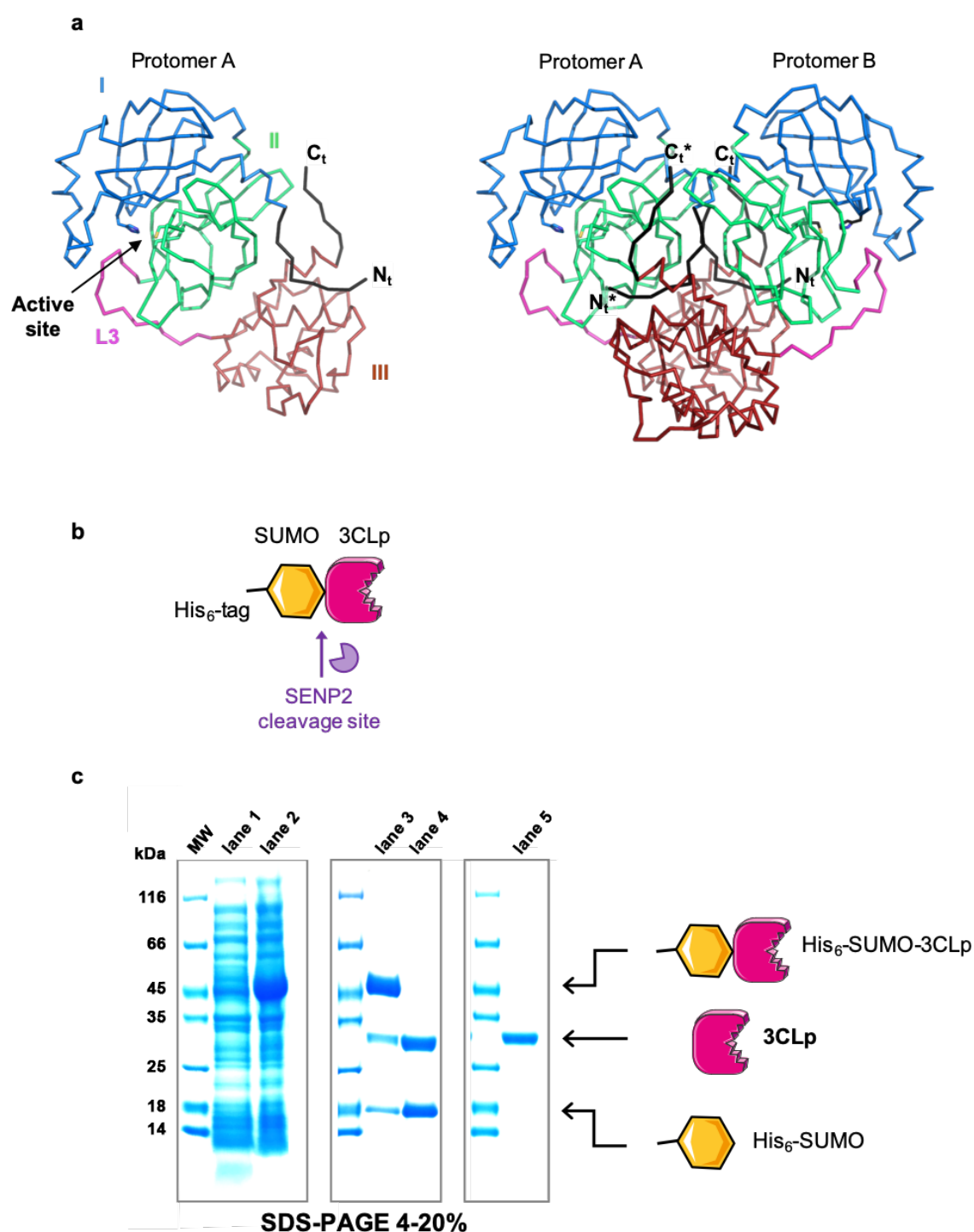


Figure S1. Expression and purification of SARS-CoV-2 3CLp. (a) Structure of 3CLp. *Left*, the domain organization of 3CLp is shown with domains I (residues 10-98), II (residues 99-184) and III (residues 199-300) coloured in blue, green and red, respectively. The linker (L3, residues 185-198) connecting domains 2 and 3 is shown in pink. *Right*, the homodimer assembly of 3CLp is shown and the N- and C-ter ends of each protomer are labeled. The extremities of protomer B are labeled with an asterisk. (b) The recombinant SARS-CoV-2 3CLp is recombinantly produced in *E. coli* as a fusion protein with a N-terminal His₆-SUMO tag, which can be cleaved thanks to a SENP2 cleavage site. (c) SDS-PAGE analyses of: **Lane 1**: *E. coli* BL21(DE3) cells before induction of protein production; **Lane 2**: *E. coli* BL21(DE3) cells after induction of protein production; **Lane 3**: Purified His₆-SUMO-3CLp fusion protein; **Lane 4**: Cleavage of the fusion protein by SENP2 protease; **Lane 5**: Purified SARS-CoV-2 3CLp.

SUPPORTING INFORMATION

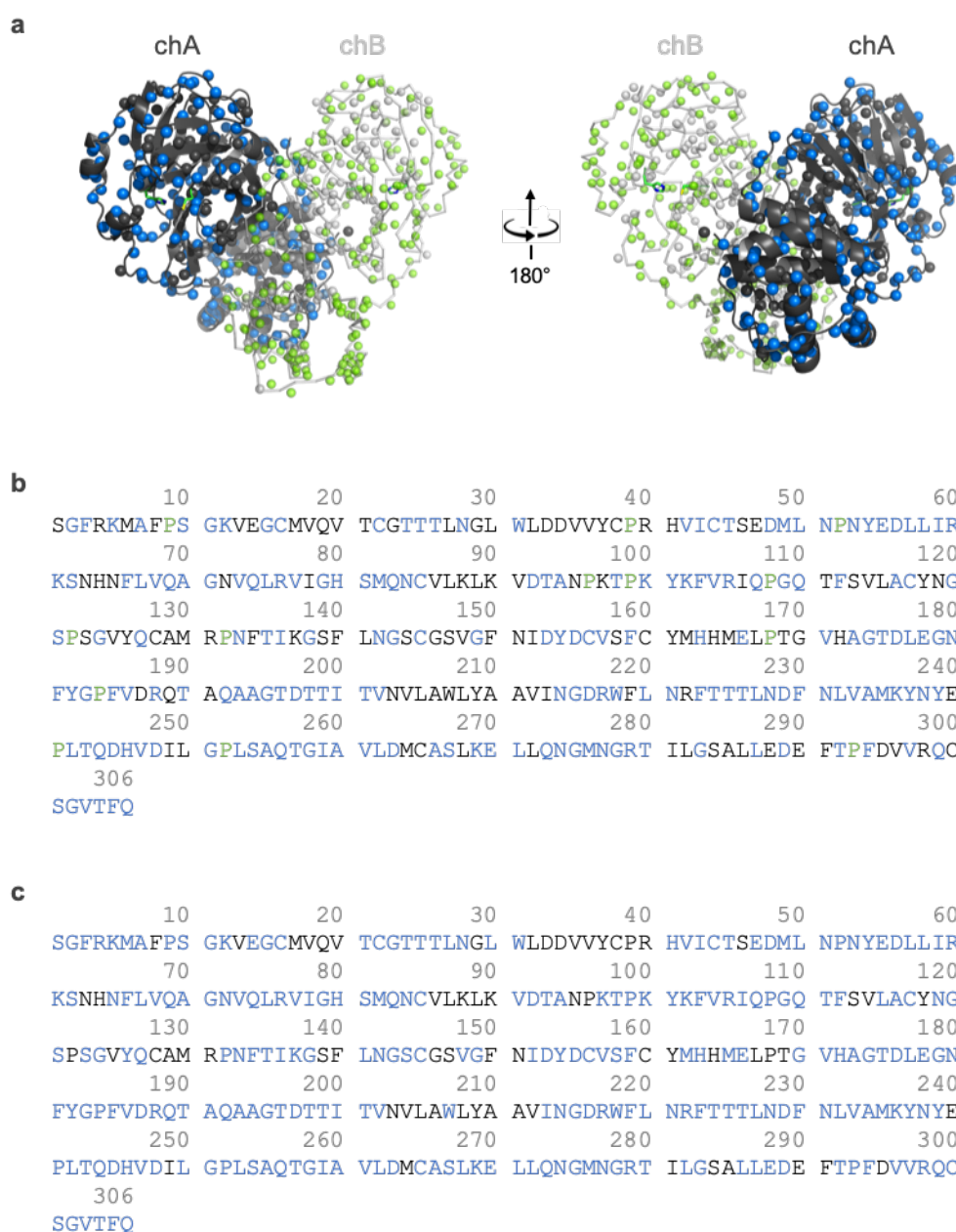


Figure S2. NMR assignment of SARS-CoV-2 3CLp. (a) The dimeric structure of 3CLp is shown as cartoon and ribbon representations for chain A (in dark grey) and chain B (in light grey), respectively. The backbone amide protons are shown as spheres and coloured in blue (chain A) or green (chain B) if their corresponding NMR resonances were assigned. Unassigned backbone amide protons are displayed in the same colour as the backbone. (b) Sequence of the SARS-CoV-2 3CLp where the residues for which their backbone proton amide correlation has been assigned are coloured in blue (183/293 non-proline residues). It corresponds to the structure shown in (a). The proline residues are shown in green. (c) Sequence of the 3CLp where the residues for which we assigned some of their NMR signals are coloured in blue (239/306 Ca, 207/280 Cb (non-glycine residues) and 234/306 C' chemical shifts).

SUPPORTING INFORMATION

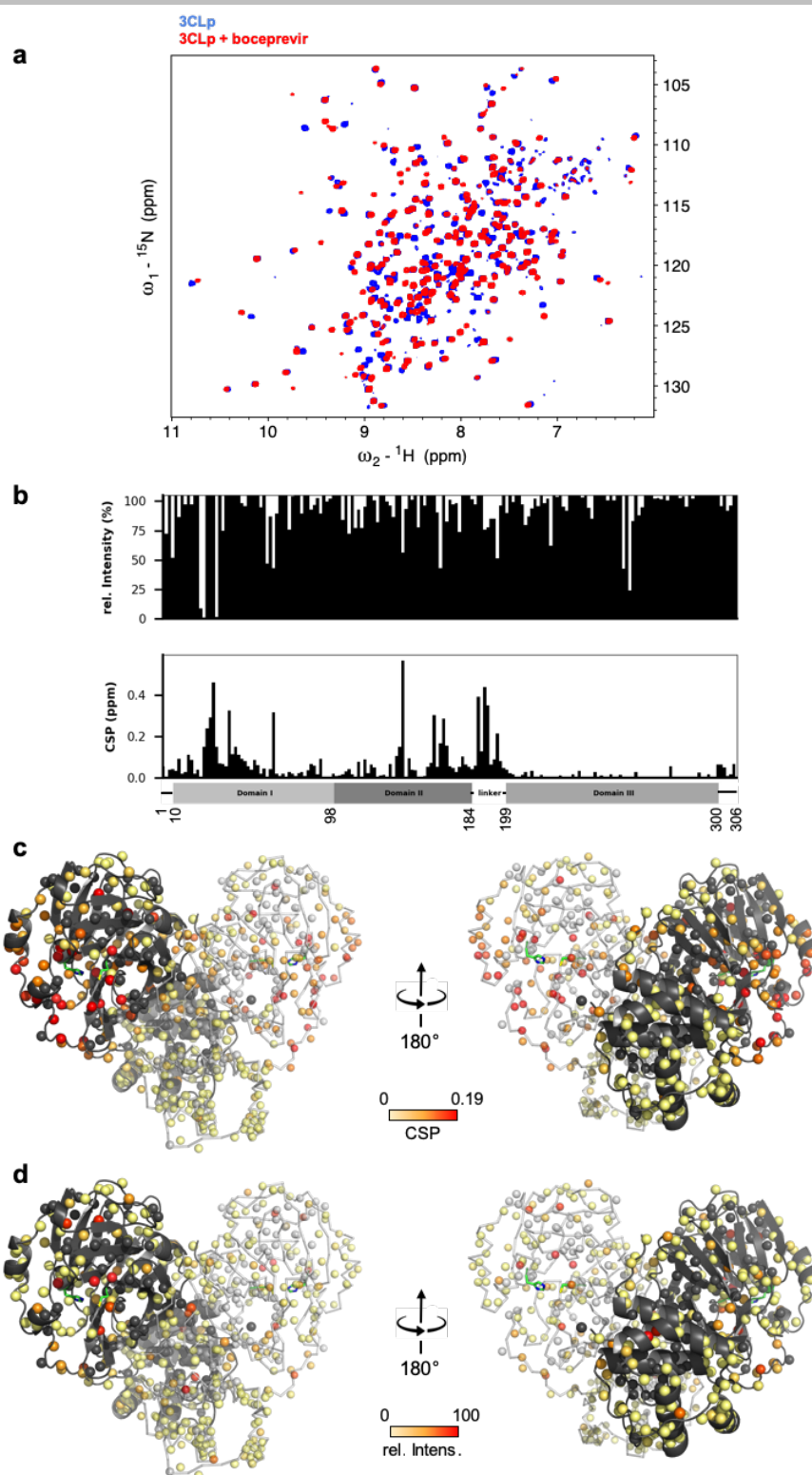


Figure S3. Boceprevir binding to $^2\text{H},^{15}\text{N}$ -3CLp. (a) Overlay of two 2D $^1\text{H},^{15}\text{N}$ -TROSY-HSQC spectra acquired on a $^2\text{H},^{15}\text{N}$ -3CLp sample (100 μM) in the absence (in blue) and in the presence (in red) of boceprevir (2 mM) (3% DMSO- d_6). The spectra were acquired at 305K on a 900 MHz spectrometer. (b) The spectral perturbations in the 3CLp spectrum induced upon boceprevir binding, corresponding both to the broadening of the resonances (rel. intens.) and to their ^1H and ^{15}N -combined chemical shift perturbations (CSP), are shown along the 3CLp sequence. (c, d) Structure of the SARS-CoV-2 3CLp dimer (PDB: 7k3t), with protomers A and B shown in black and light grey, respectively. Each small ball represents a proton amide of the protein and thus should correspond to a resonance in the $^1\text{H},^{15}\text{N}$ 2D spectrum. The ^1H and ^{15}N -combined CSPs and the relative intensities, from (a) and shown in (b), have been color coded (from light yellow to red) and are displayed on these balls in (d) and (c), respectively. Unassigned residues were kept in the original color of the protomer. The side chains of the two catalytic residues, His41 and Cys145, are shown in green.

SUPPORTING INFORMATION

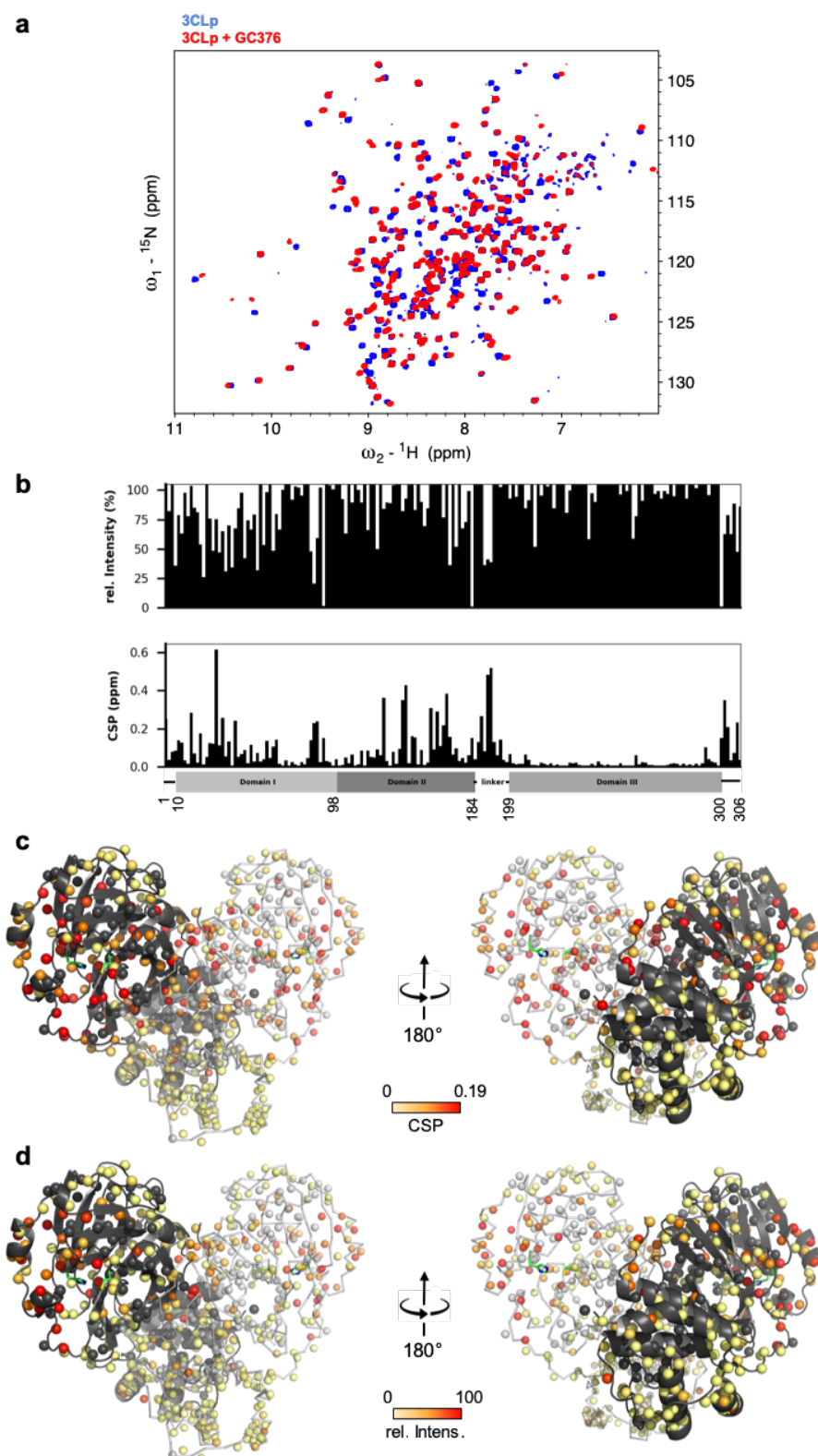


Figure S4. GC376 binding to $^2\text{H}, ^{15}\text{N}$ -3CLp. (a) Overlay of two 2D $^1\text{H}, ^{15}\text{N}$ -TROSY-HSQC spectra acquired on a $^2\text{H}, ^{15}\text{N}$ -3CLp sample (100 μM) in the absence (in blue) and in the presence (in red) of GC376 (2 mM) (3% DMSO- d_6). The spectra were acquired at 305K on a 900 MHz spectrometer. (b) The spectral perturbations in the 3CLp spectrum induced upon GC376 binding, corresponding both to the broadening of the resonances (rel. intens.) and to their ^1H and ^{15}N -combined chemical shift perturbations (CSP), are shown along the 3CLp sequence. (c, d) Structure of the SARS-CoV-2 3CLp dimer (PDB: 7k3t), with protomers A and B shown in black and light grey, respectively. Each small ball represents a proton amide of the protein and thus should correspond to a resonance in the $^1\text{H}, ^{15}\text{N}$ 2D spectrum. The ^1H and ^{15}N -combined CSPs and the relative intensities, from (a) and shown in (b), have been color coded (from light yellow to red) and are displayed on these balls in (d) and (c), respectively. Unassigned residues were kept in the original color of the protomer. The side chains of the two catalytic residues, His41 and Cys145, are shown in green.

SUPPORTING INFORMATION

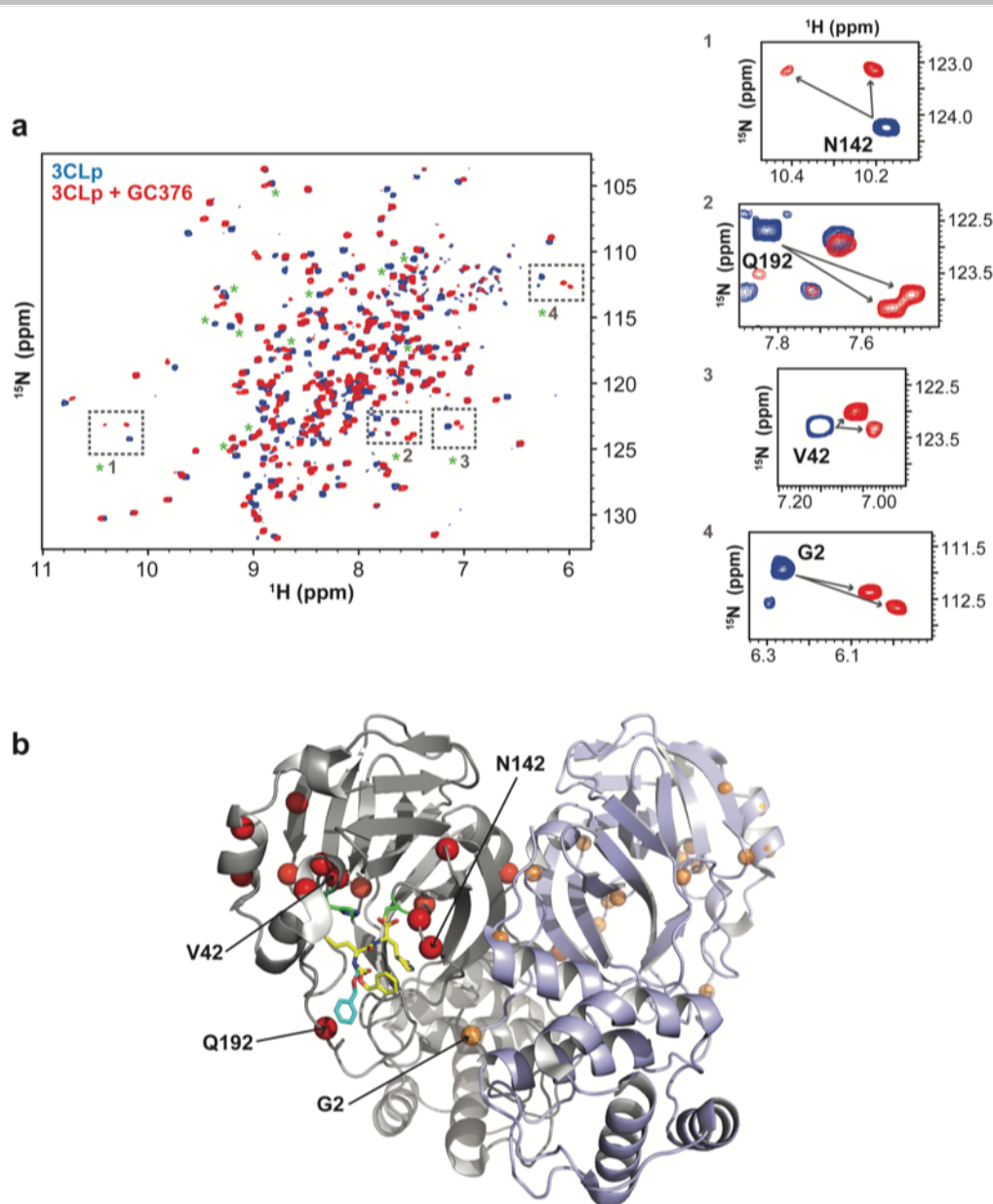


Figure S5. GC376 binding to ^2H , ^{15}N -3CLp. (a) Overlay of two 2D ^1H , ^{15}N -TROSY-HSQC spectra acquired on a ^2H , ^{15}N -3CLp sample (100 μM) in the absence (in blue) and in the presence (in red) of GC376 (2 mM) (3% DMSO- d_6). The spectra were acquired at 305K on a 900 MHz spectrometer. In the presence of GC376 several 3CLp resonances, highlighted with a green asterisk, split into two new resonances. The split resonances include Asn142, Gln192, Val42 and Gly2, that are shown in insets 1, 2, 3 and 4, respectively. (b) The split resonances in (a) are shown on the structure of the SARS-CoV-2 GC376-bound 3CLp dimer (PDB: 7d1m)^[5], with protomers A and B displayed in grey and violet, respectively. Each ball represents a proton amide of the protein for which a split resonance has been observed in the ^1H , ^{15}N 2D spectrum following GC376 binding. The red and orange balls correspond to proton amide in the protomers A and B, respectively. The black labels indicate the residues highlighted in the insets 1-4. The side-chains of the two catalytic residues, His41 and Cys145, are shown in green.

SUPPORTING INFORMATION

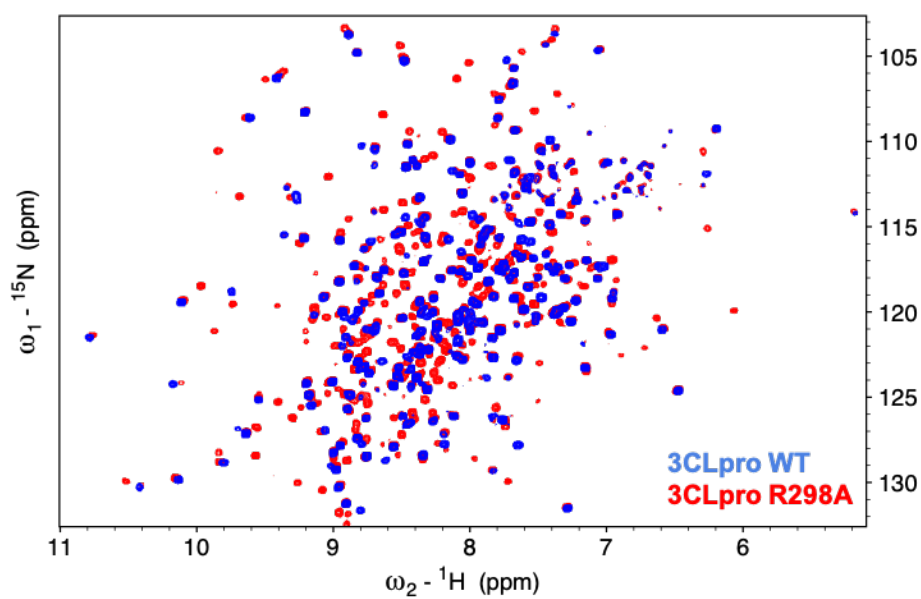


Figure S6. Comparison of the 2D NMR spectra of dimeric and monomeric SARS-CoV-2 3CLp. Overlay of two 2D ^1H , ^{15}N -TROSY-HSQC spectra acquired either on a dimeric ^2H , ^{15}N -3CLp WT sample (100 μM) or on a monomeric ^2H , ^{15}N -3CLp R298A mutant (100 μM). The spectra were acquired at 305K on a 900 MHz spectrometer. Whereas 293 resonances, corresponding to non-proline residues, could be expected, ~ 280 and ~ 398 can be observed in the spectra of 3CLp WT and 3CLp R298A mutant, respectively.

SUPPORTING INFORMATION

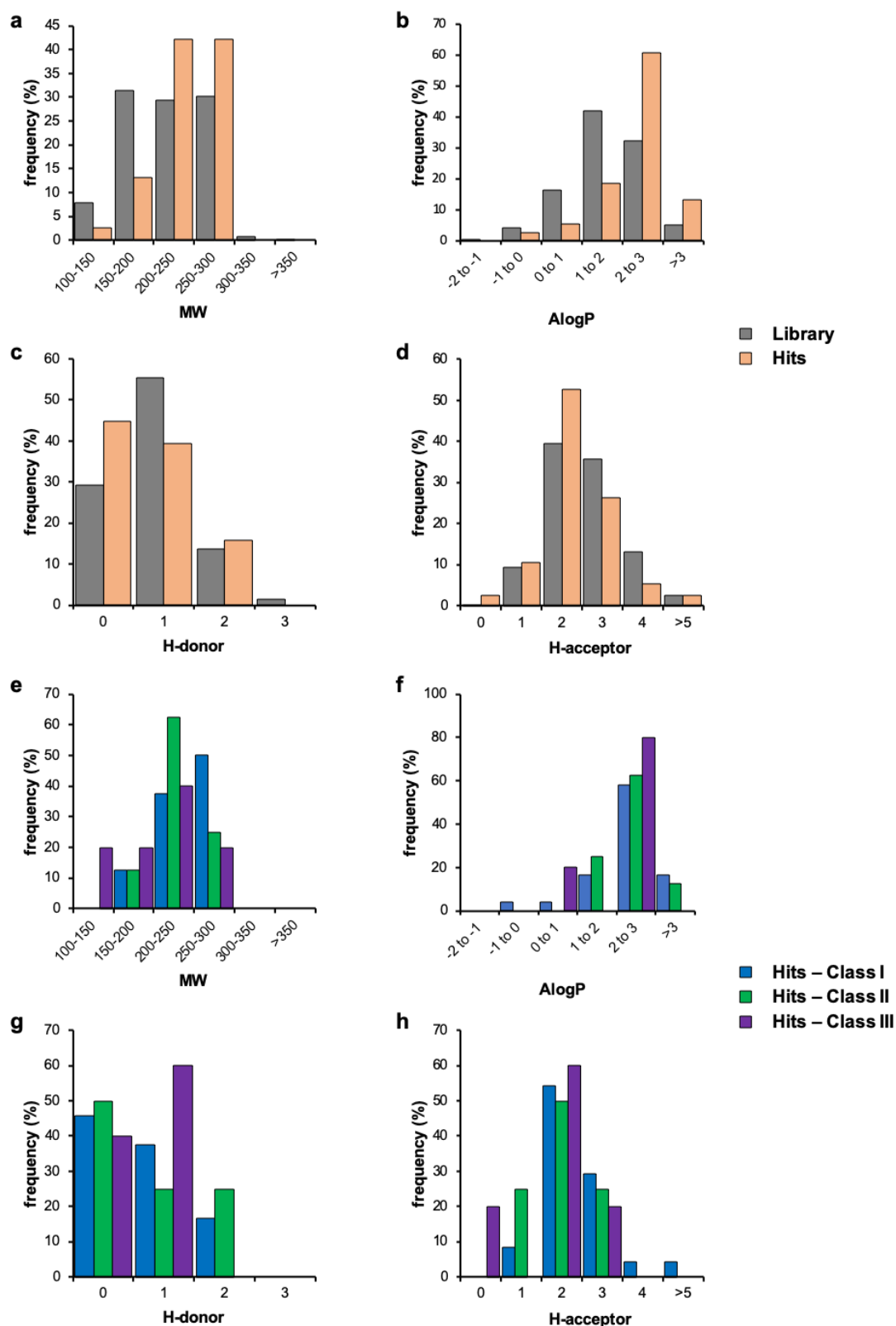


Figure S7. Fragment's chemical properties. (a-d) Distributions of (a) the molecular weight (MW, Da), (b) the lipophilicity (AlogP), (c) the number of H-bond donor, and (d) the number of H-bond acceptor of both the 960 fragments in the library used in this work (in grey) and the 38 fragment hits selected (in salmon). (e-h) Distributions of (e) the molecular weight (MW, Da), (f) the lipophilicity (AlogP), (g) the number of H-bond donor, and (h) the number of H-bond acceptor of the fragment hits grouped in Class I (in blue), Class II (in green), and Class III (in violet) according to the NMR CSPs (see Figure S8). The chemical properties have been calculated using Pipeline Pilot 8.5 (Dassault Systems).

SUPPORTING INFORMATION

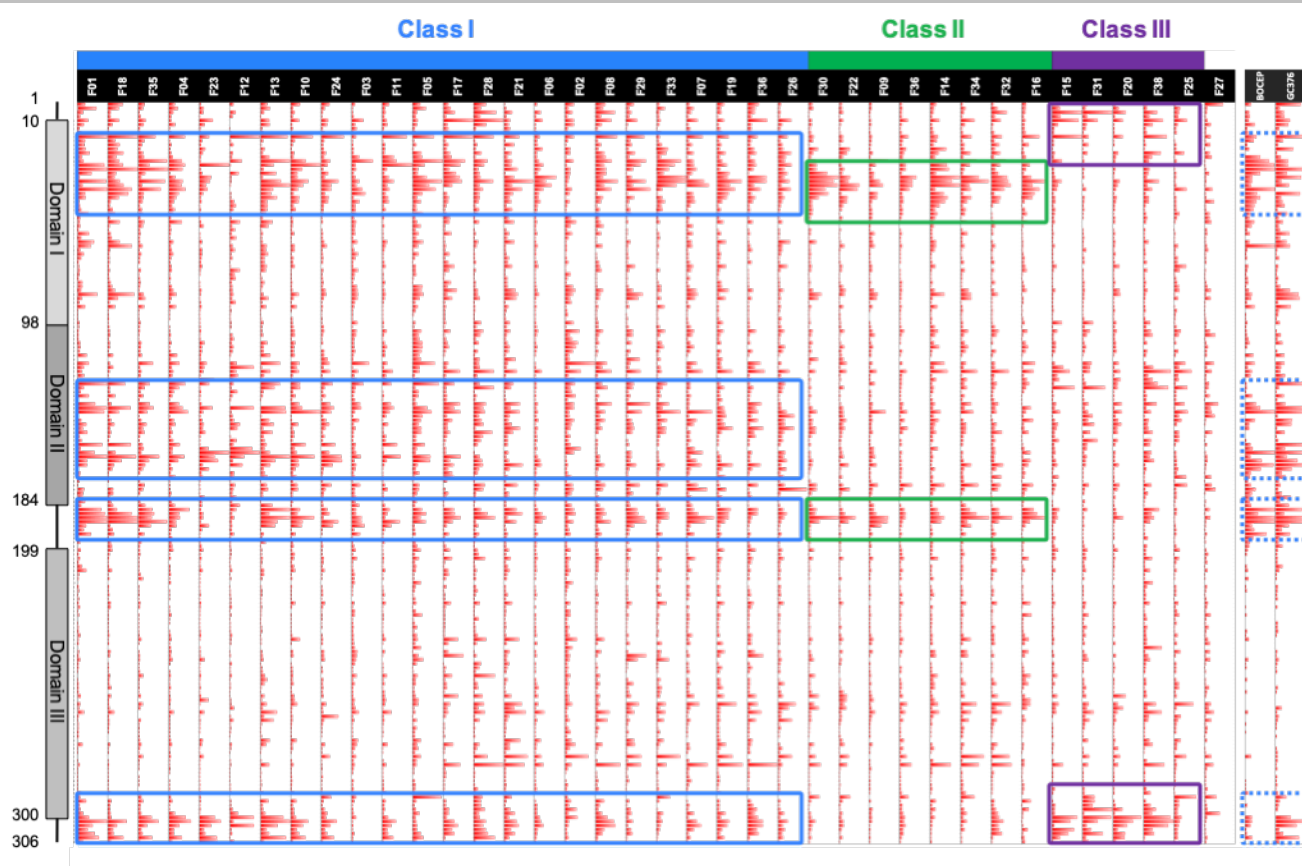


Figure S8. NMR chemical shift perturbations (CSPs) induced upon fragments binding grouped into three classes. The CSPs induced in the ^1H , ^{15}N -TROSY-HSQC spectrum of 3CLp following binding of each of the 38 fragment hits are shown in red along the 3CLp sequence. The organization of 3CLp domains is shown on the left, with their boundaries numbered according to the amino acid sequence. The CSP patterns corresponding to the Class I, Class II, and Class III are highlighted in blue, green and violet, respectively. On the right, the CSP patterns corresponding to the binding of two reference molecules, boceprevir and GC376, are shown (in dotted blue).

SUPPORTING INFORMATION

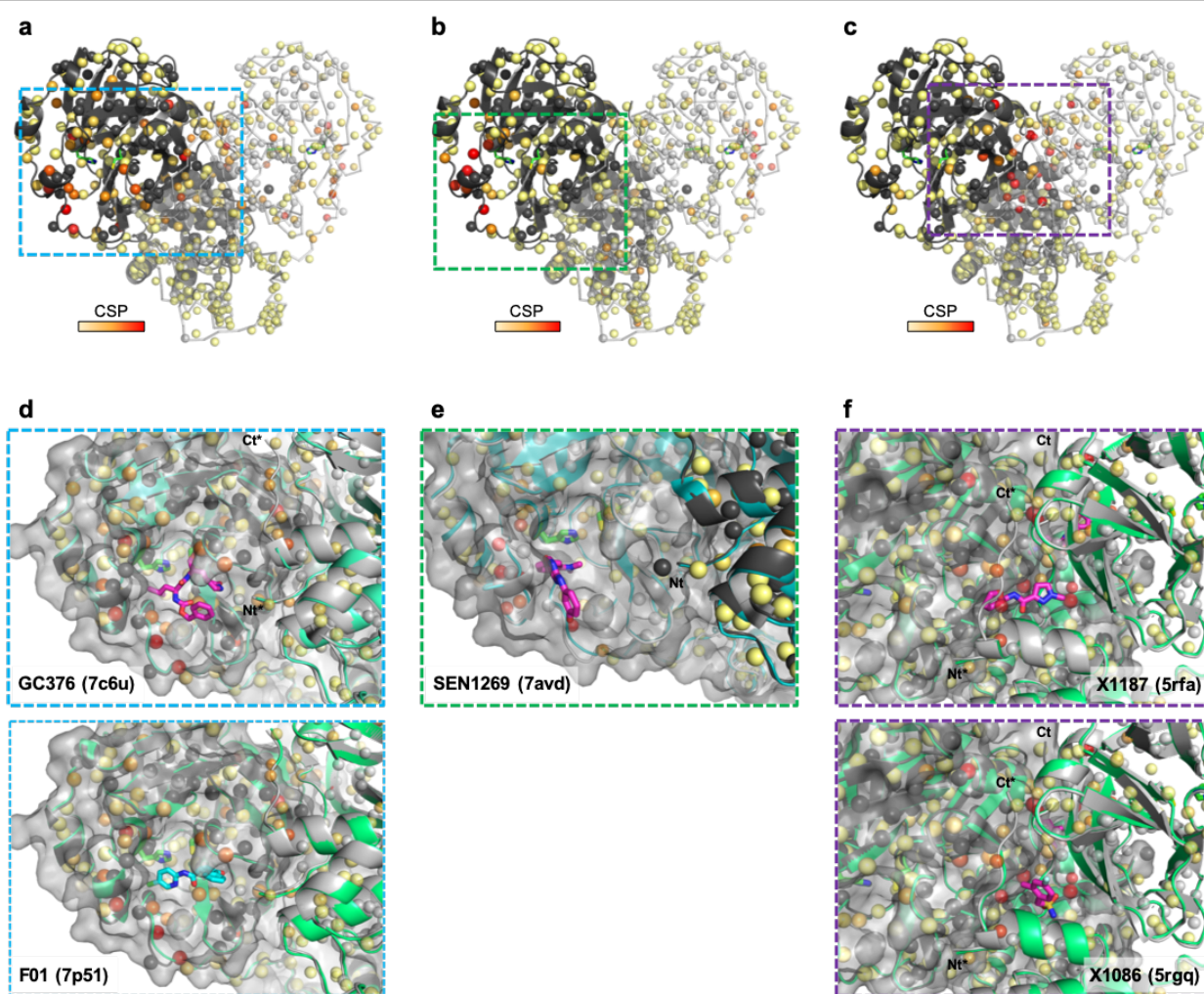


Figure S9. The 38 hits identified in the NMR screening can be grouped into three classes according to the CSPs they induced on the 2D NMR spectrum of ^2H , ^{15}N -3CLp upon binding. The representation is similar to that in Figure 4, and the ^1H and ^{15}N -combined CSPs following fragments binding have been color coded (from light yellow to red). The molecular surface of one of the 3CLp protomer is shown in translucent grey. (a) Class I (**F01**) - The CSPs are distributed in all the active site cleft, including the S1-S4 substrate pockets, and extend toward the dimerization interface of the protease. (b) Class II (**F30**) - The CSPs induced correspond to a binding of the fragments in the S2 and S3 pockets, with the highest perturbations observed for residues located in a short α -helix (Ser46-Leu50), (c) Class III (**F15**) - Upon binding these fragments induce CSPs at the dimerization interface of 3CLp. In panels (d), (e), and (f), the 3CLp structures with the NMR CSPs shown in (a), (b), and (c) have been overlaid with X-ray structures of 3CLp (in green or cyan) bound to molecules or fragments (in pink) that could induced similar CSPs. (d) Class I CSPs could match with the binding of GC376 (PDB: 7c6u) (top) and are in agreement with the crystal structure of 3CLp bound to fragment **F01** (in cyan) (PDB: 7np51) (bottom). (e) Class II CSPs could match with the binding of SEN1269 (PDB: 7avd). (e) Class III CSPs could match with the binding of two fragments, X1187 and X1086 (PDB: 5rfa and 5rgq). Figures were prepared using PyMOL (PyMOL Molecular Graphics System, version 1.8, Schrodinger, LLC).

SUPPORTING INFORMATION

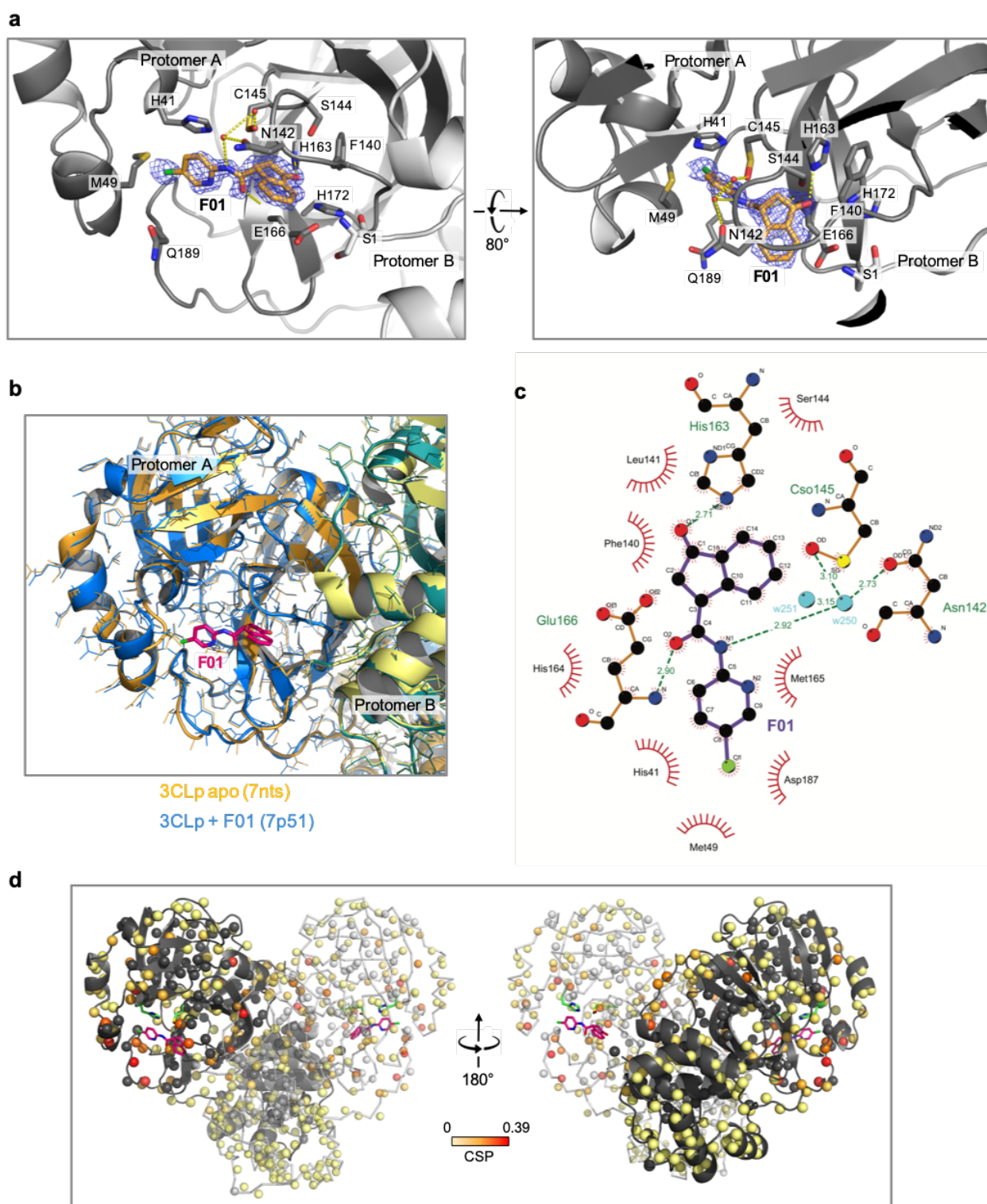


Figure S10. Crystal structure of 3CLp bound to fragment F01. (a) Structure of fragment F01 bound to SARS-CoV-2 3CLp (PDB: 7p51). The $2F_o - F_c$ electron-density map, contoured at 1σ , is shown as blue mesh around F01. Hydrogen bonds and water molecules are shown as yellow dashes and red spheres, respectively. (b) Structure of 3CLp bound to F01 (PDB: 7p51; in blue for the protein and in pink for F01) compared to that of 3CLp apo (PDB: 7nts; in orange). (c) Interactions of the fragment F01 (violet bonds) with 3CLp generated by the LigPlot[®] program^[31]. (d) Crystal structure of the SARS-CoV-2 3CLp bound to fragment F01 (PDB: 7p51) on which the NMR CSPs measured upon F01 binding (see Figure 4) have been color coded (from light yellow to red) and are displayed on the balls that represent the proton amides of the protein. Protomers A and B shown in black and light grey, respectively. Unassigned residues were kept in the original color of the protomer. The side-chains of the two catalytic residues, His41 and Cys145, are shown in green.

SUPPORTING INFORMATION

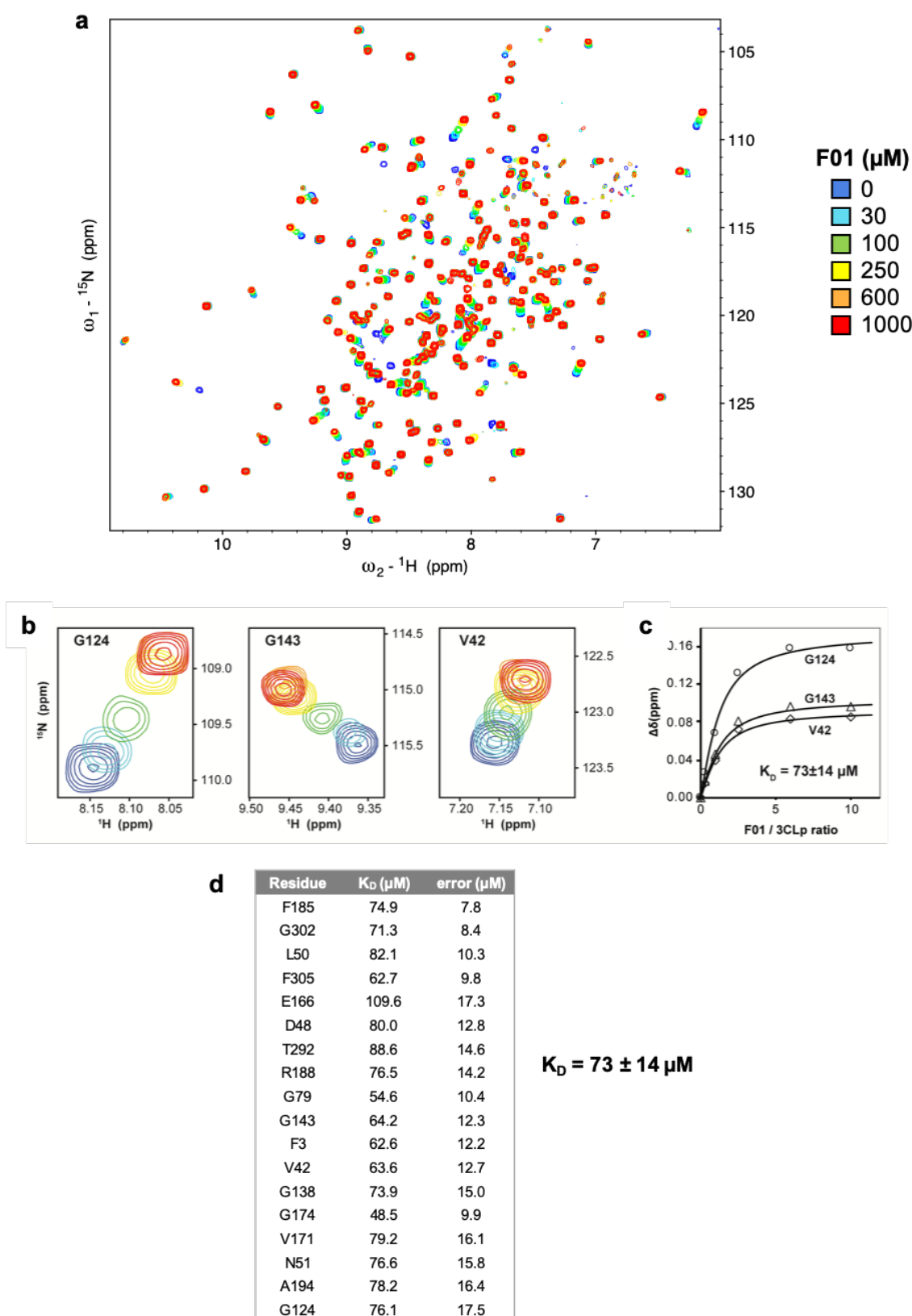


Figure S11. Interaction between F01 and 3CLp. The affinity of this interaction has been measured by NMR spectroscopy. (a) Overlays of 2D ^1H , ^{15}N -TROSY-HSQC spectra of 3CLp (100 μM) acquired in the absence (in blue) and the presence of increasing concentrations of F01 (30, 100, 250, 600 and 1000 μM , from cyan to red). (b) The spectra are centered on Val42 (next to catalytic residue His41), G143 (in the oxyanion loop around the S1 pocket) and G124 (located in the dimerization interface). (c) Titration curves corresponding to spectra in (b) and where the ^1H , ^{15}N -combined CSPs ($\Delta\delta$, ppm) were plotted as a function of the F01/3CLp ratios. (d) Table with the dissociation constants (K_D , μM) that have been measured for 18 3CLp residues. The K_D value (μM) on the right corresponds to the mean ($\pm\text{SD}$) calculated over the 18 3CLp resonances.

SUPPORTING INFORMATION

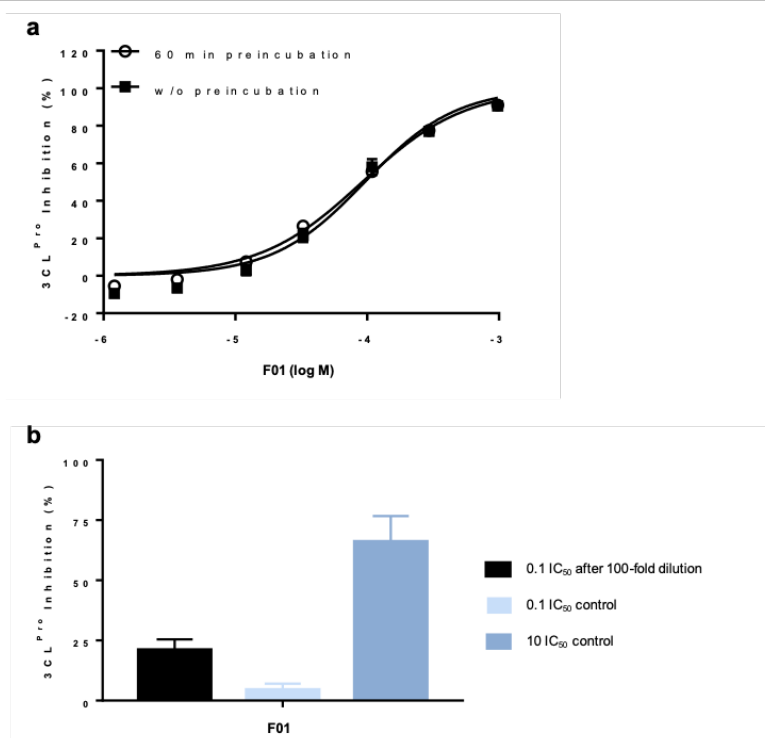


Figure S12. F01 is a reversible inhibitor of 3CLp. (a) The inhibition of the enzymatic activity of 3CLp by F01 was assessed without and with a 60 minutes preincubation of the compound with 3CLp before the addition of the fluorogenic peptide substrate (quenched FRET assay). (b) Jump dilution assay. F01 was preincubated at 10 times its IC₅₀ with 3CLpro. Then, the incubate was quickly diluted by 100-fold with substrate solution before measuring the fluorescence kinetics. Inhibition is compared to control standard incubations at 10x IC₅₀ and 0.1x IC₅₀ final concentrations of compounds. Final enzyme and substrate concentrations are 15 nM and 10 μM, respectively. Inhibitions (%) are shown as mean ± SD of triplicate from a representative experiment.

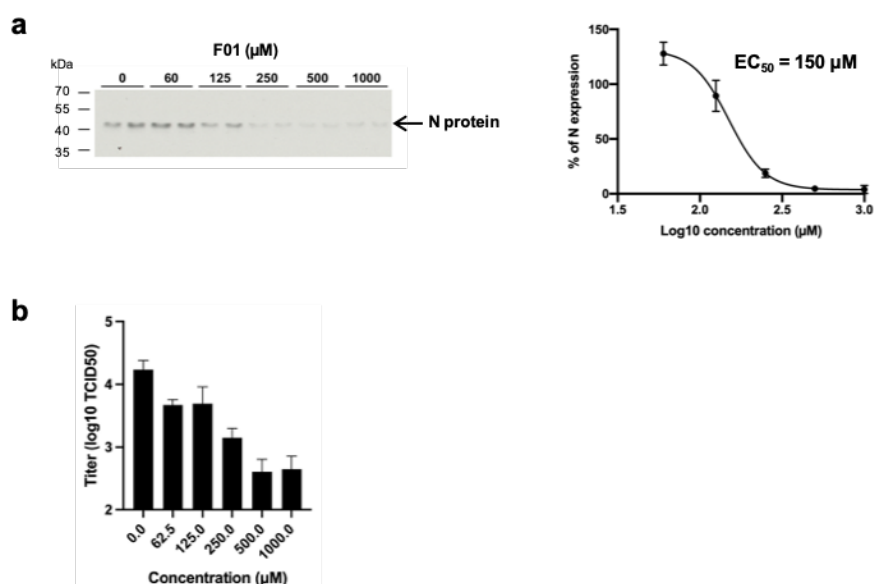
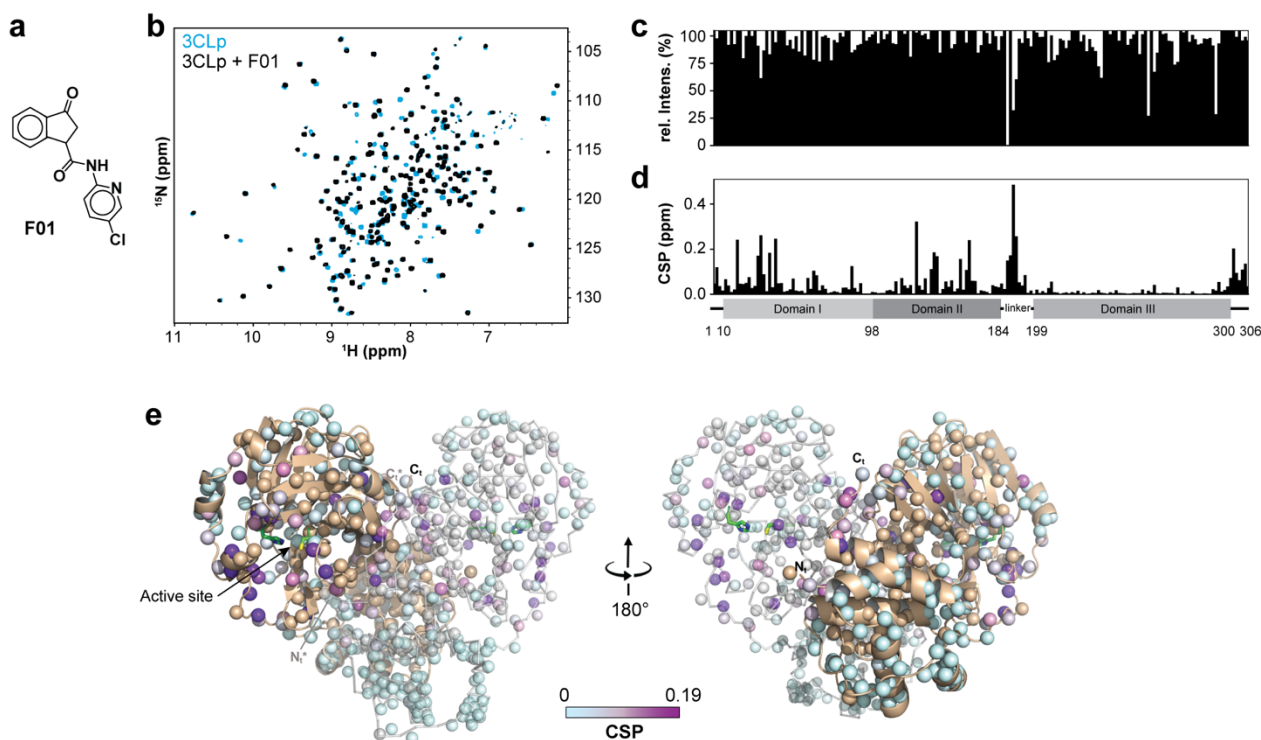


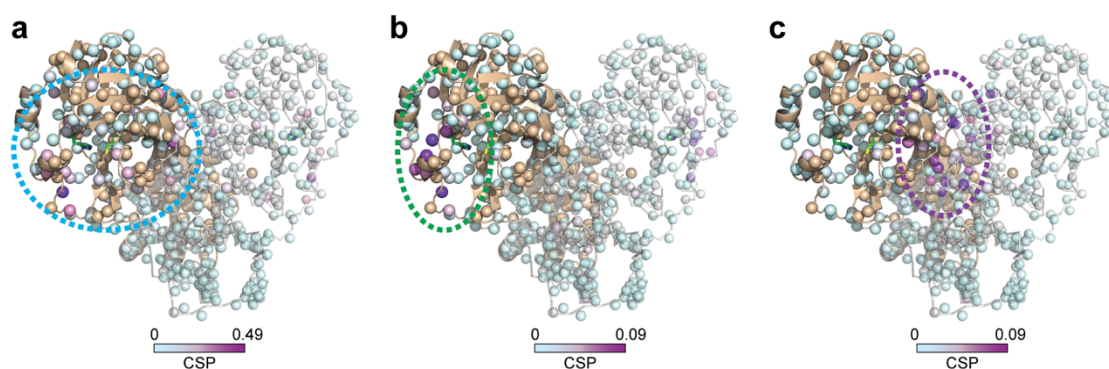
Figure S13. F01 is active against SARS-CoV-2 in Vero 81 cells. (a) The antiviral activity of F01 against SARS-CoV-2 has been tested on Vero-81 infected cells. After infection, in the presence of increasing F01 concentrations, the cells were lysed (t = 16 h) and the viral N-protein content was quantified (western blot) and was used to determine the half-maximal effective concentration (EC₅₀). (b) Viral titers were also measured in the cell supernatants by the TCID₅₀/ml method. A dose-dependent inhibition in the production of progeny virions was observed with a maximal decrease of 1.63 log₁₀ at the highest concentrations.

SUPPORTING INFORMATION

Adapted version of Figure 4 and Figure 5 for color-blind persons



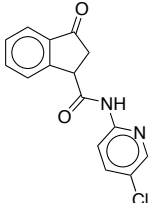
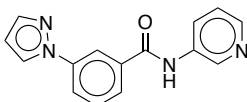
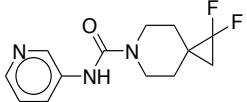
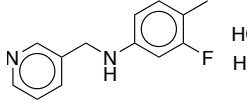
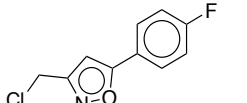
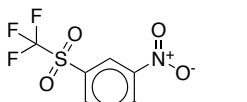
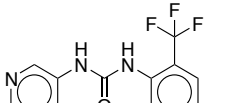
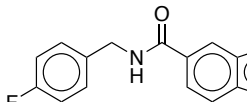
Adapted Figure 4. Protein-based NMR secondary screening. (a) Fragment **F01** structure. (b) Overlay of two 2D ^1H , ^{15}N -TROSY-HSQC spectra acquired on 3CLp in the absence (in light blue) and in the presence (in black) of fragment **F01**. The broadening of the resonances and ^1H and ^{15}N -combined CSPs induced upon fragment binding are shown along the 3CLp sequence in (c) and (d), respectively. (e) Structure of the 3CLp dimer (PDB: 7k3t), with protomers A and B shown in light brown and light grey, respectively. Each small ball represents a proton amide and thus should correspond to a resonance in the ^1H , ^{15}N 2D spectrum. The CSPs, shown in (d), have been color coded (from pale cyan to purple) and are displayed on these balls. Unassigned residues were kept in the original color of the protomer. Catalytic His41 and Cys145, are shown in green. See Scheme S2 and Table S2 for other hits.



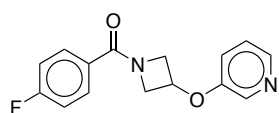
Adapted Figure 5. The 38 hits identified in the NMR screening can be grouped into three classes according to the CSPs they induced on the 2D NMR spectrum of ^2H , ^{15}N -3CLp upon binding. The representation is similar to that in Adapted Figure 4e. (a) Class I (**F01**) - The CSPs are distributed in all the active site cleft, including the S1-S4 substrate pockets, and extend toward the dimerization interface of the protease. (b) Class II (**F30**) - The CSPs induced correspond to a binding of the fragments in the S2 and S3 pockets, with the highest perturbations observed for residues located in a short α -helix (Ser46-Leu50). (c) Class III (**F15**) - Upon binding these fragments induce CSPs at the dimerization interface of 3CLp. See Figs. S8-S9.

SUPPORTING INFORMATION

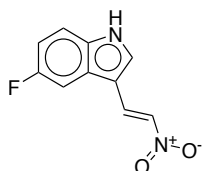
Table S1. Fragment hits identified.

| Fragment Hit | SMILES | MW (Da) | Supplier | Ref |
|--------------|---|---------|----------------|------------|
| F01 | <chem>Clc1ccc(NC(=O)C2CC(=O)c3ccccc23)nc1</chem> | 286.71 | Life Chemicals | F5735-0021 |
| |  | | | |
| F18 | <chem>O=C(Nc1cccnc1)c2cccc(c2)n3cccn3</chem> | 264.28 | Life Chemicals | F6665-6159 |
| |  | | | |
| F35 | <chem>FC1(F)CC12CCN(CC2)C(=O)Nc3ccnc3</chem> | 267.27 | Life Chemicals | F6573-6532 |
| |  | | | |
| F04 | <chem>Cl.C1Cc1ccc(NCc2ccccc2)cc1F</chem> | 216.26 | Life Chemicals | F6616-6309 |
| |  | | | |
| F23 | <chem>Fc1ccc(cc1)c2onc(CCl)c2</chem> | 211.62 | Life Chemicals | F2161-0038 |
| |  | | | |
| F12 | <chem>[O-][N+](=O)c1cc(ccc1Cl)S(=O)(=O)C(F)(F)F</chem> | 289.62 | Life Chemicals | F0293-0022 |
| |  | | | |
| F13 | <chem>Cc1ccncc1NC(=O)Nc2ccccc2C(F)(F)F</chem> | 295.26 | Life Chemicals | F6666-4121 |
| |  | | | |
| F10 | <chem>Fc1ccc(CNC(=O)c2ccc3nc[nH]c3c2)cc1</chem> | 269.27 | Life Chemicals | F3055-0113 |
| |  | | | |
| F24 | <chem>Fc1ccc(cc1)C(=O)N2CC(C2)Oc3ccnc3</chem> | 272.27 | Life Chemicals | F6360-9264 |

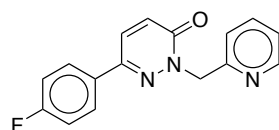
SUPPORTING INFORMATION



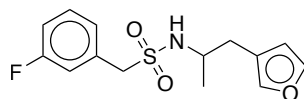
F03 [O-][N+](=O)\C=Cc1c[nH]c2ccc(F)cc12 206.17 Life Chemicals F1913-1352



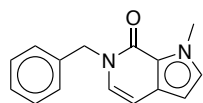
F11 Fc1ccc(cc1)C2=NN(Cc3ccccc3)C(=O)C=C2 281.28 Life Chemicals F2774-1665



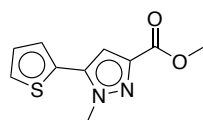
F05 CC(Cc1ccccc1)NS(=O)(=O)Cc2ccc(F)c2 297.35 Life Chemicals F6356-0753



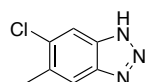
F17 Cn1ccc2C=CN(Cc3ccccc3)C(=O)c12 238.28 Life Chemicals F6660-2788



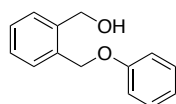
F28 COC(=O)c1cc(c2cccs2)n(C)n1 222.26 Life Chemicals F1911-2943



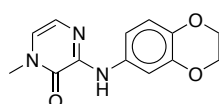
F21 CC1=C(Cl)C=C(NN=N2)C2=C1 167.60 MayBridge RDP00249R3



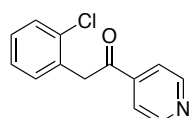
F06 C1=CC=C(C=C1)OCC2=CC=CC=C2CO 214.26 MayBridge CC58509R3



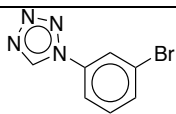
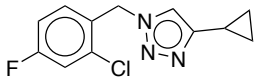
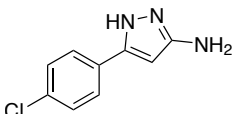
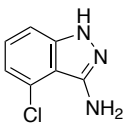
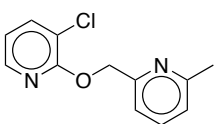
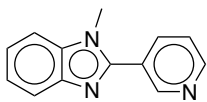
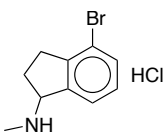
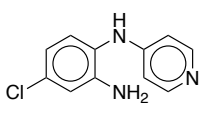
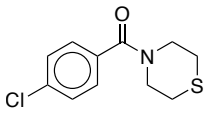
F02 CN1C=CN=C(Nc2ccc3OCCOc3c2)C1=O 259.26 Life Chemicals F6616-6302



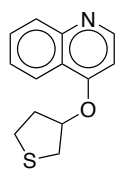
F08 C1=CC=C(C(=C1)CC(=O)C2=CC=NC=C2)Cl 231.68 MayBridge MO00227R3



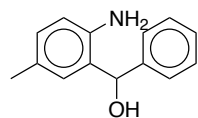
SUPPORTING INFORMATION

| | | | | |
|---|---|---------|----------------|------------|
| F29 | <chem>BrC1CCCC(C1)N2CN=CN2</chem> | 225.05 | Life Chemicals | F2158-0656 |
|  | | | | |
| F33 | <chem>Fc1ccc(Cn2cc(nn2)C3CC3)c(Cl)c1</chem> | 251.69 | Life Chemicals | F6511-9283 |
|  | | | | |
| F07 | <chem>C1=CC(=CC=C1C2=CC(=NN2)N)Cl</chem> | 193.63 | MayBridge | KM01548R3 |
|  | | | | |
| F19 | <chem>C1=CC2=C(C(=C1)Cl)C(=NN2)N</chem> | 167.60 | MayBridge | CD05676R3 |
|  | | | | |
| F36 | <chem>Cc1cccc(COc2ncccc2Cl)n1</chem> | 234.68 | Life Chemicals | F6608-1709 |
|  | | | | |
| F26 | <chem>Cn1c(nc2ccccc12)c3ccncc3</chem> | 209.25 | Life Chemicals | F0919-0210 |
|  | | | | |
| F30 | <chem>Cl.CNC1CCc2c(Br)cccc12</chem> | 226.12 | Life Chemicals | F2167-9584 |
|  | | | | |
| F22 | <chem>Nc1cc(Cl)ccc1Nc1ccncc1</chem> | 167.60 | MayBridge | HTS11483R3 |
|  | | | | |
| F09 | <chem>c1cc(ccc1C(=O)N2CCSCC2)Cl</chem> | 241.737 | MayBridge | CD07932R3 |
|  | | | | |
| F36 | <chem>C1CC(CS1)Oc2ccnc3ccccc23</chem> | 231.31 | Life Chemicals | F6608-0683 |

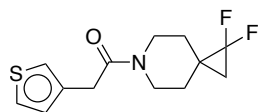
SUPPORTING INFORMATION



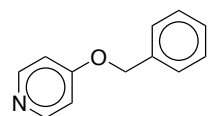
F14 Cc1ccc(N)c(c1)C(O)c2ccccc2 213.28 Life Chemicals F0013-1175



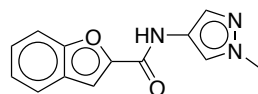
F34 FC1(F)CC12CCN(CC2)C(=O)Cc3ccsc3 271.33 Life Chemicals F6573-6442



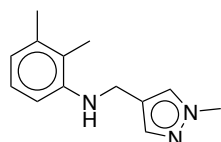
F32 C(Oc1ccncc1)c2ccccc2 185.22 Life Chemicals F2199-0592



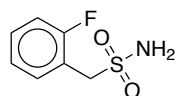
F16 Cn1cc(NC(=O)c2oc3ccccc3c2)cn1 241.25 Life Chemicals F6616-7253



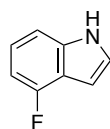
F15 Cc1cccc(NCc2cnn(C)c2)c1C 215.29 Life Chemicals F6616-6292



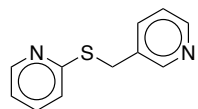
F31 NS(=O)(=O)Cc1ccccc1F 189.21 Life Chemicals F2189-0092



F20 C1=CC2=C(C=CN2)C(=C1)F 135.14 MayBridge AC39765R3

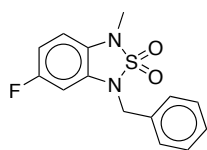


F38 C(Sc1ccccn1)c2cccnc2 202.28 Life Chemicals F6616-1307

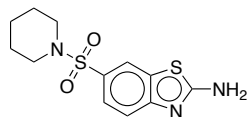


F25 CN1c2ccc(F)cc2N(Cc3ccccc3)S1(=O)=O 292.33 Life Chemicals F6510-4543

SUPPORTING INFORMATION



| | | | | |
|-----|--|--------|----------------|------------|
| F27 | <chem>Nc1nc2ccc(cc2s1)S(=O)(=O)N3CCCCC3</chem> | 297.40 | Life Chemicals | F1420-1568 |
|-----|--|--------|----------------|------------|



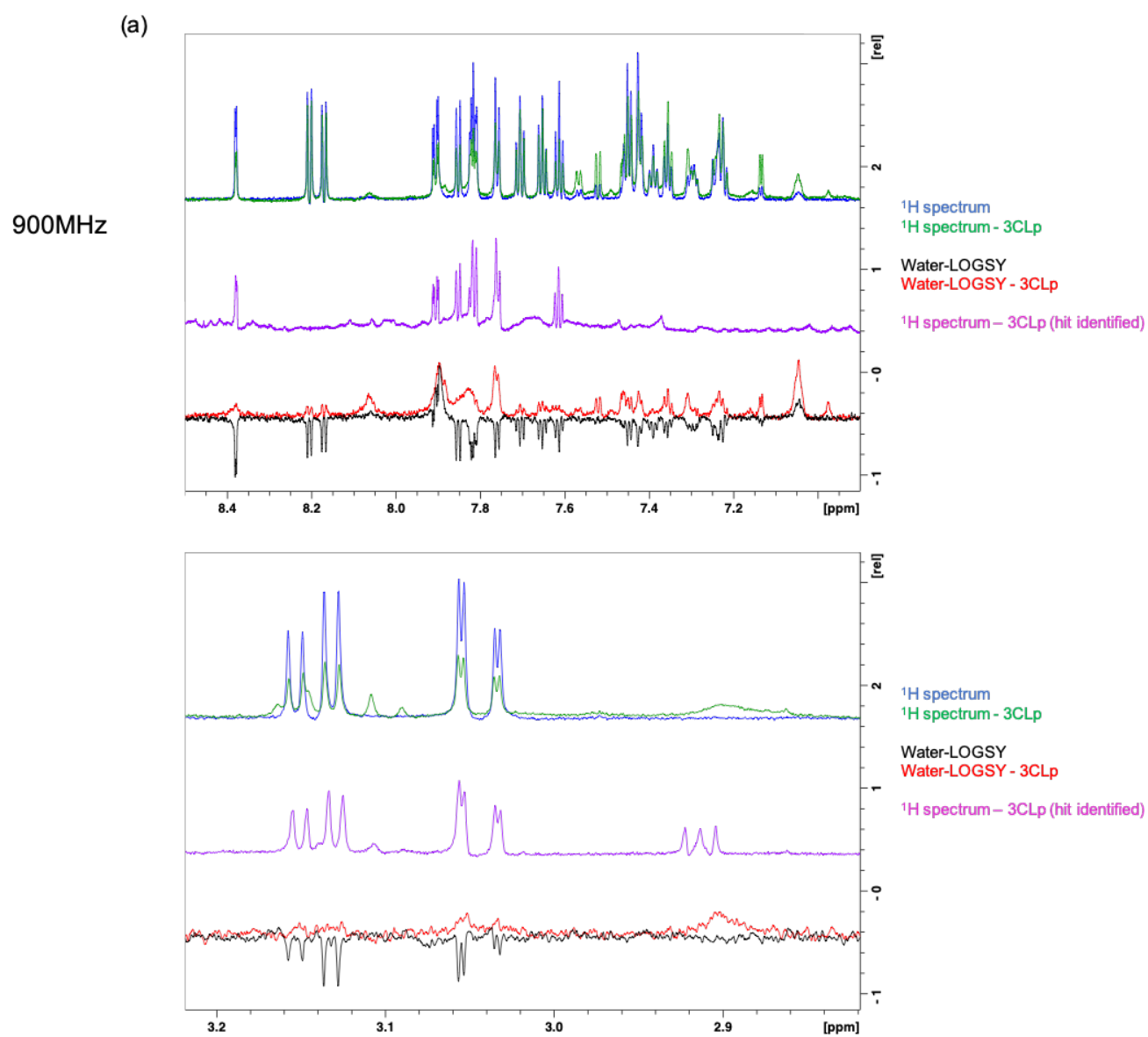
SUPPORTING INFORMATION

Scheme S1. Ligand-based NMR primary screening. Analyses of the cocktails (5 fragments, 377 μM each) containing the 38 fragment hits that have been identified in this study. (a) The 1D ^1H and ^1H Water-LOGSY spectra of the cocktail, that have been acquired in the absence and in the presence of 25 μM of unlabeled 3CLp, are shown in blue, black, green and red, respectively. The positive signals in the Water-LOGSY spectrum, acquired in the presence of 3CLp, identify the fragment(s) as direct binder(s) of the protease. The ^1H spectrum shown in pink corresponds to the identified fragment hit (2 mM) in the presence of the 3CLp (100 μM) [data from the secondary screening]. The ^1H spectra were acquired, using 5 mm tubes, at 293K on either a 600 MHz or a 900 MHz spectrometer equipped with a ^1H cryogenic probe. Low intensity signals from binders are marked with an asterisk. (b) Analysis of the same cocktail using 1D ^{19}F -NMR, in the absence (in blue) and in the presence (in green) of 3CLp. The inset highlights the signal corresponding to the fragment that binds to 3CLp (chemical shift perturbation and/or signal broadening). All ^{19}F spectra were acquired, using 5 mm tubes, at 293K on a 600 MHz spectrometer equipped with a ^1H and ^{19}F cryogenic probe.

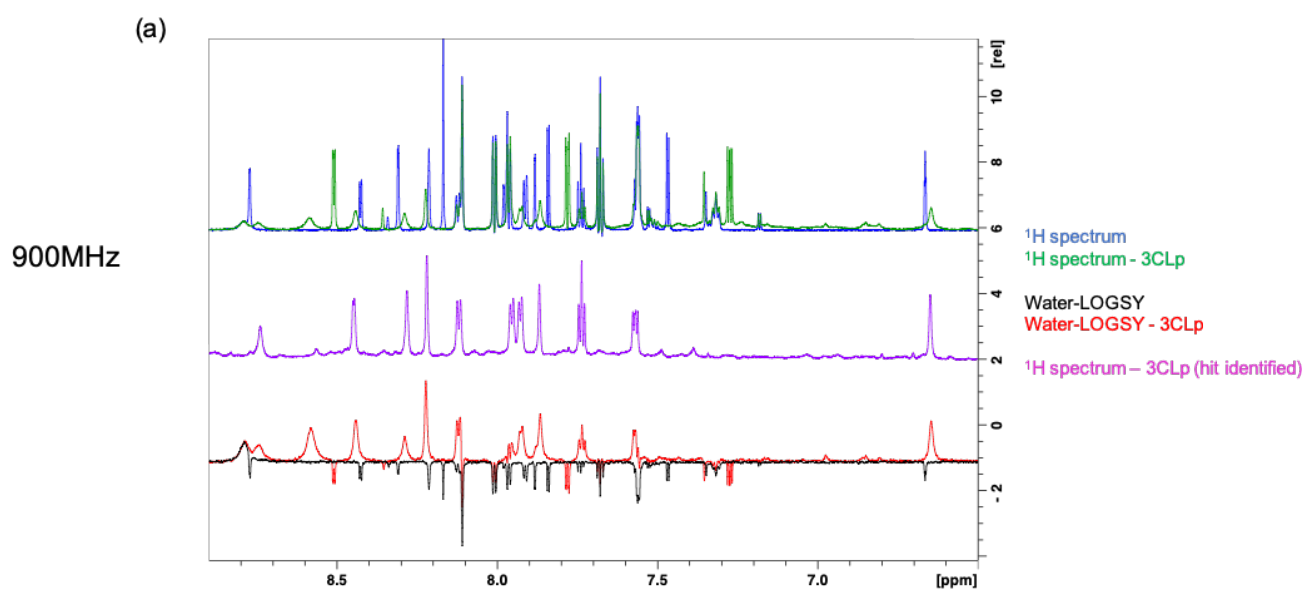
| | |
|-----|----|
| F01 | 25 |
| F18 | 26 |
| F35 | 27 |
| F04 | 28 |
| F23 | 29 |
| F12 | 30 |
| F13 | 31 |
| F10 | 32 |
| F24 | 33 |
| F03 | 34 |
| F11 | 35 |
| F05 | 36 |
| F17 | 37 |
| F28 | 38 |
| F21 | 39 |
| F06 | 40 |
| F02 | 41 |
| F08 | 42 |
| F29 | 42 |
| F33 | 43 |
| F07 | 44 |
| F19 | 44 |
| F36 | 45 |
| F26 | 46 |
| F15 | 47 |
| F31 | 48 |
| F20 | 49 |
| F38 | 50 |
| F25 | 51 |
| F30 | 52 |
| F22 | 53 |
| F09 | 54 |
| F36 | 55 |
| F14 | 56 |
| F34 | 57 |
| F32 | 58 |
| F16 | 59 |
| F27 | 60 |

SUPPORTING INFORMATION

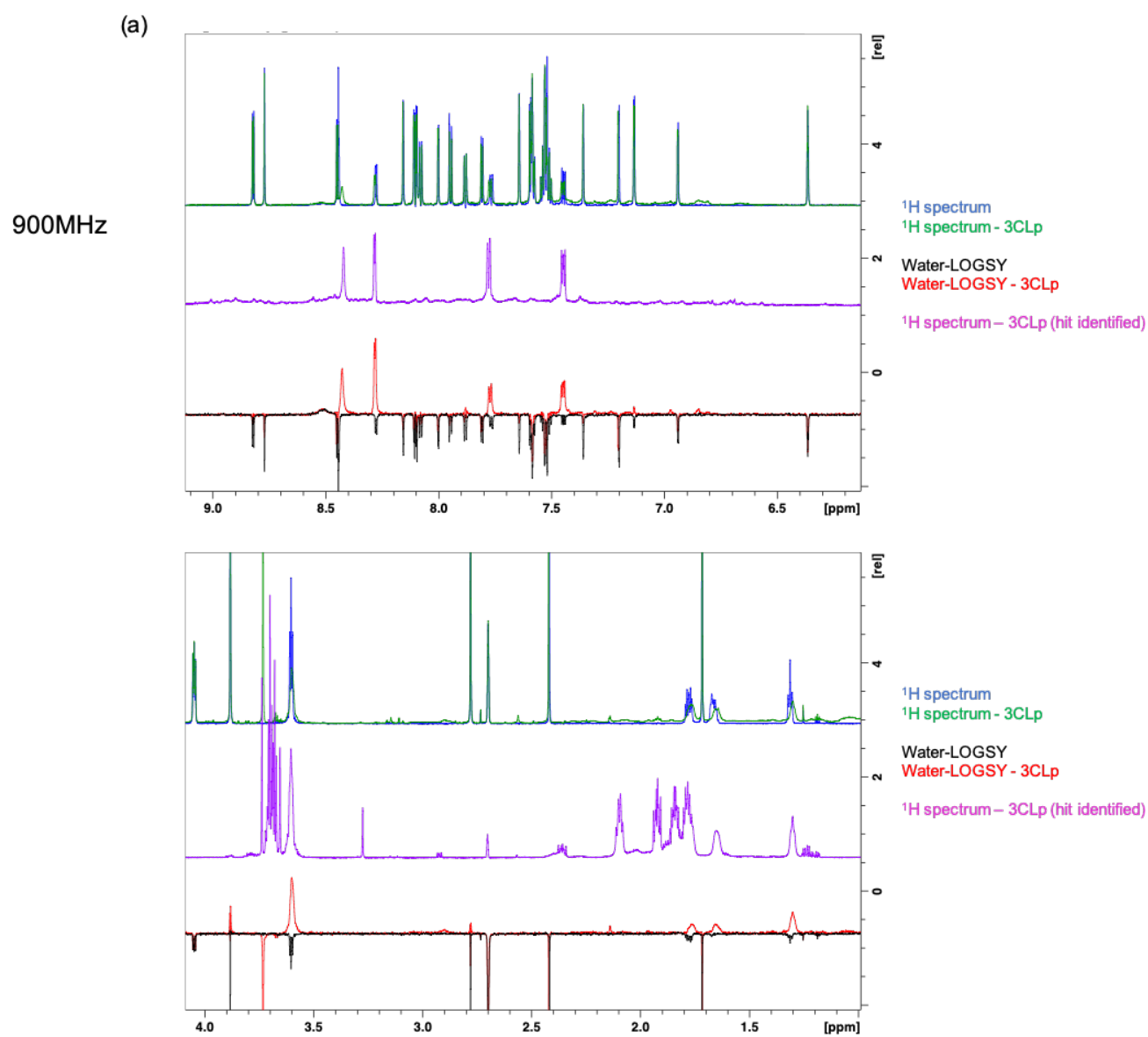
F01



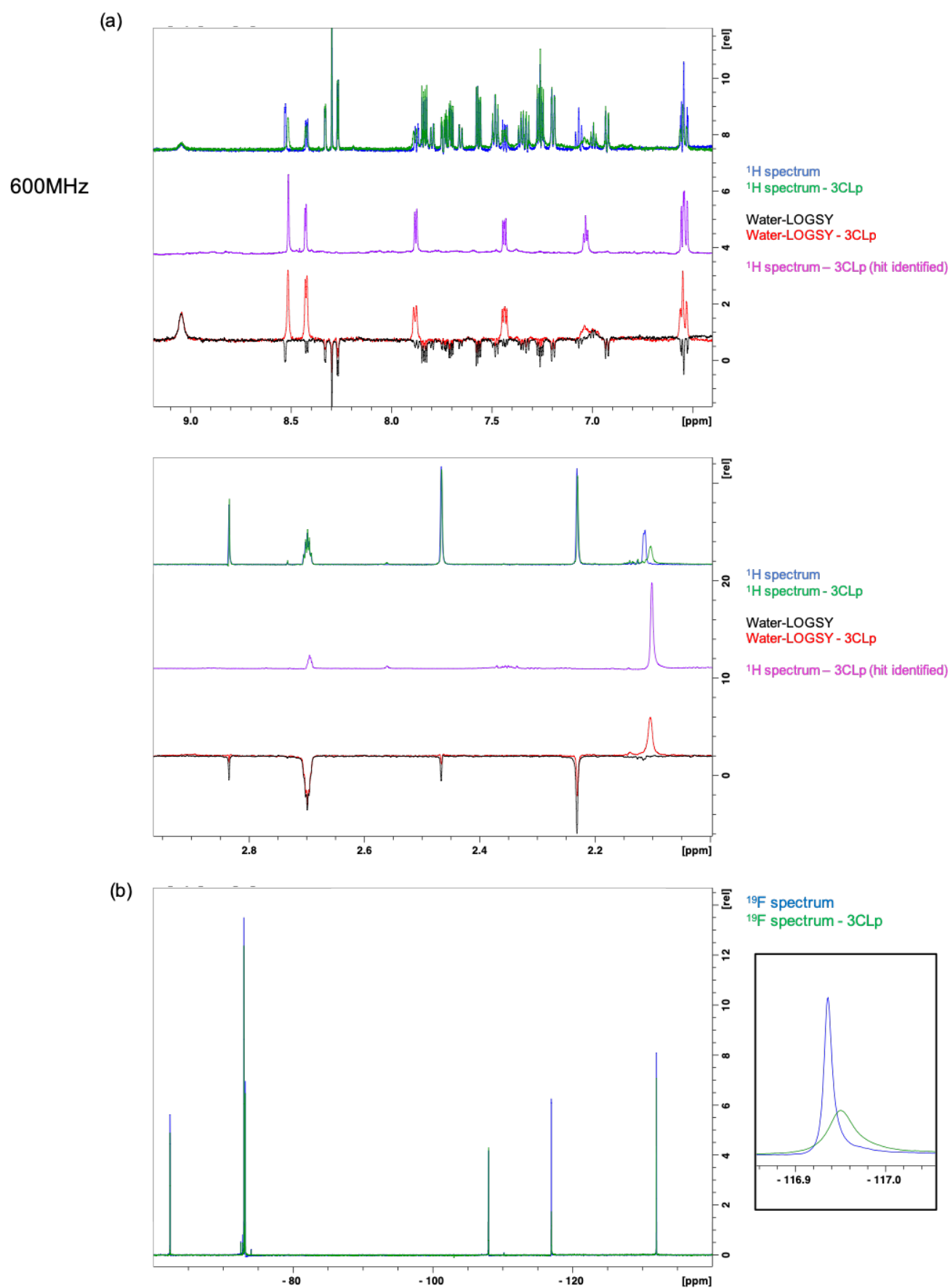
F18



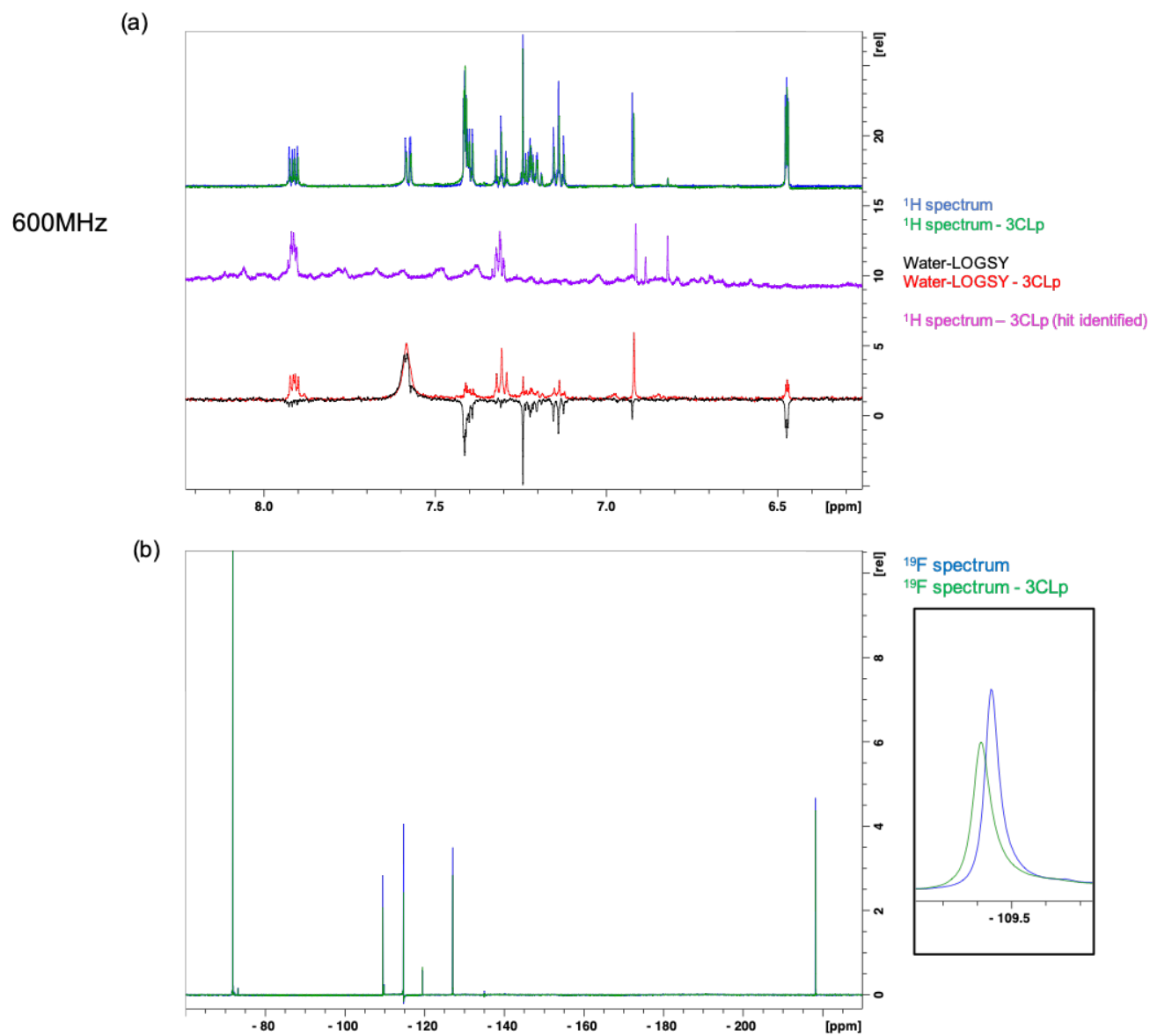
F35



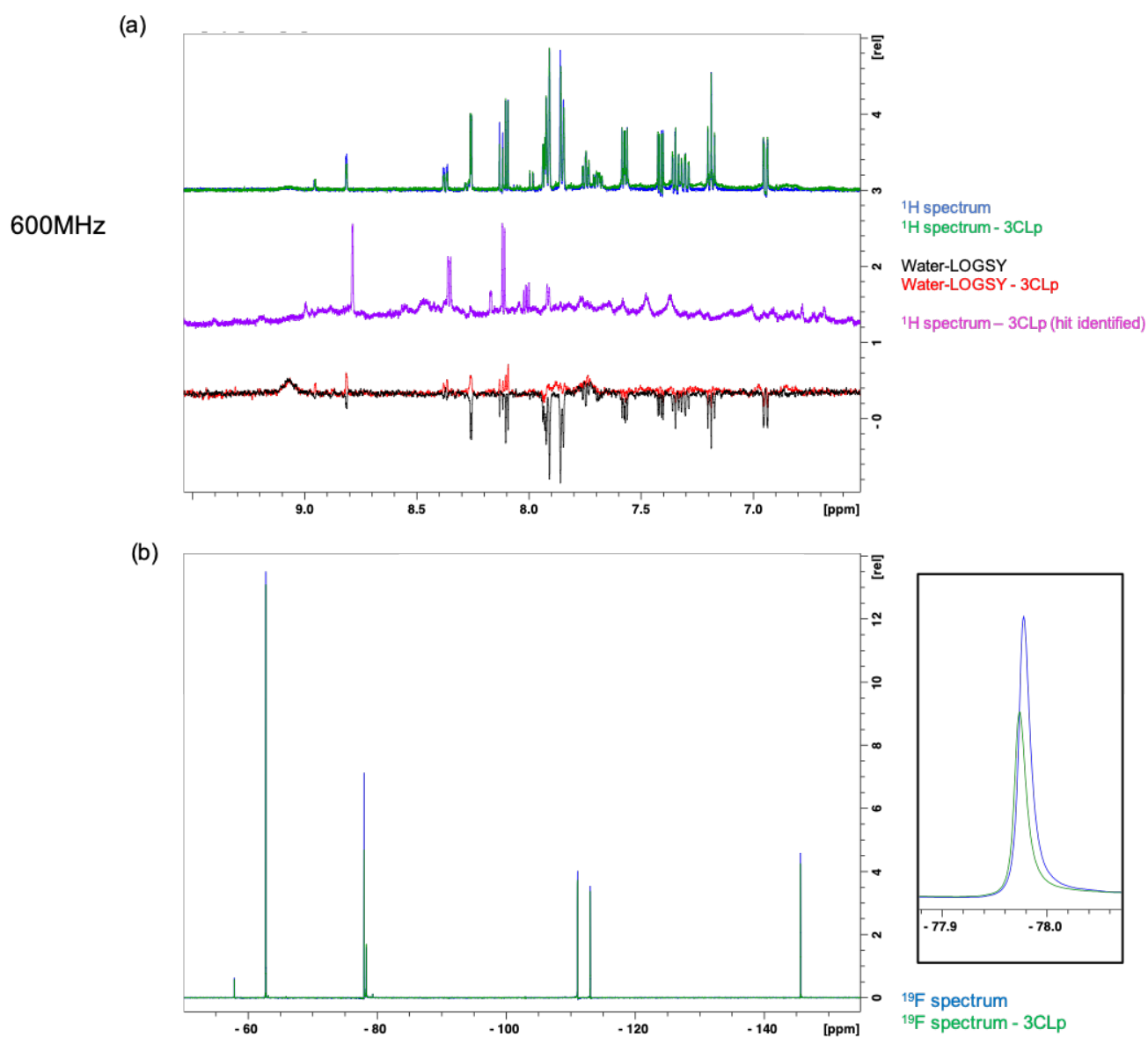
F04



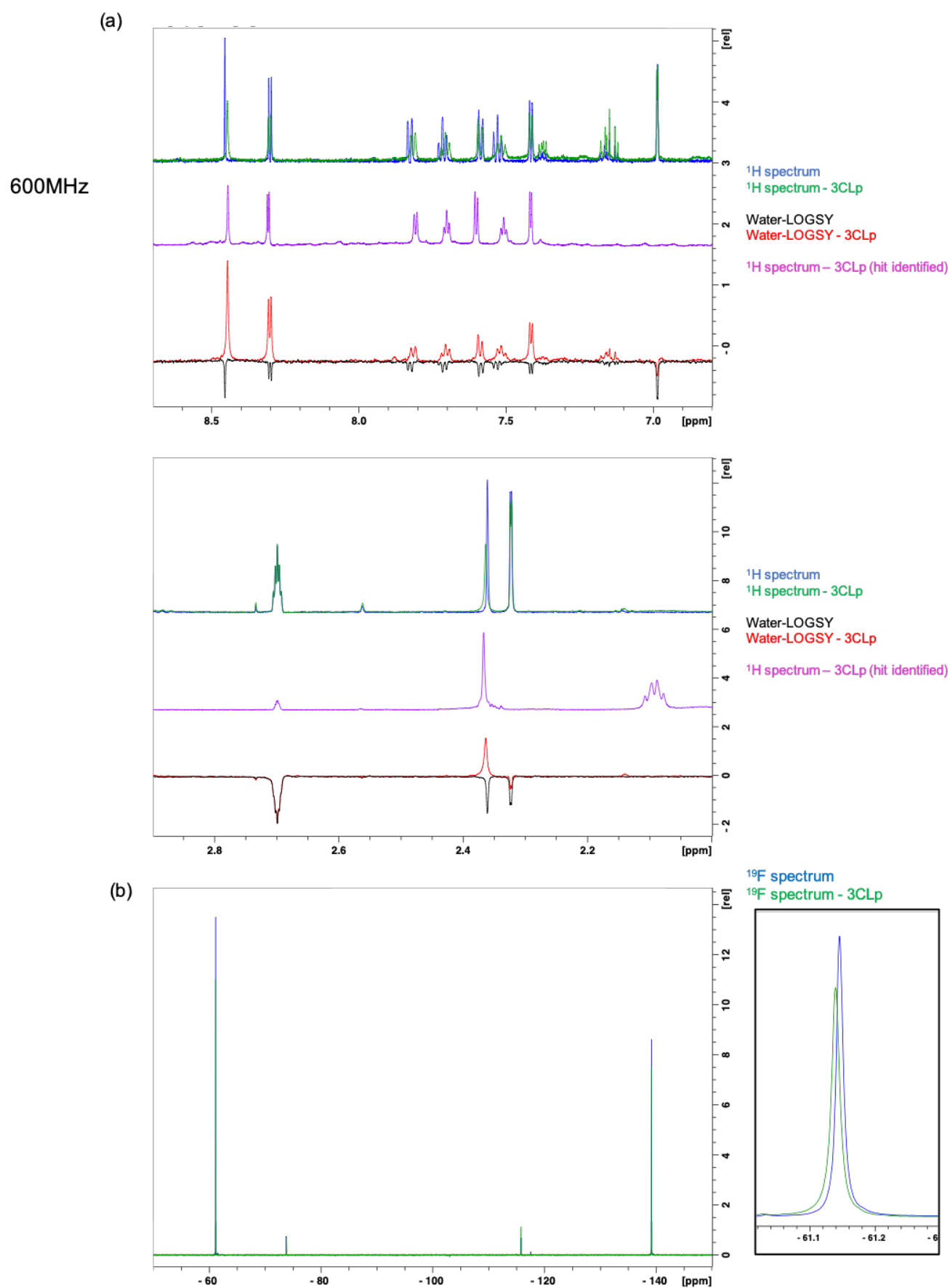
F23



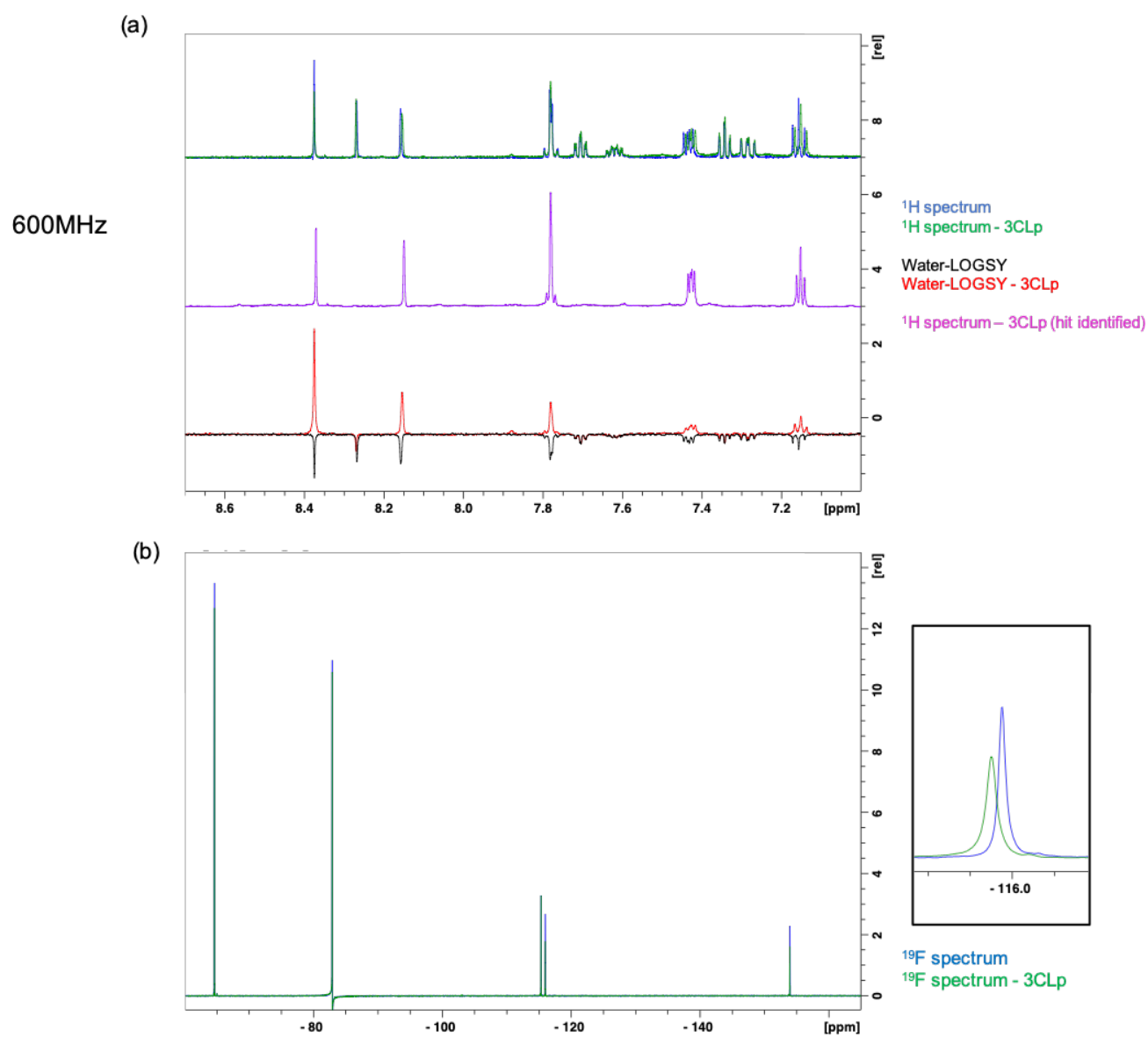
F12



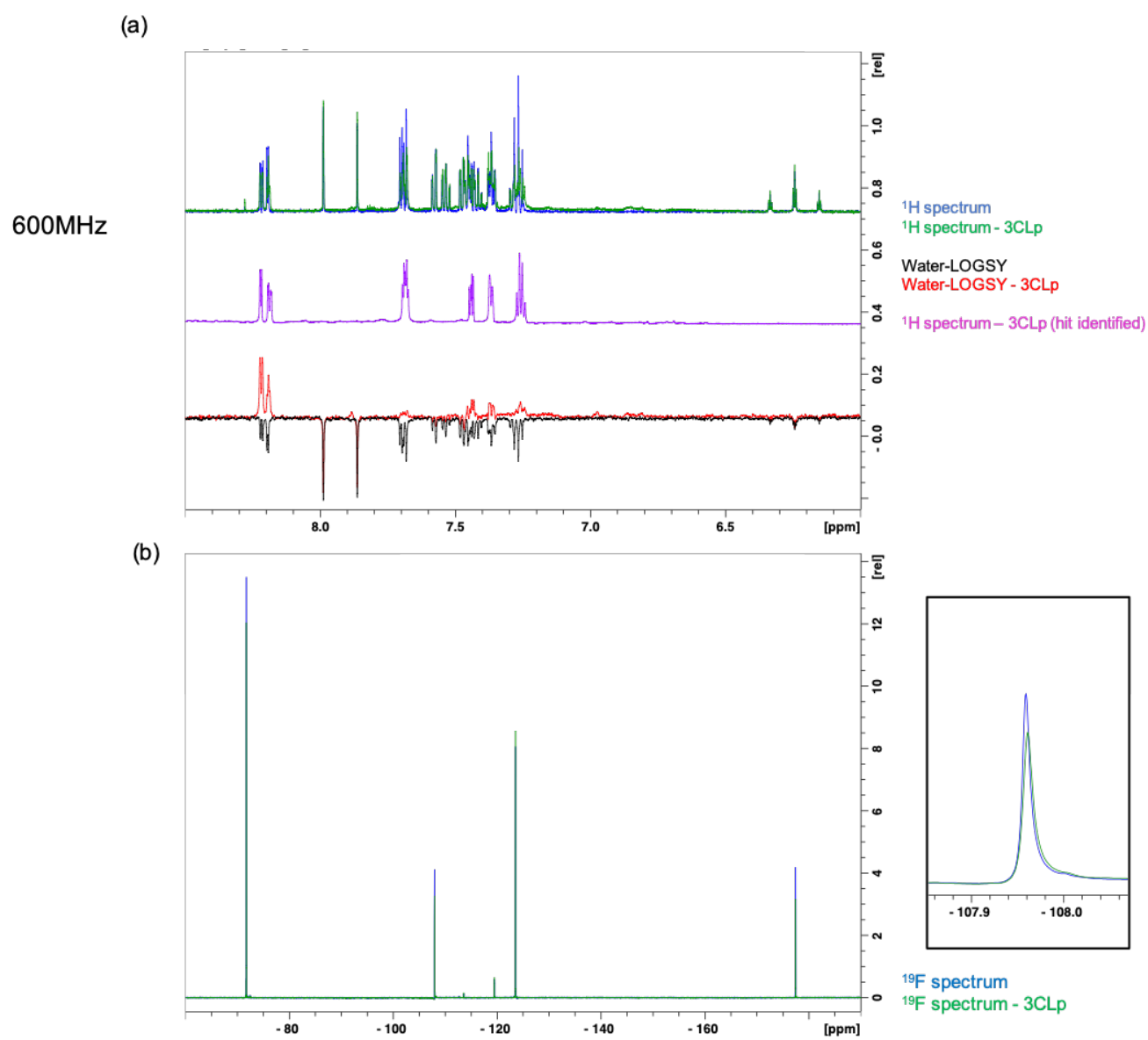
F13



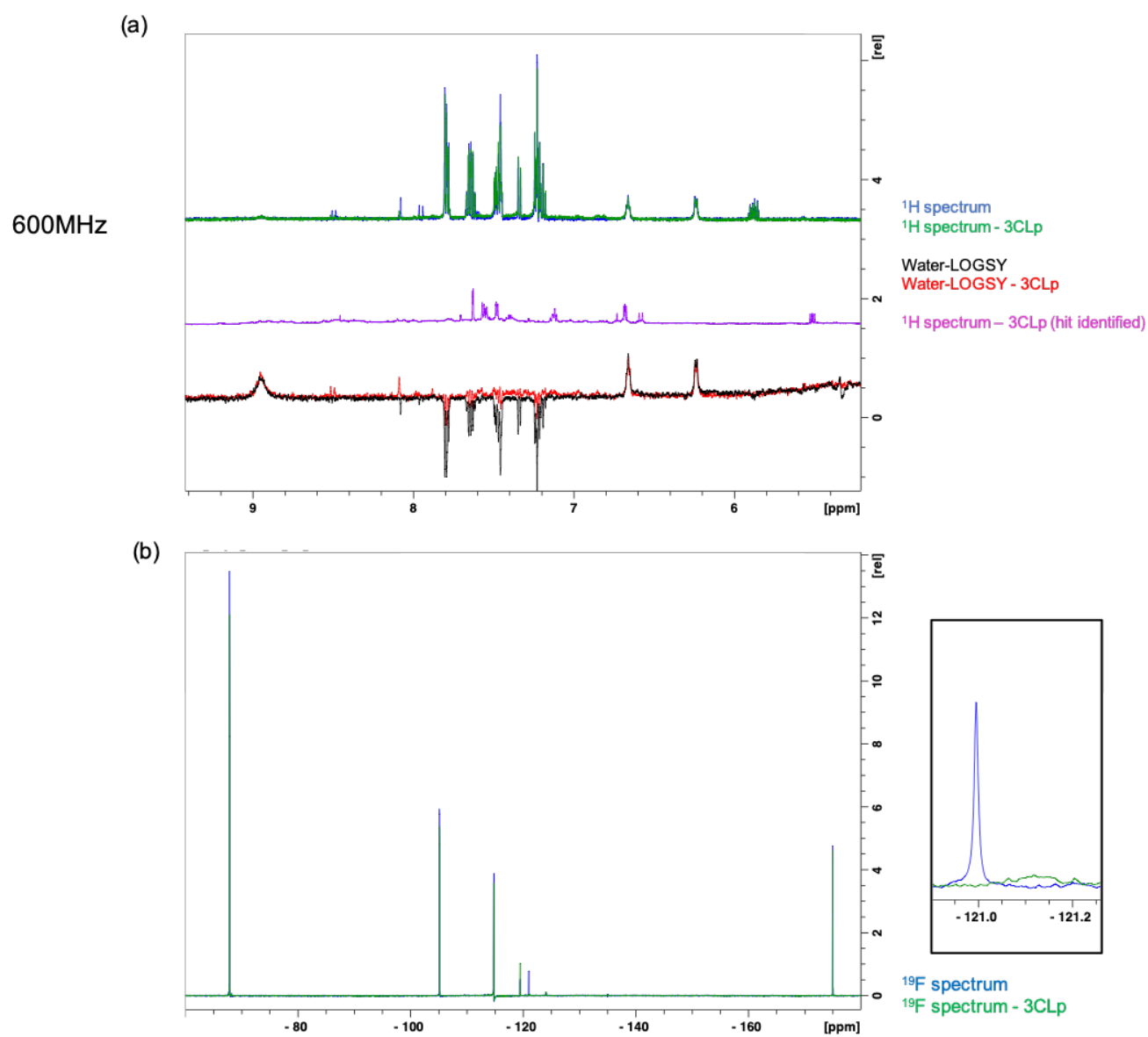
F10



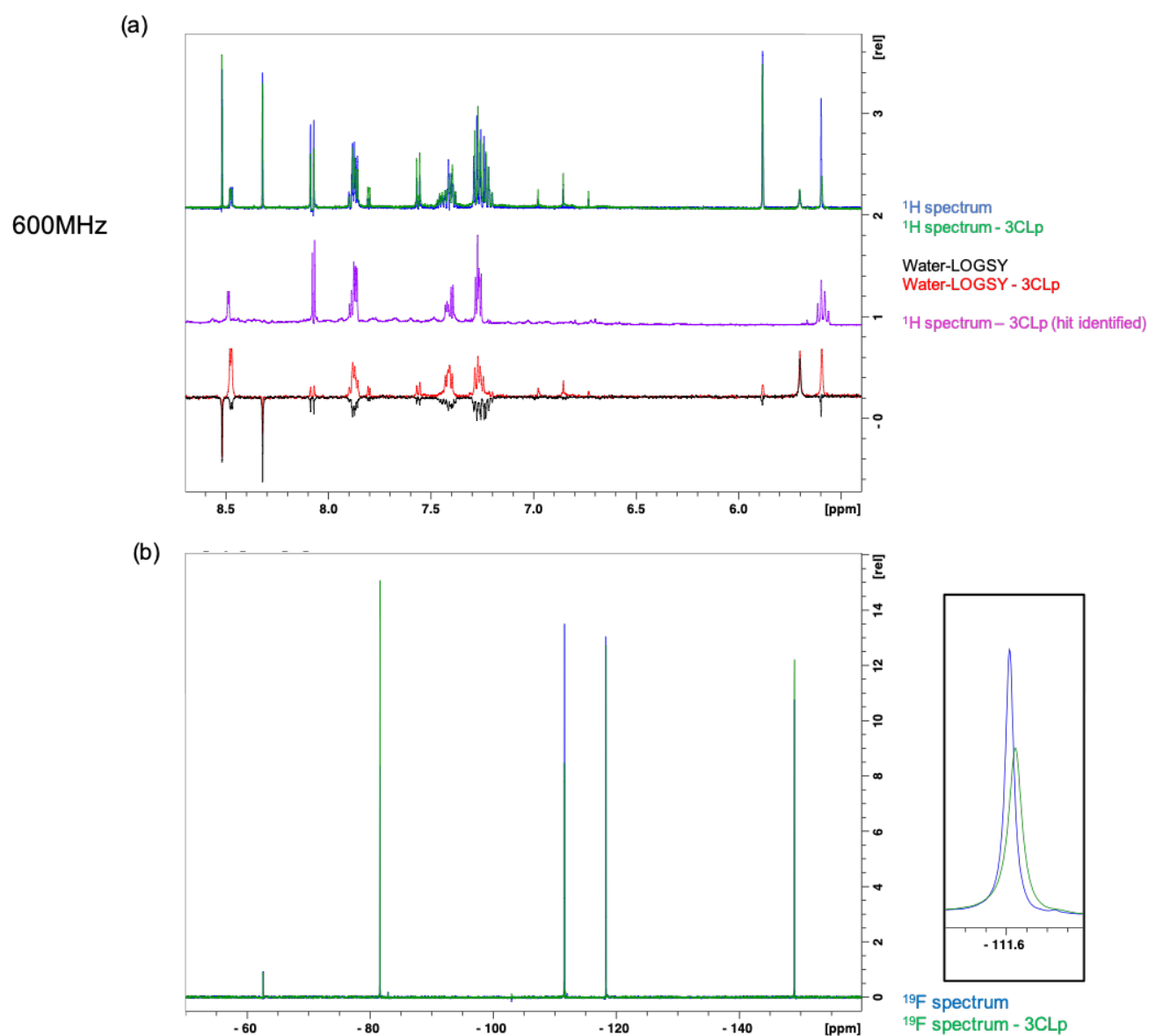
F24



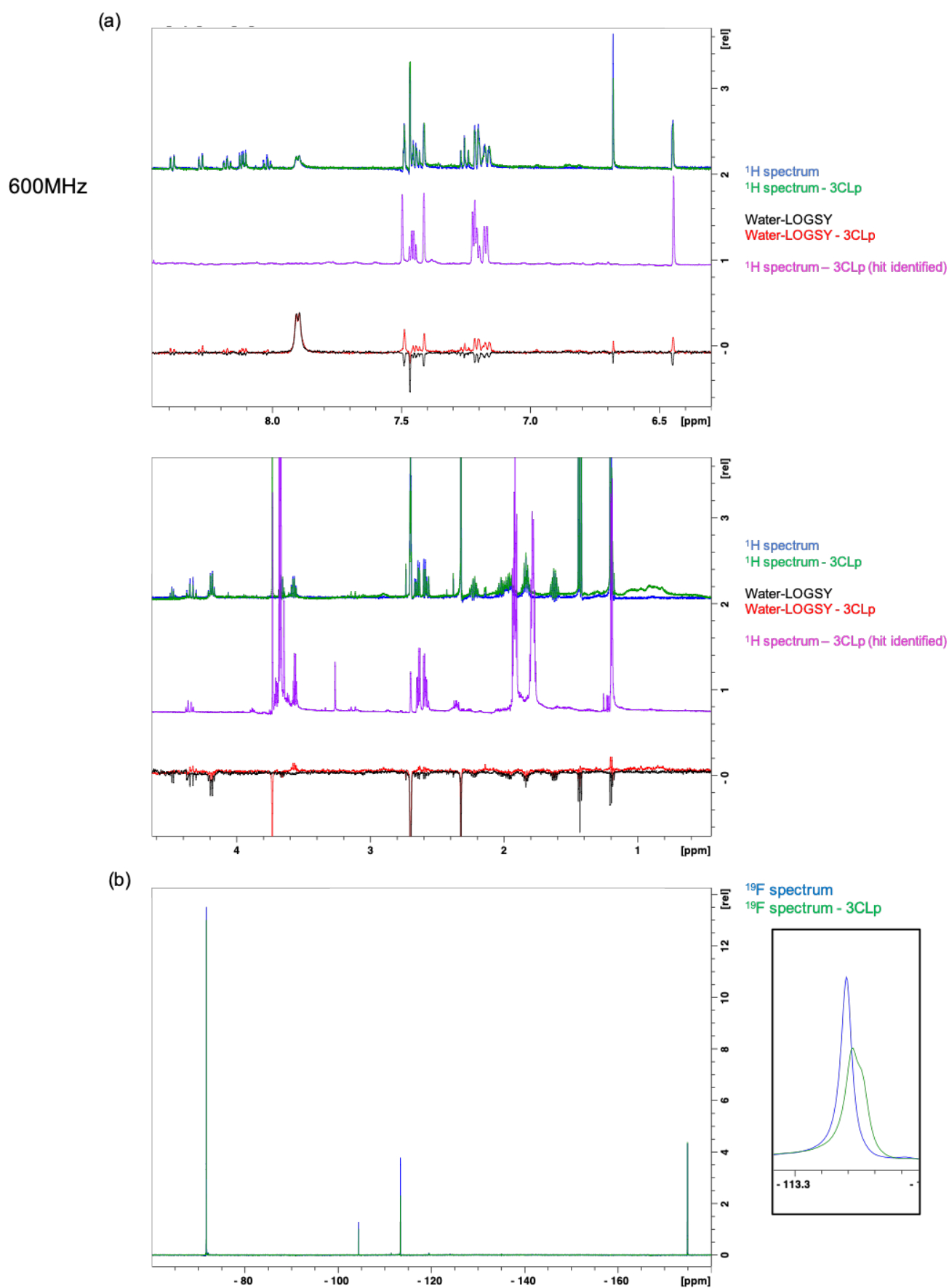
F03



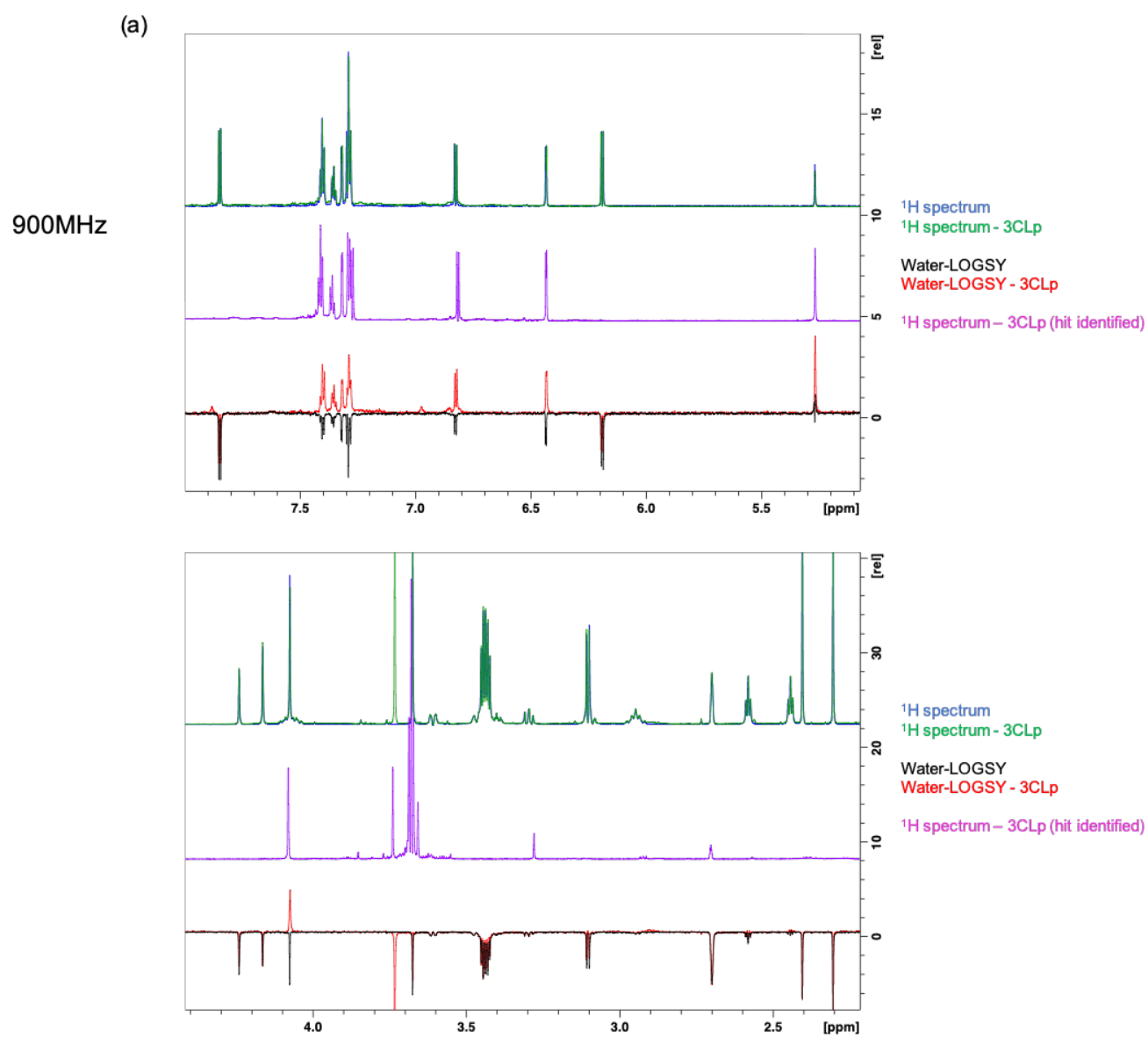
F11



F05

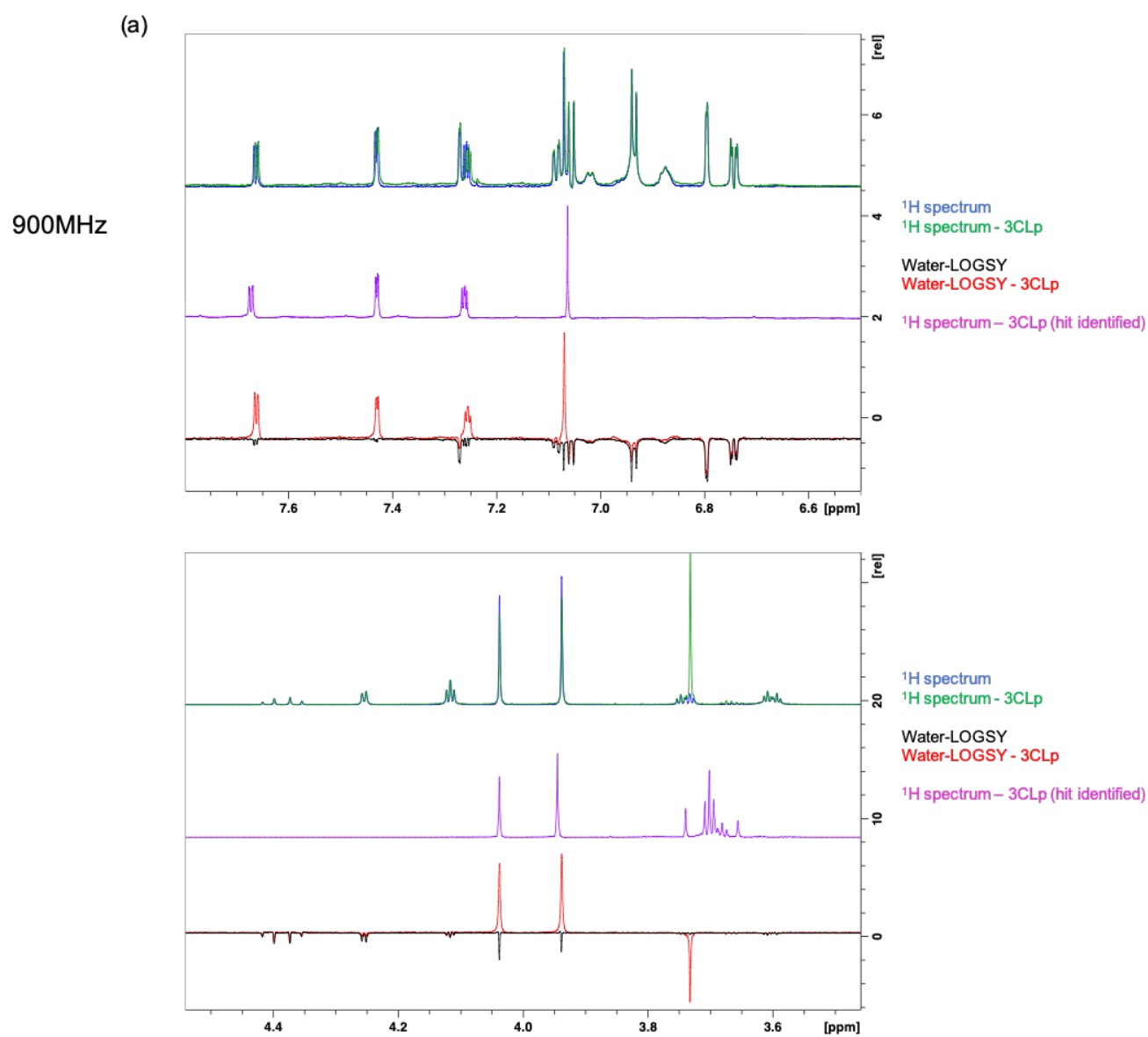


F17

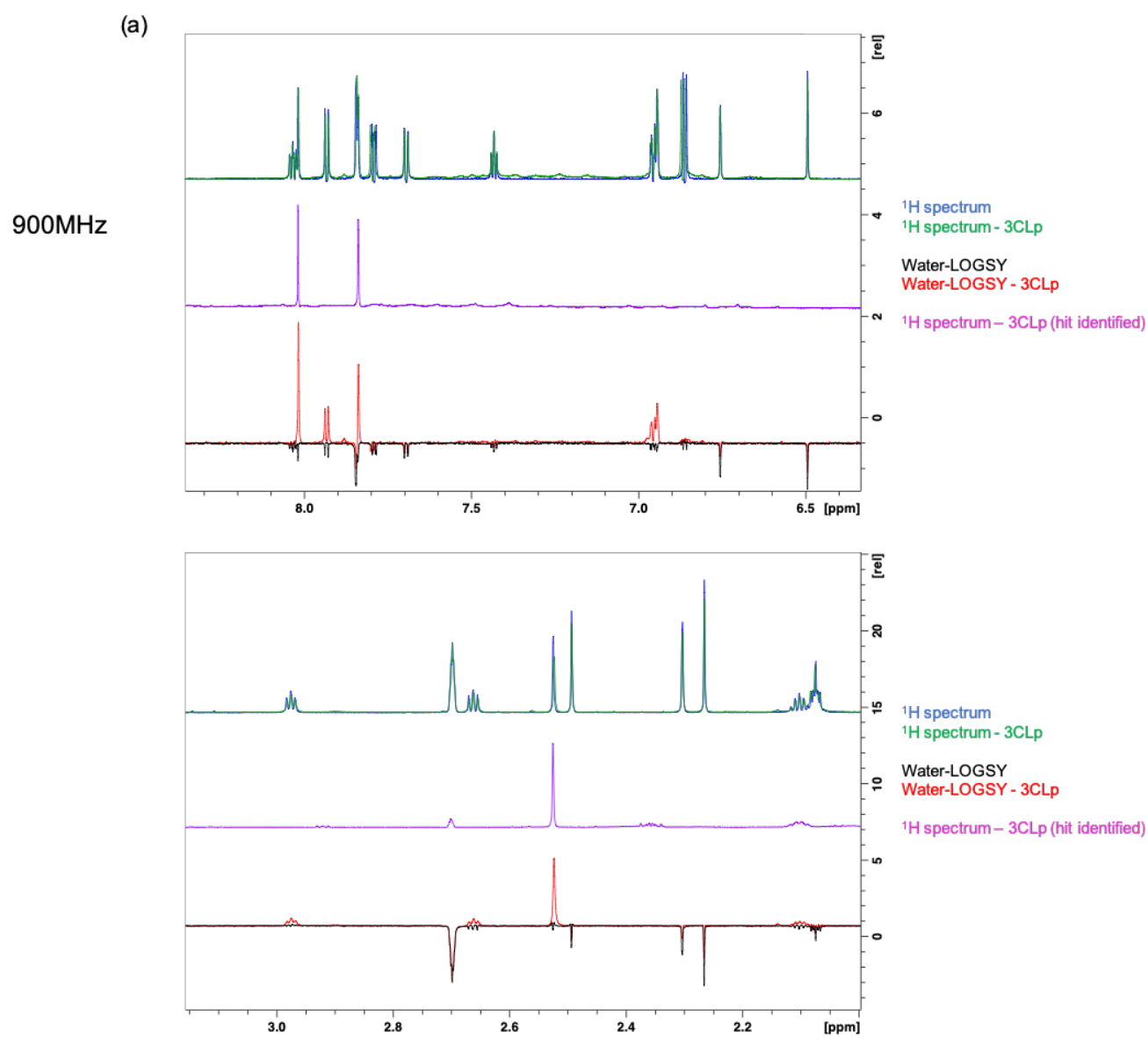


SUPPORTING INFORMATION

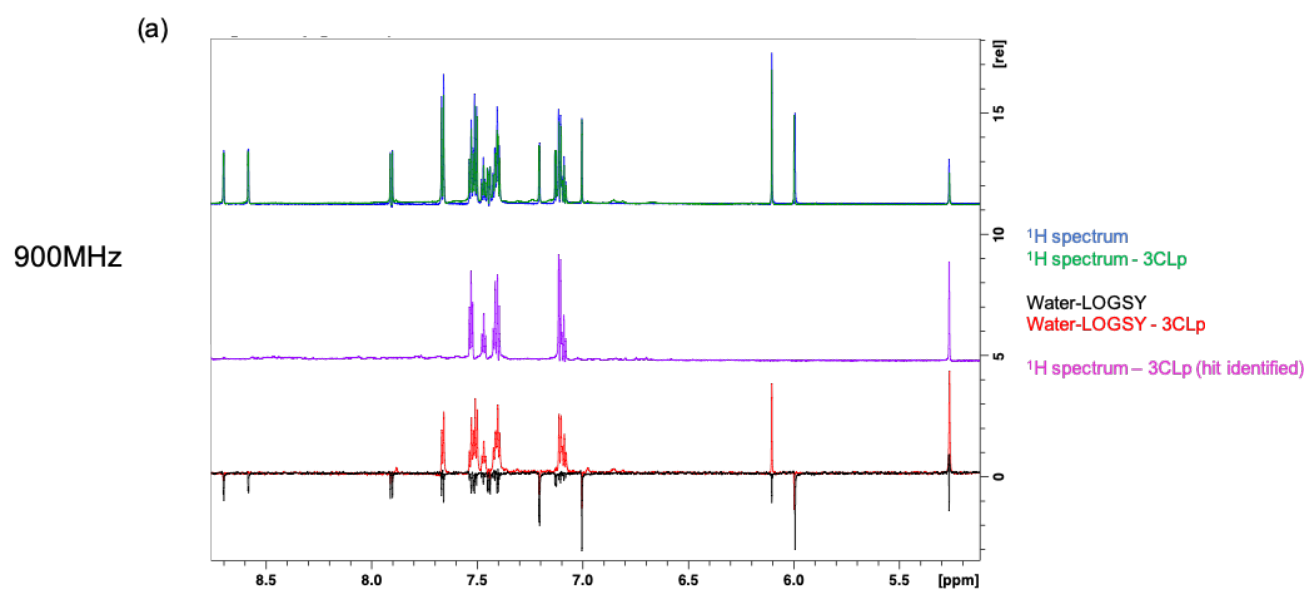
F28



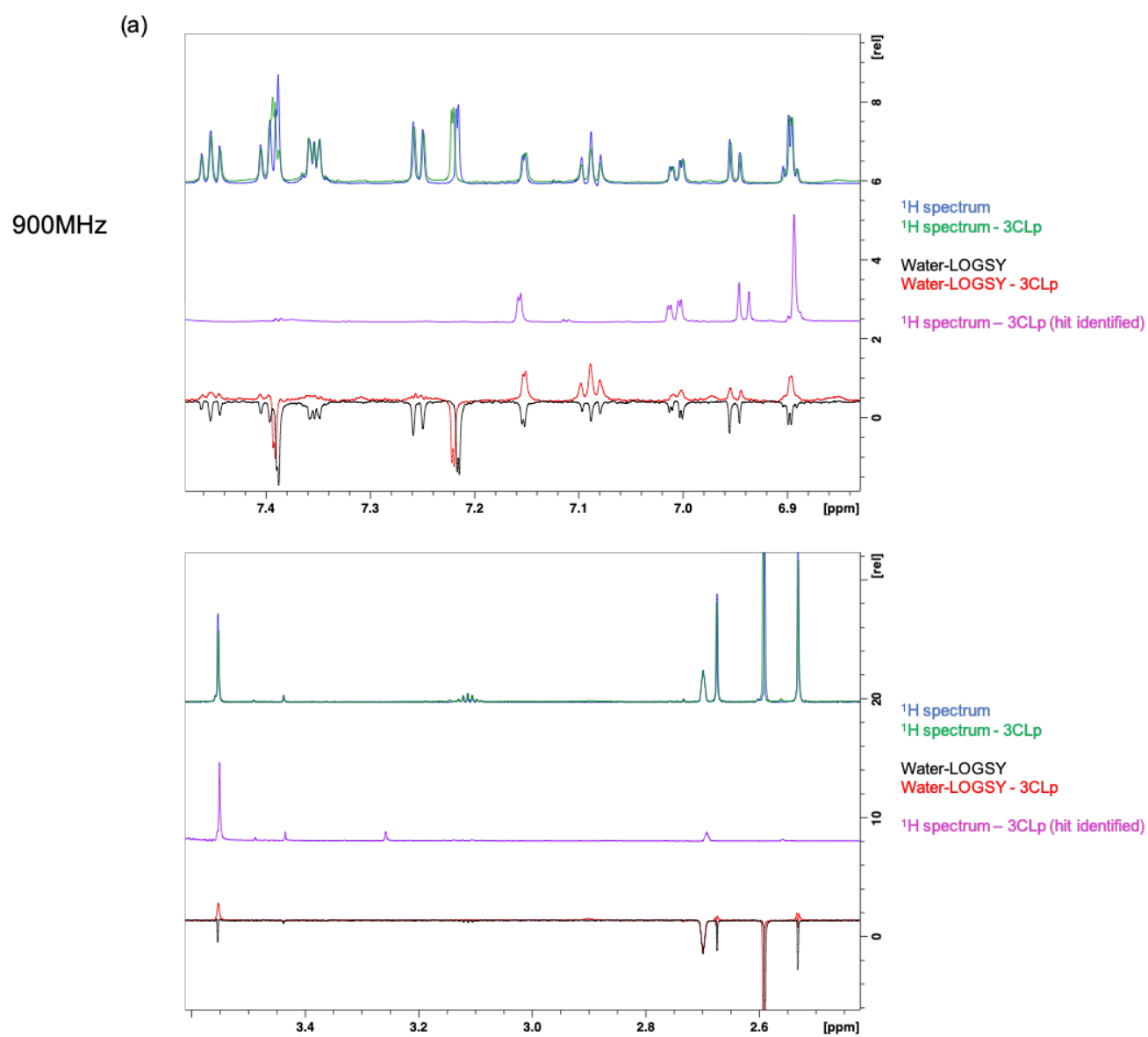
F21



F06

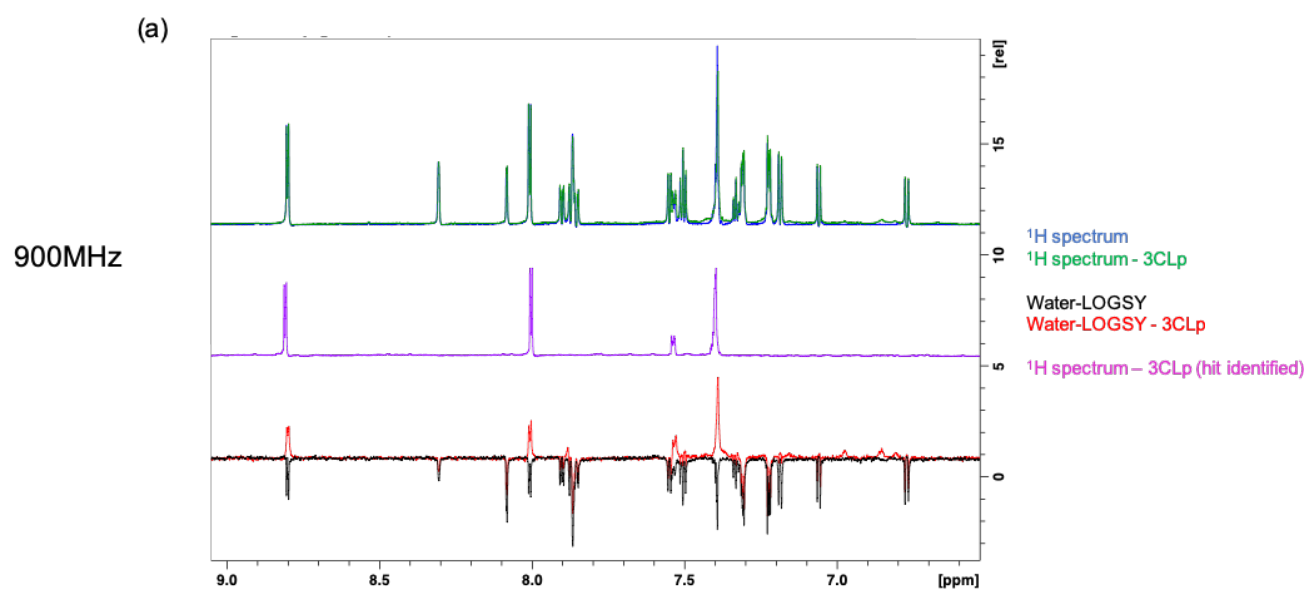


F02

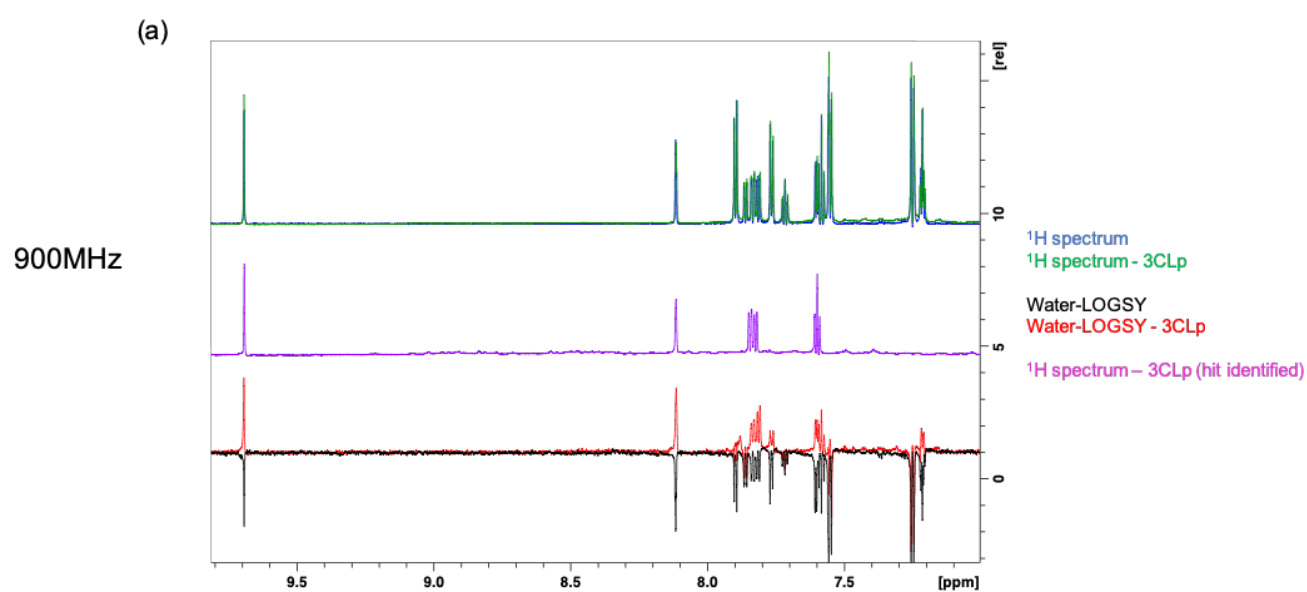


SUPPORTING INFORMATION

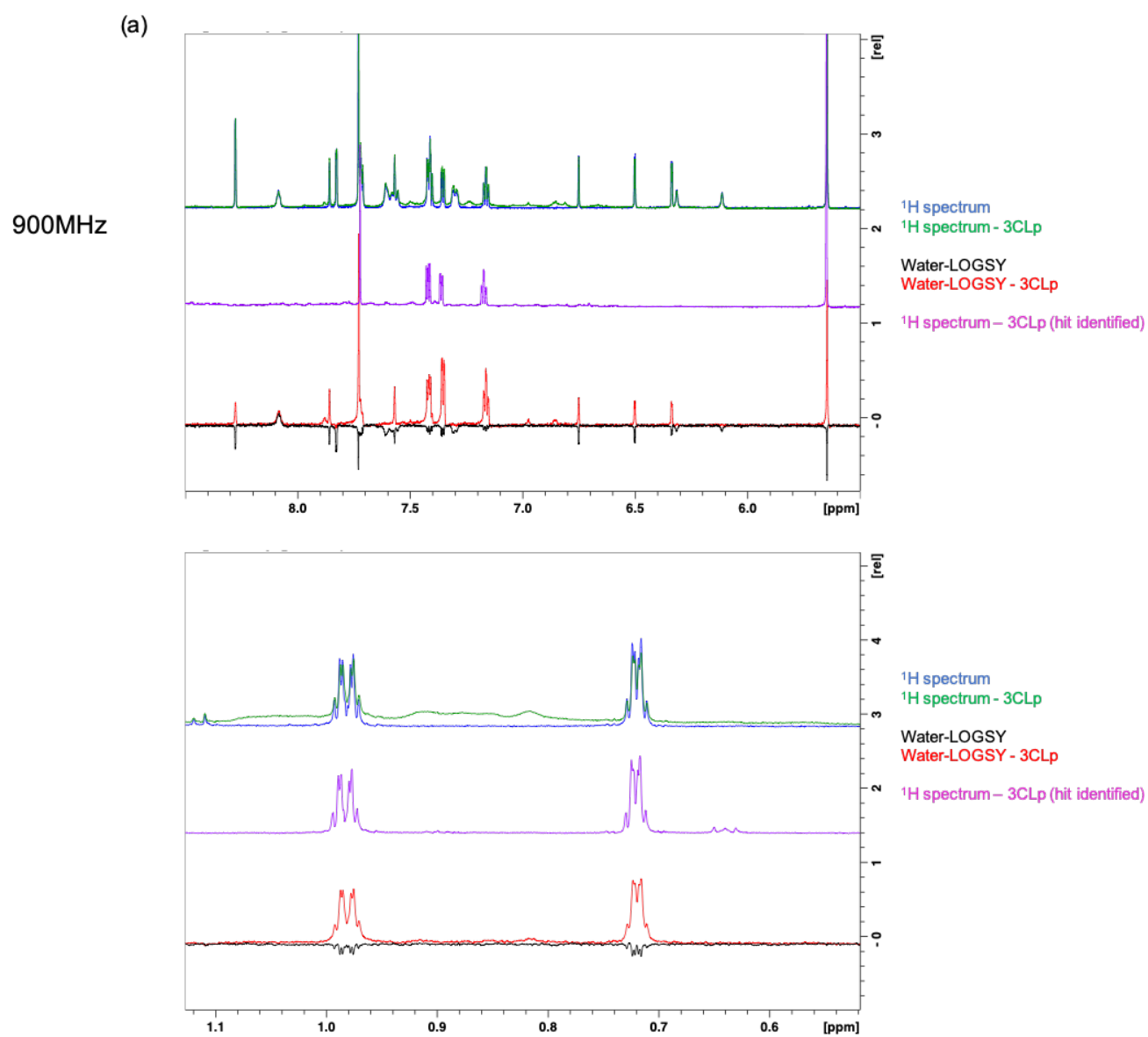
F08



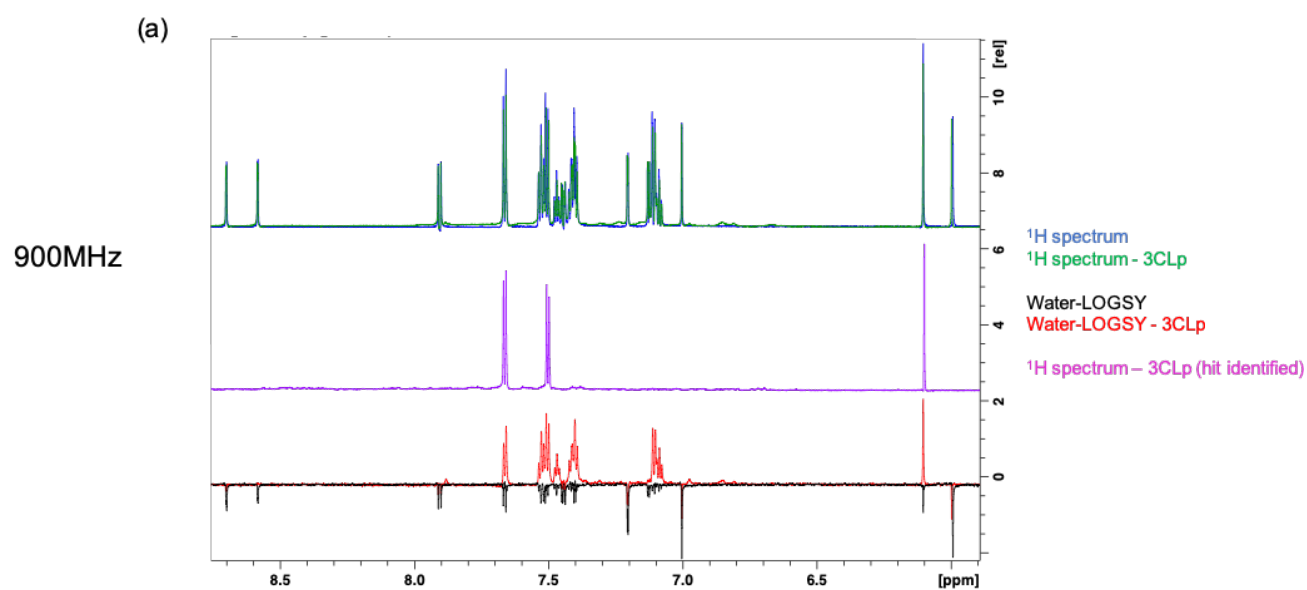
F29



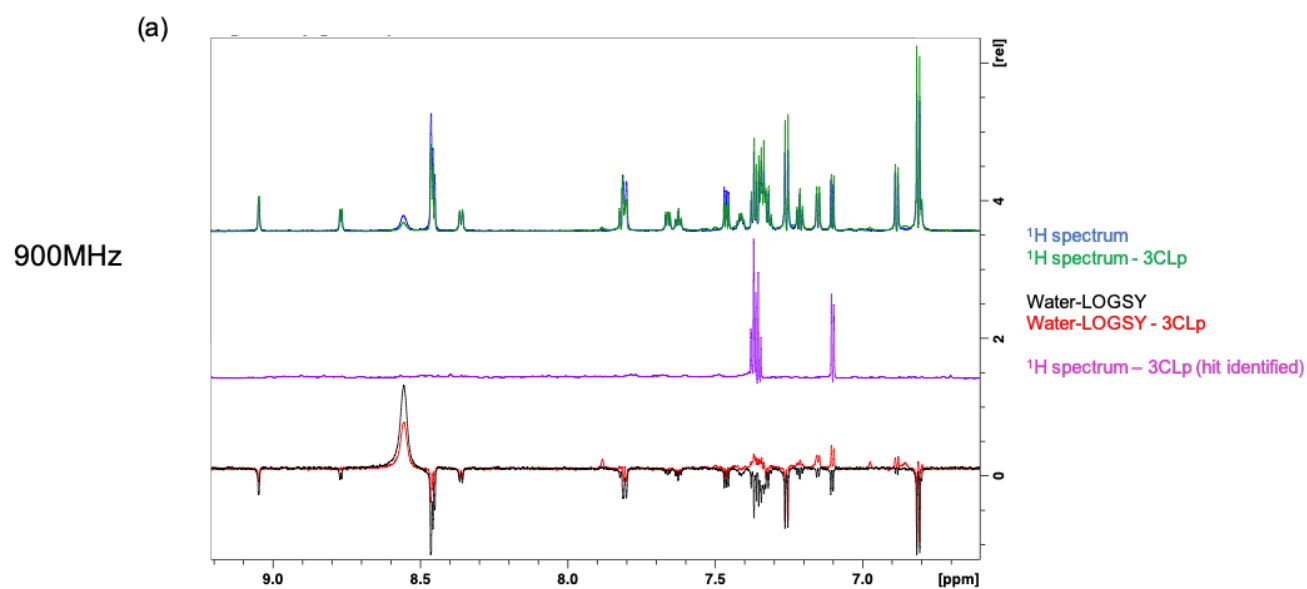
F33



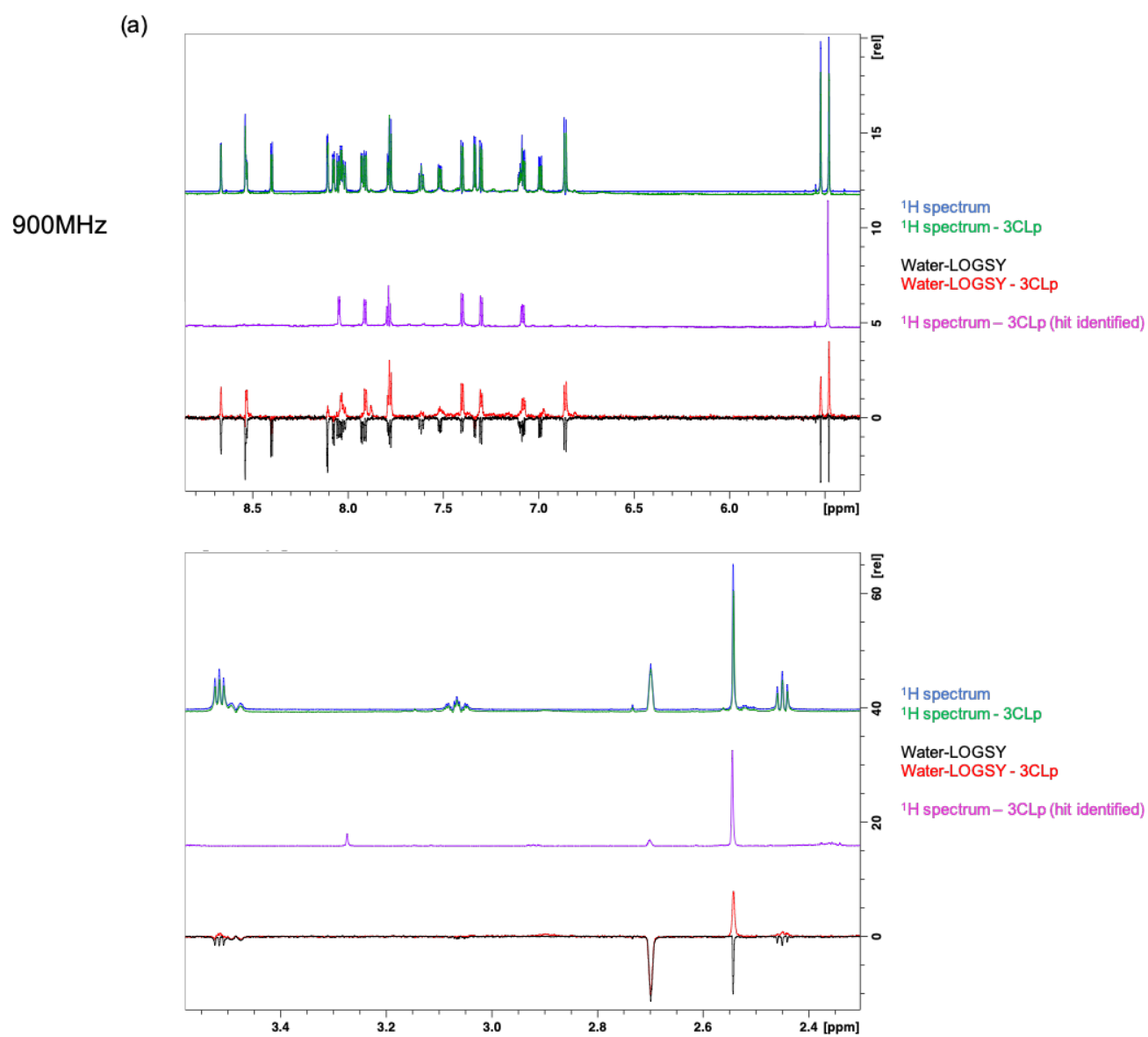
F07



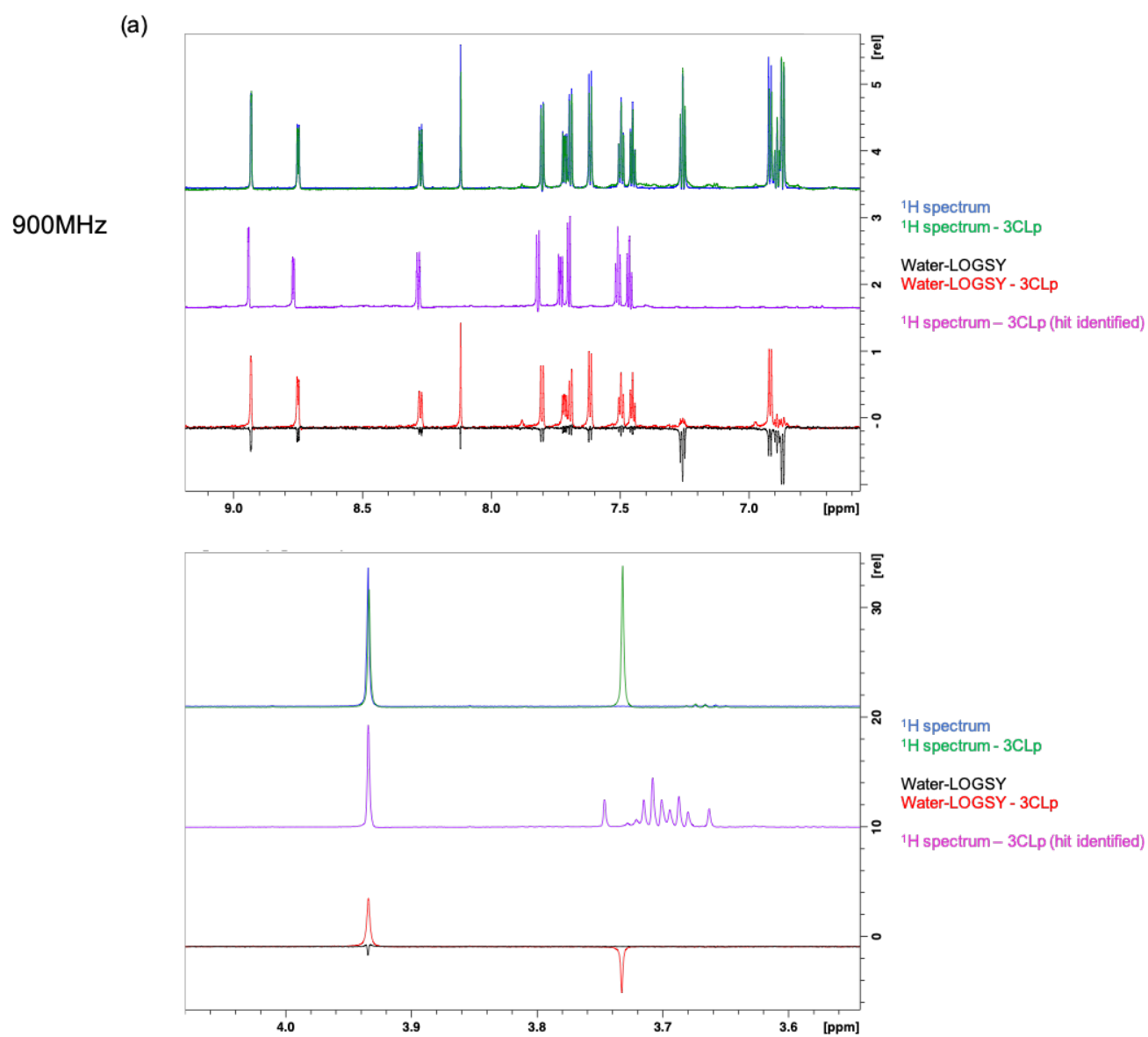
F19



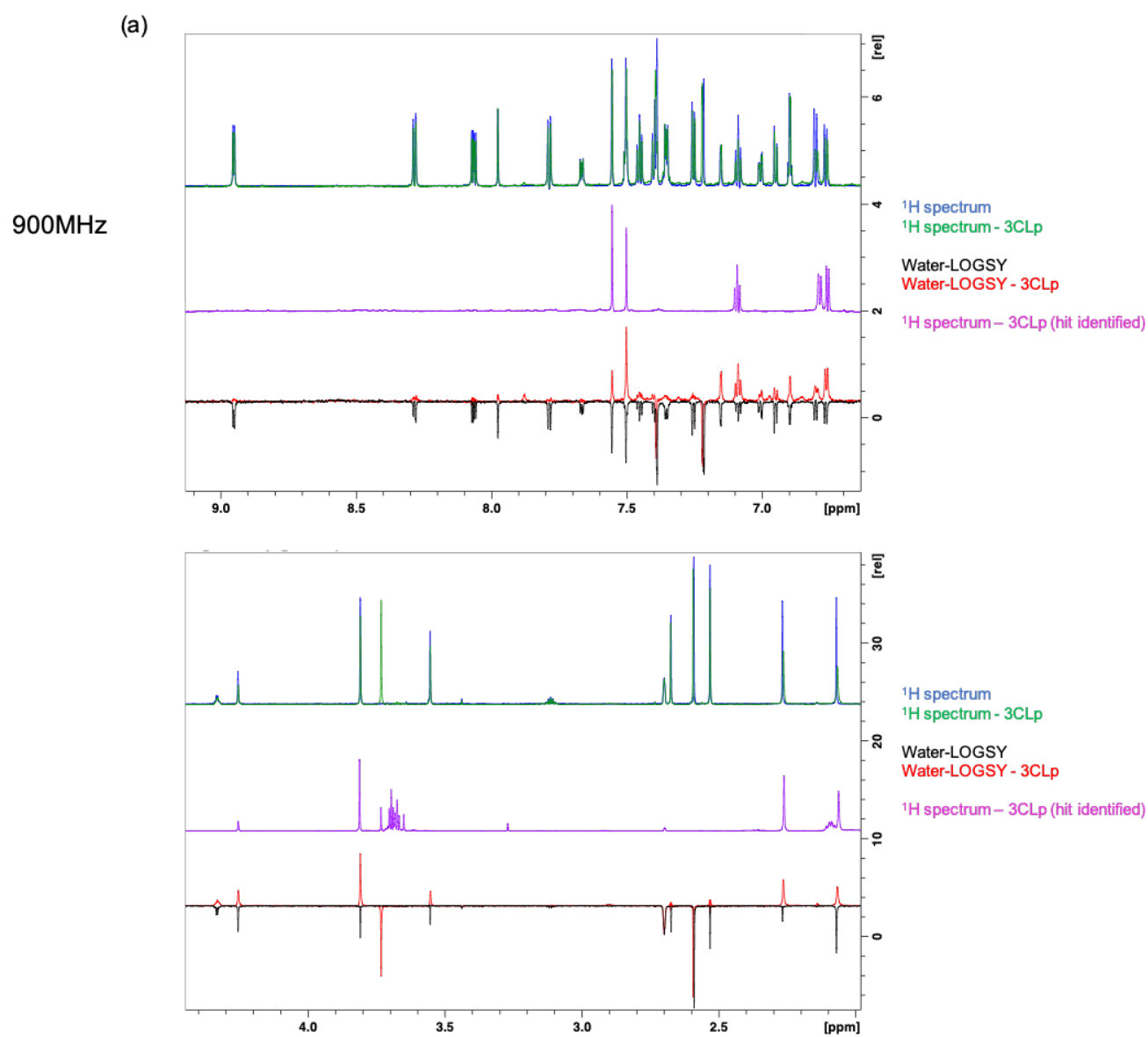
F36



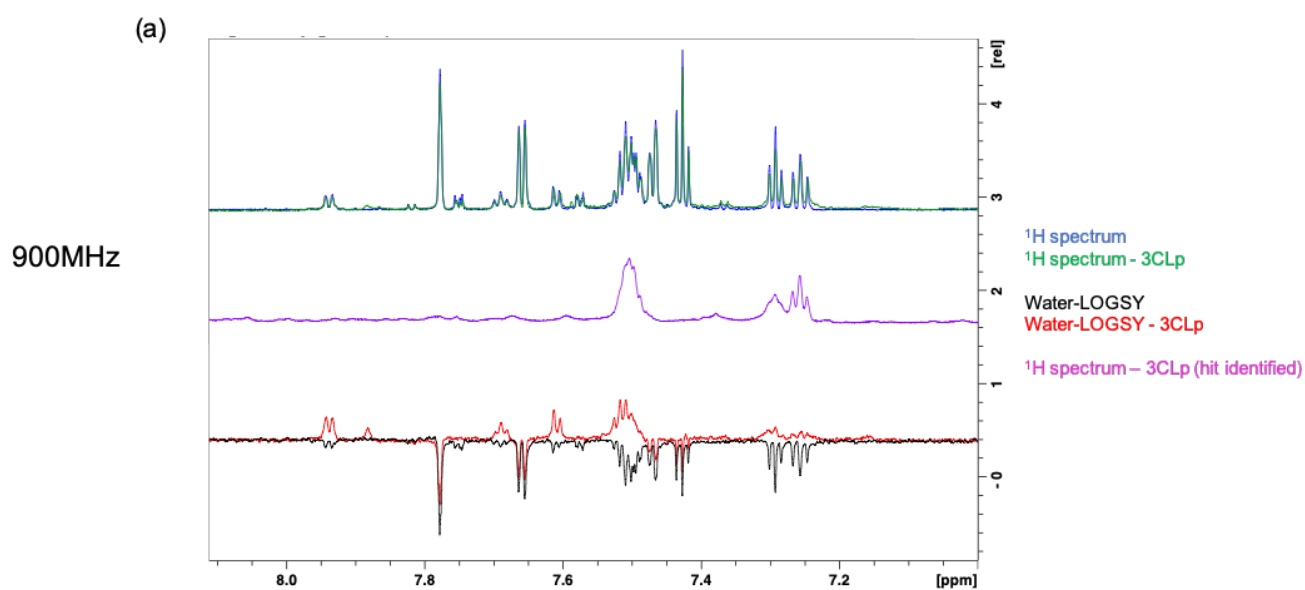
F26



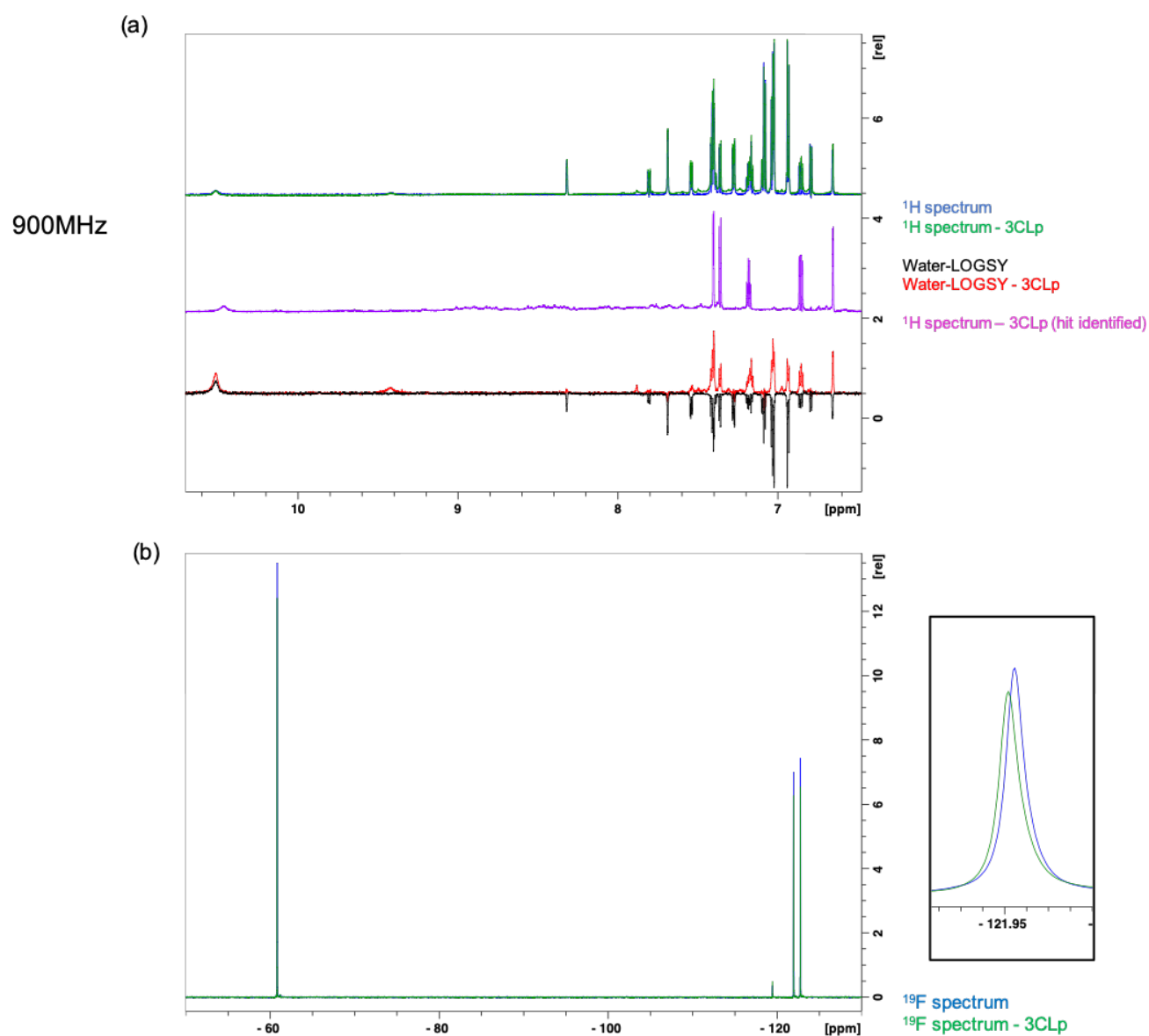
F15



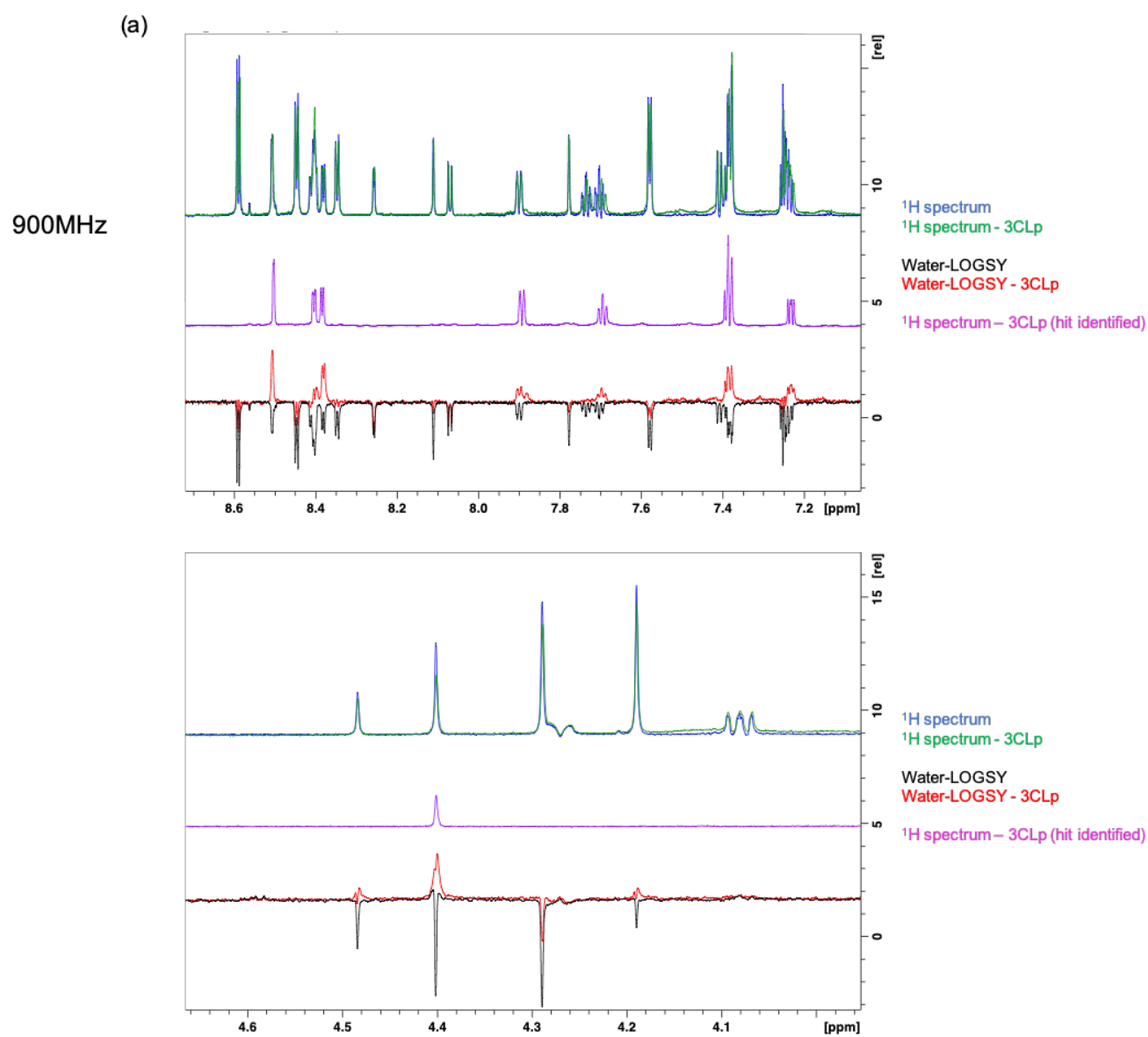
F31



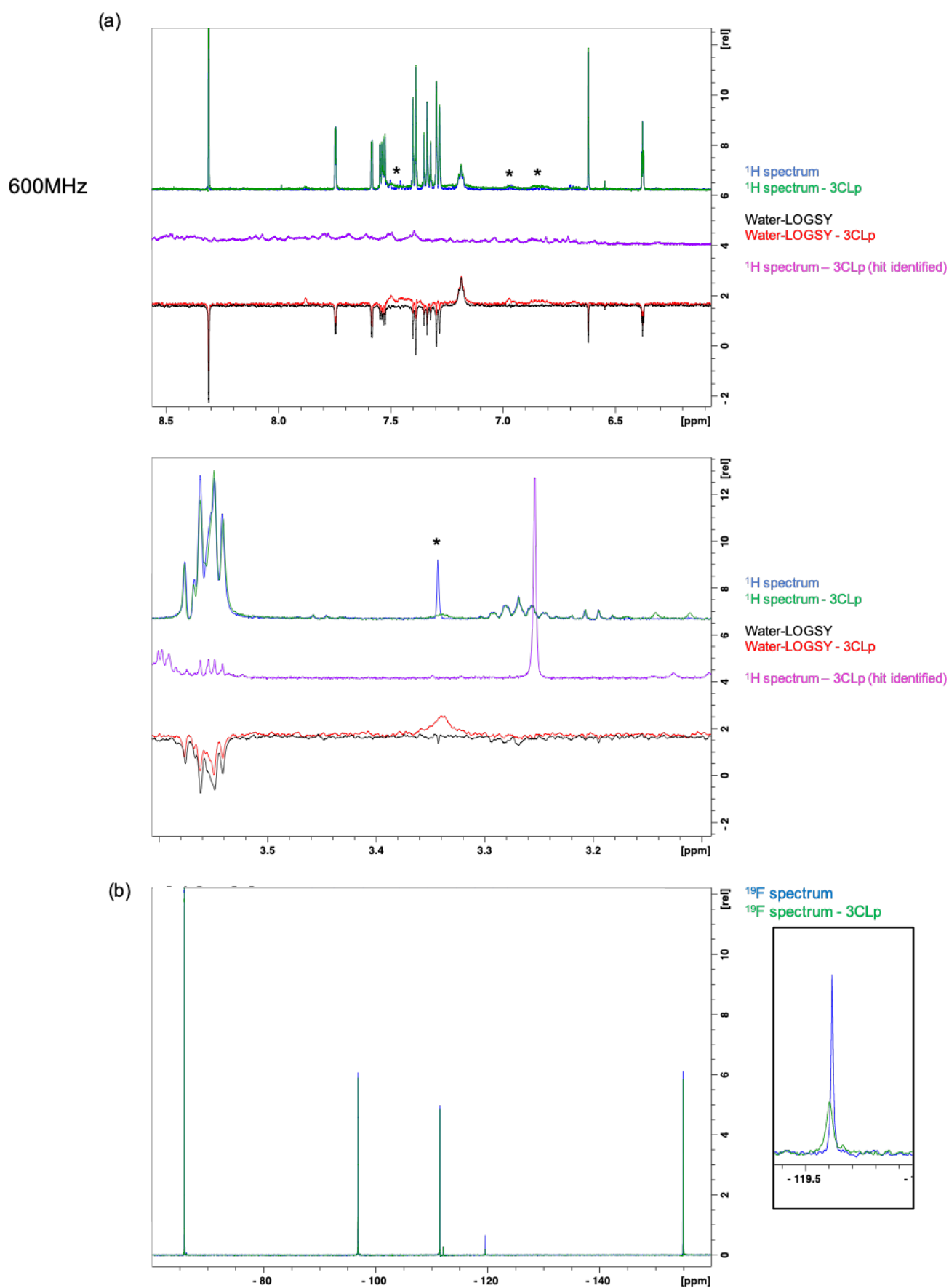
F20



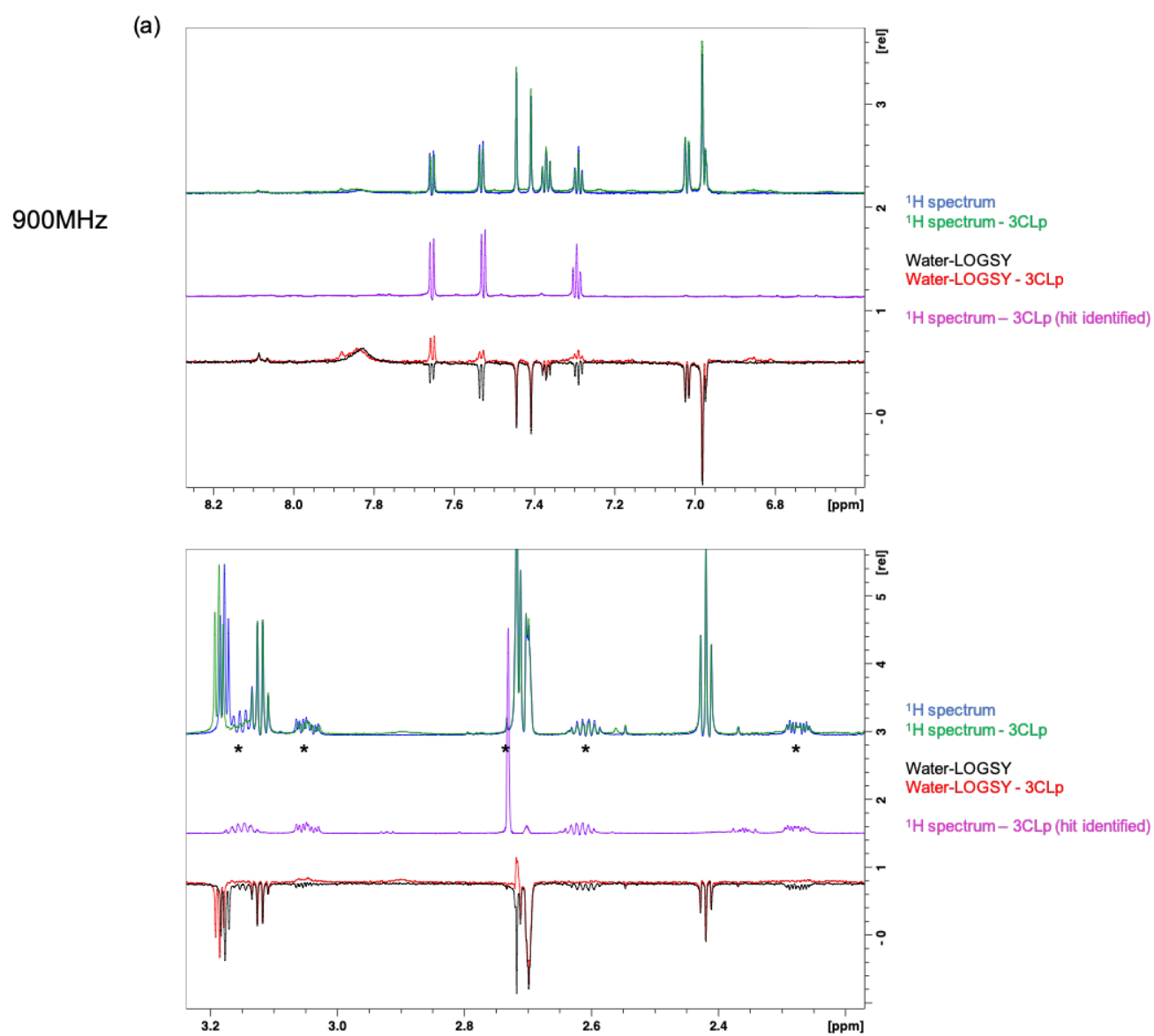
F38



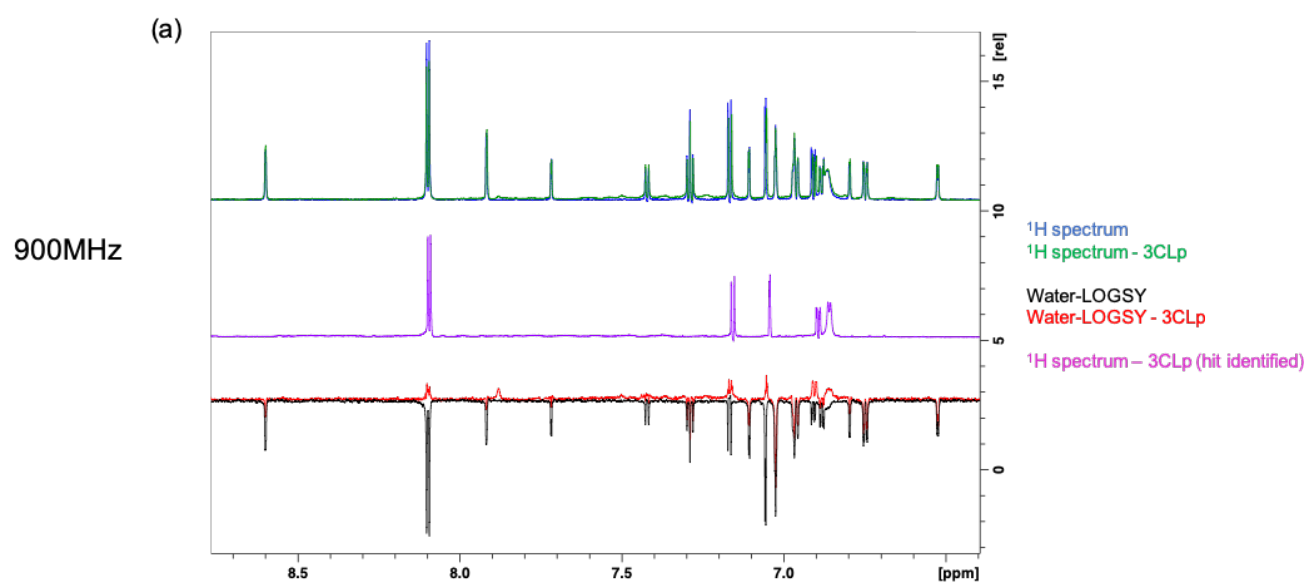
F25



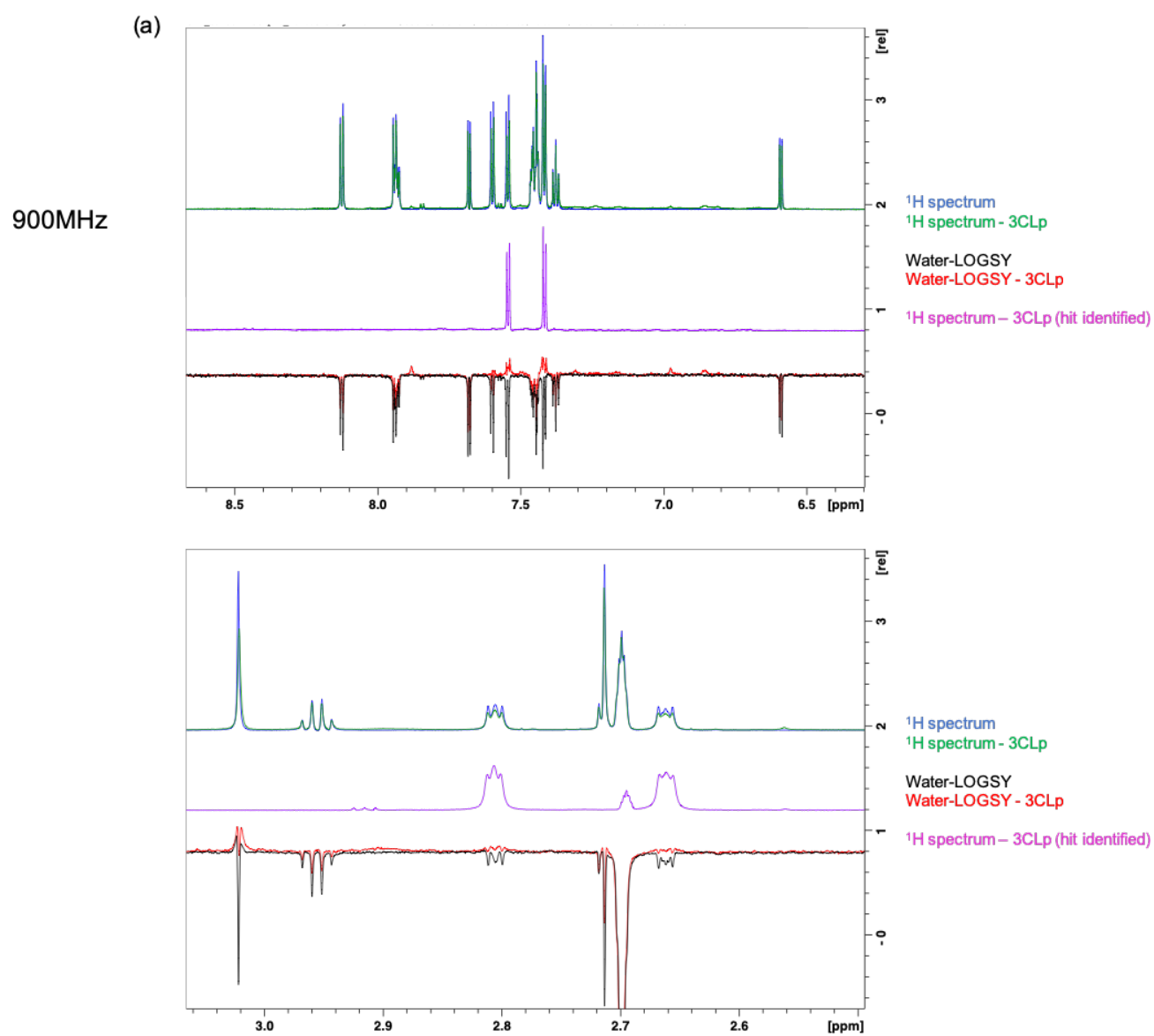
F30



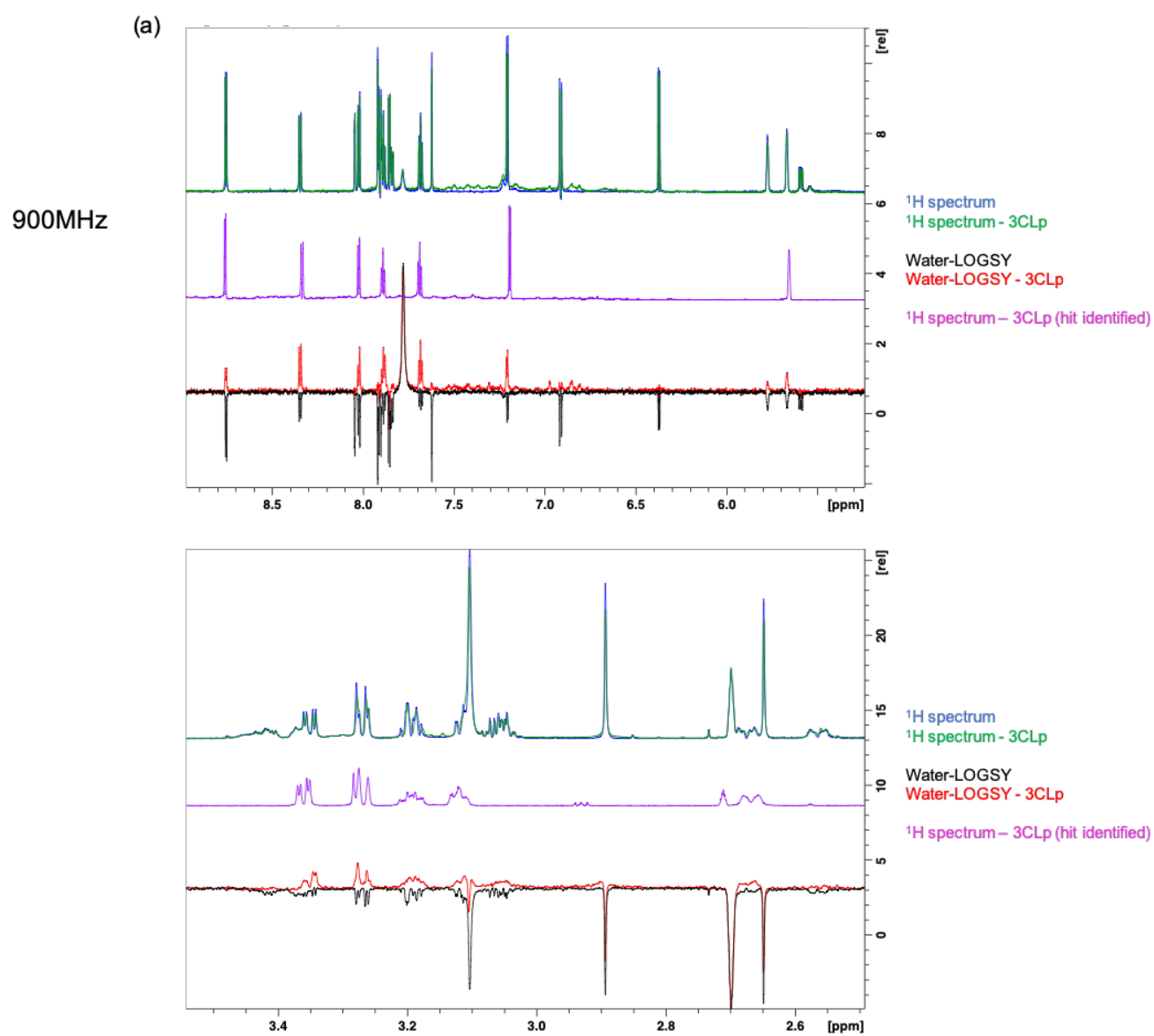
F22



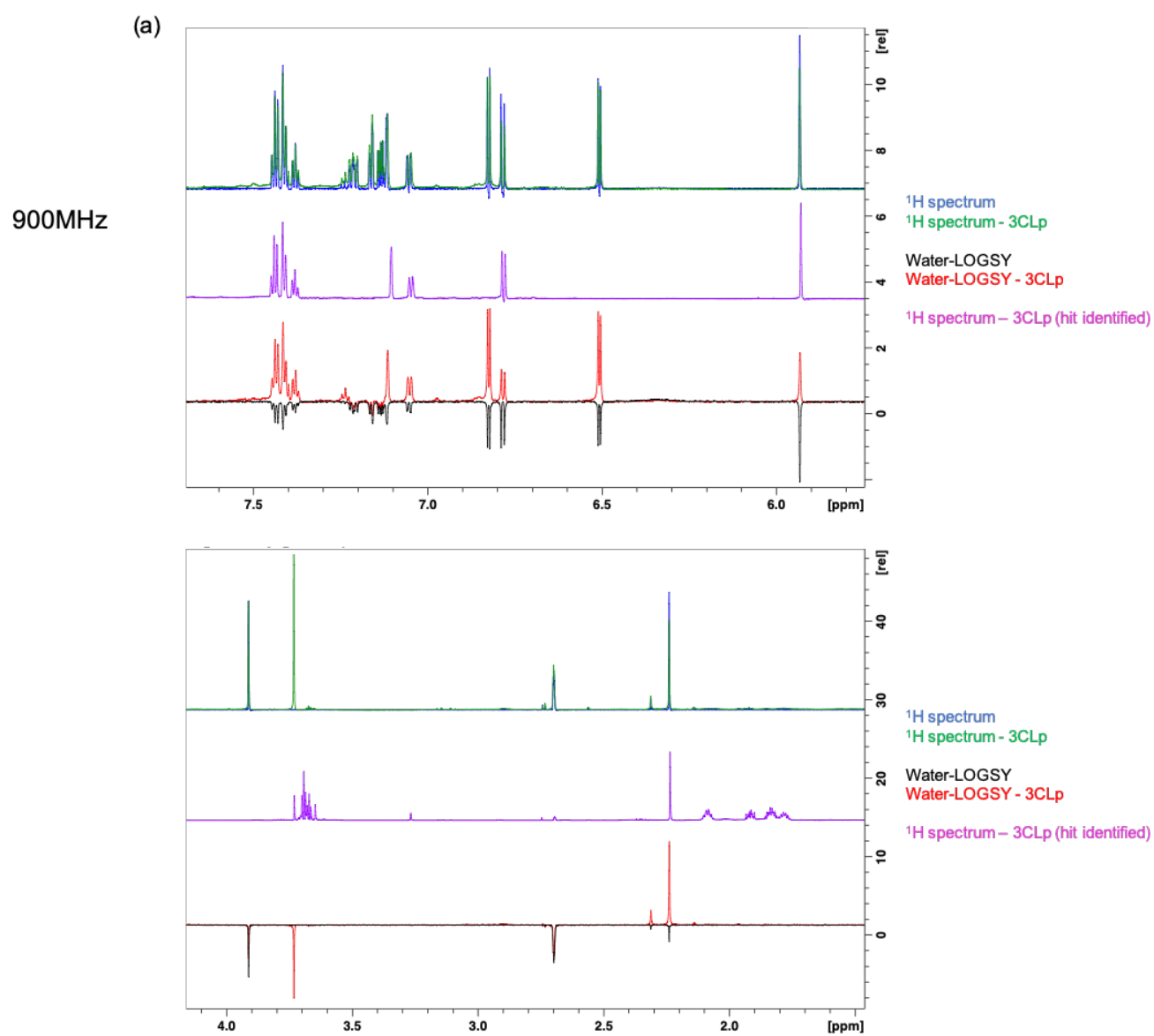
F09



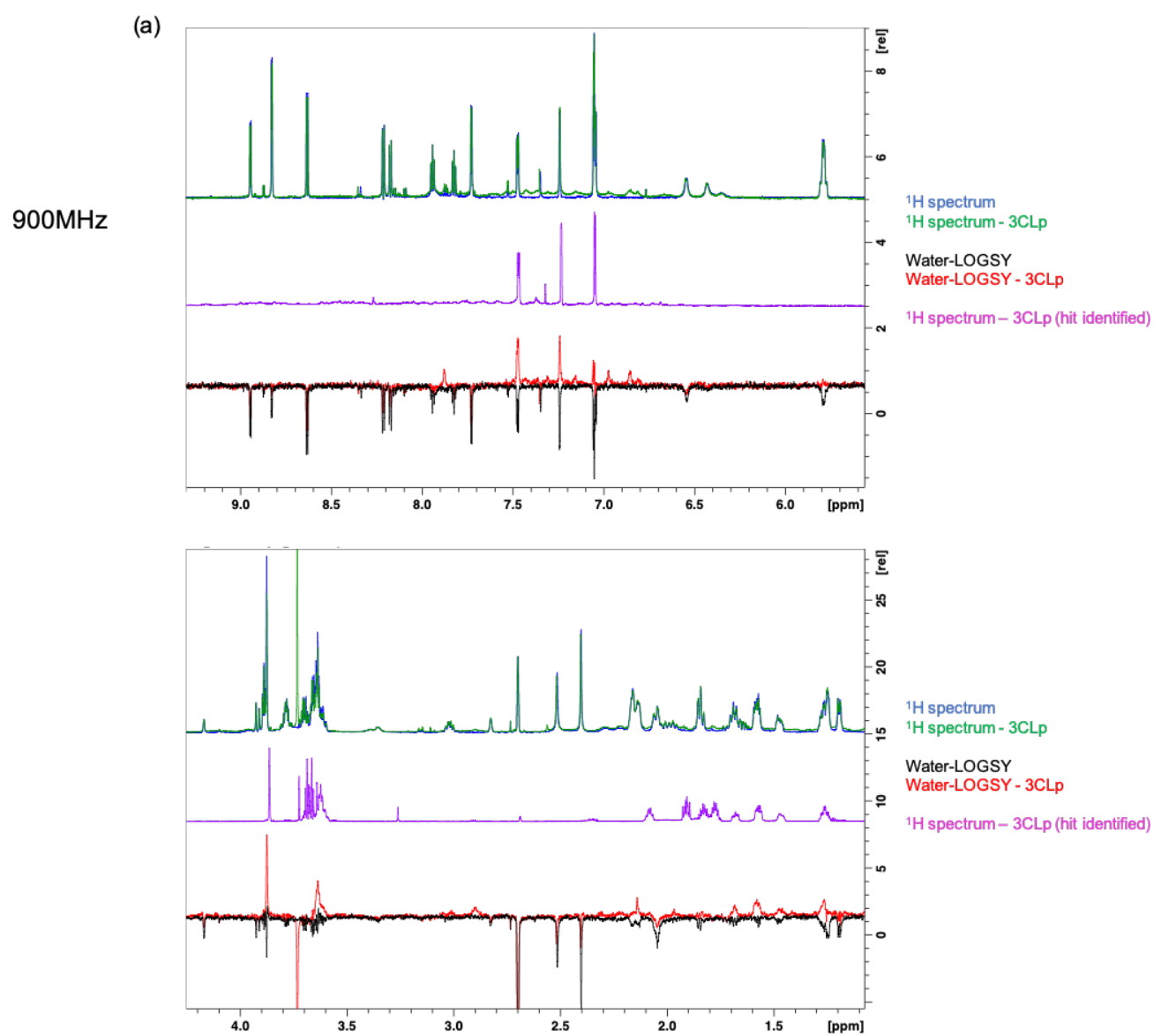
F36



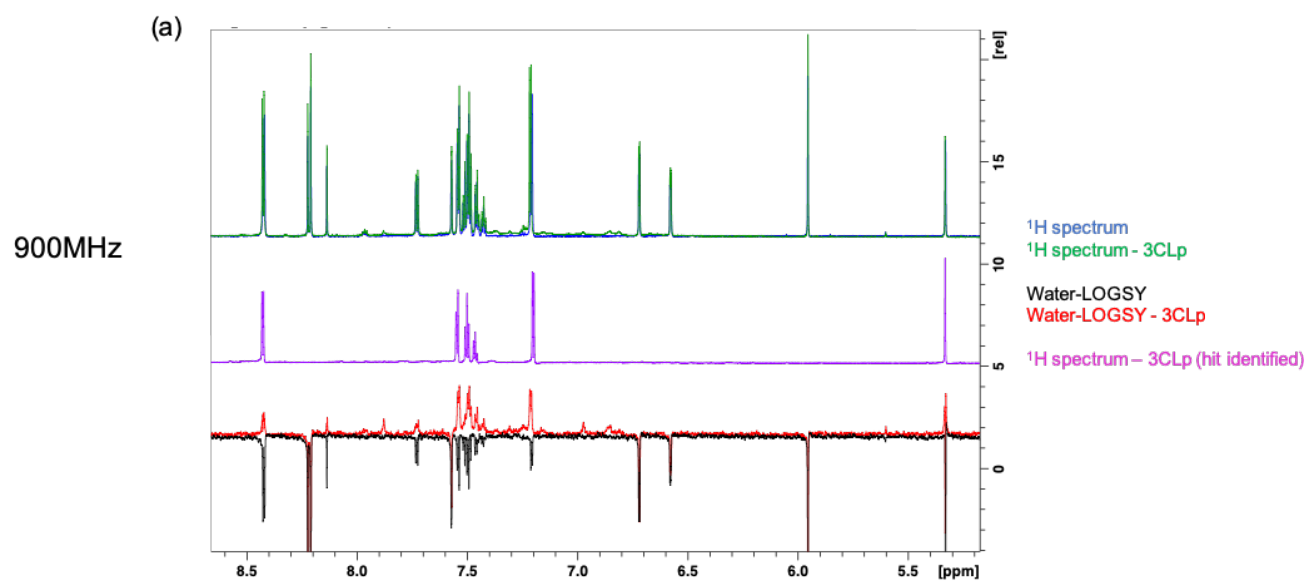
F14



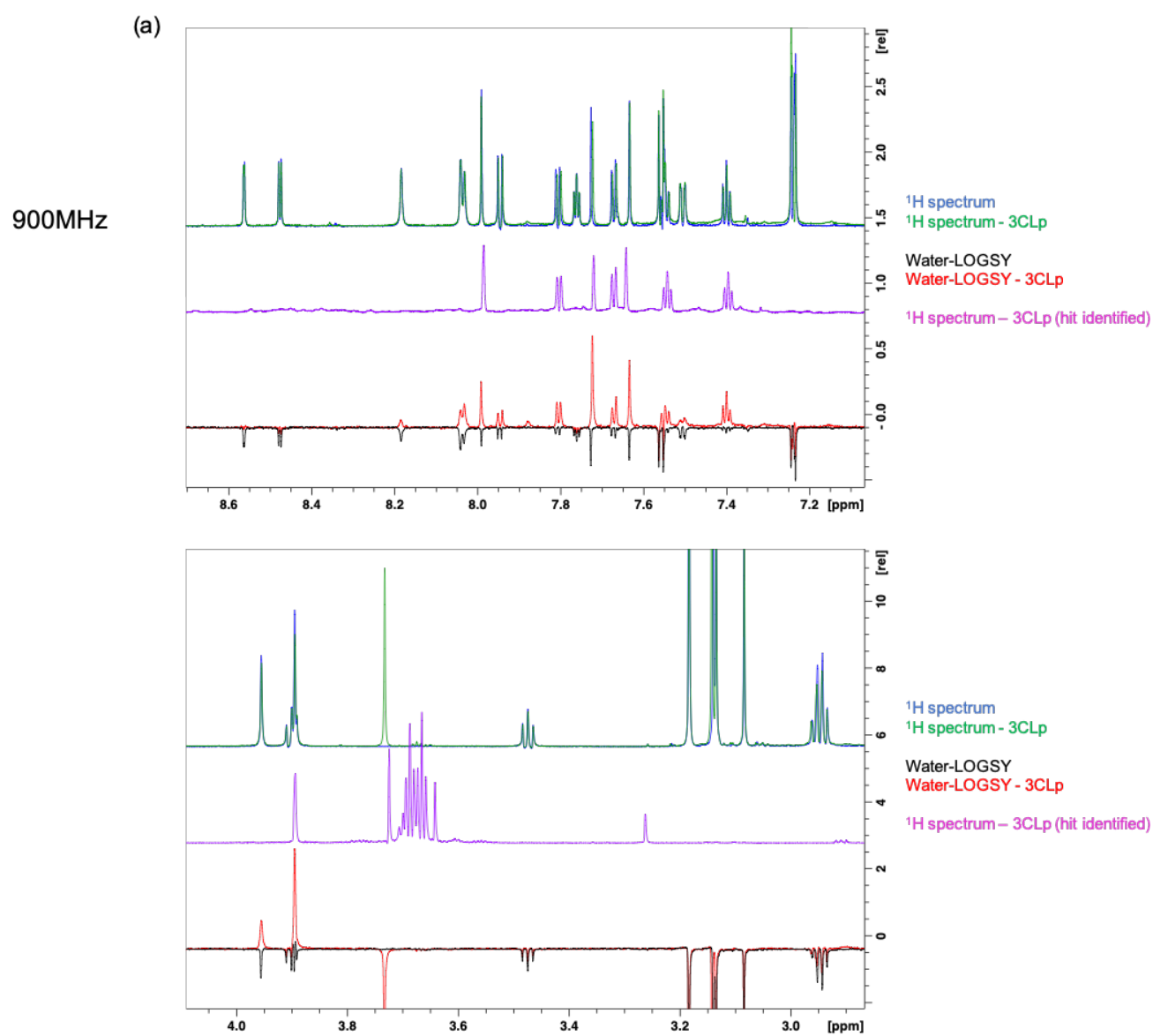
F34



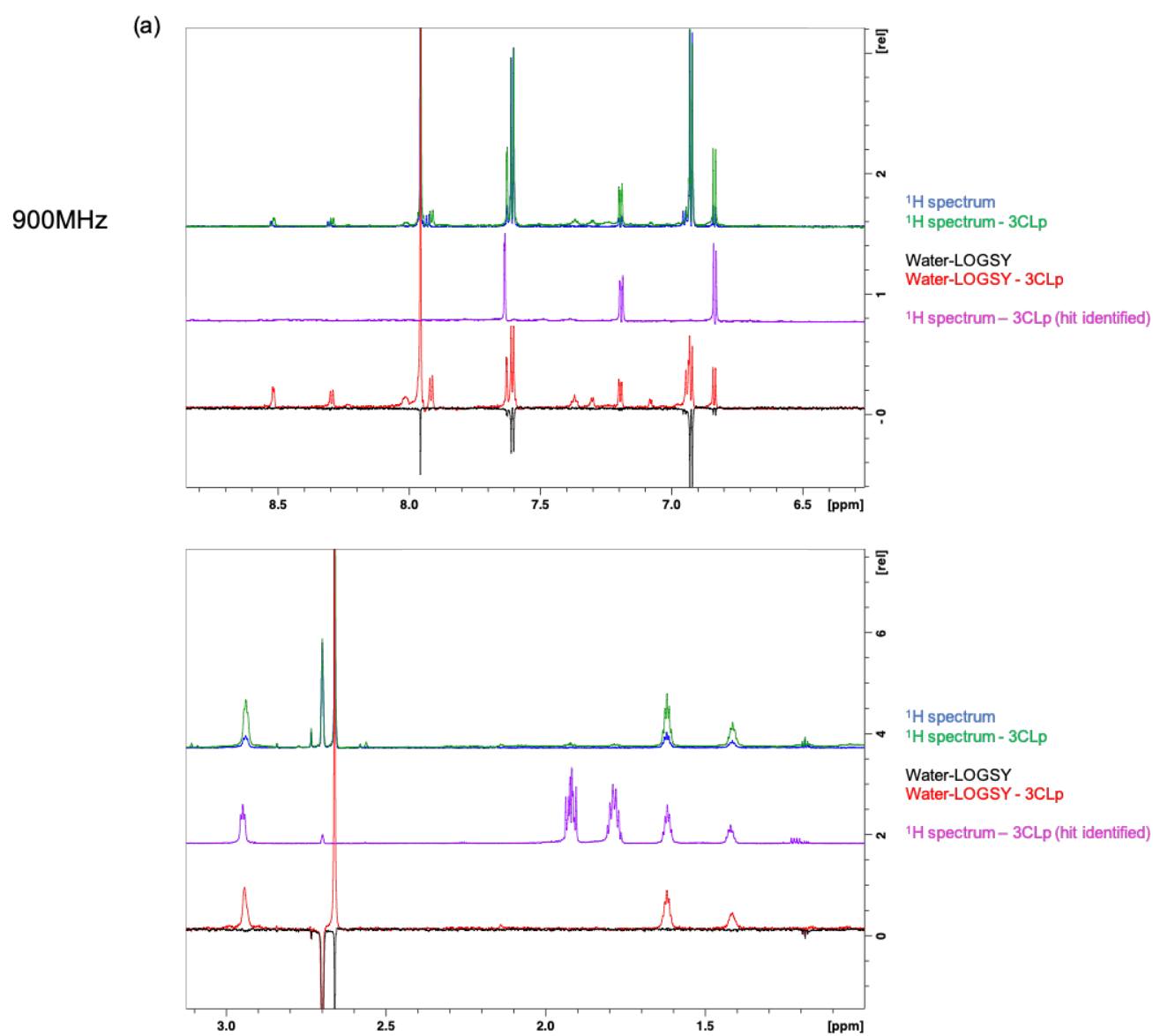
F32



F16



F27



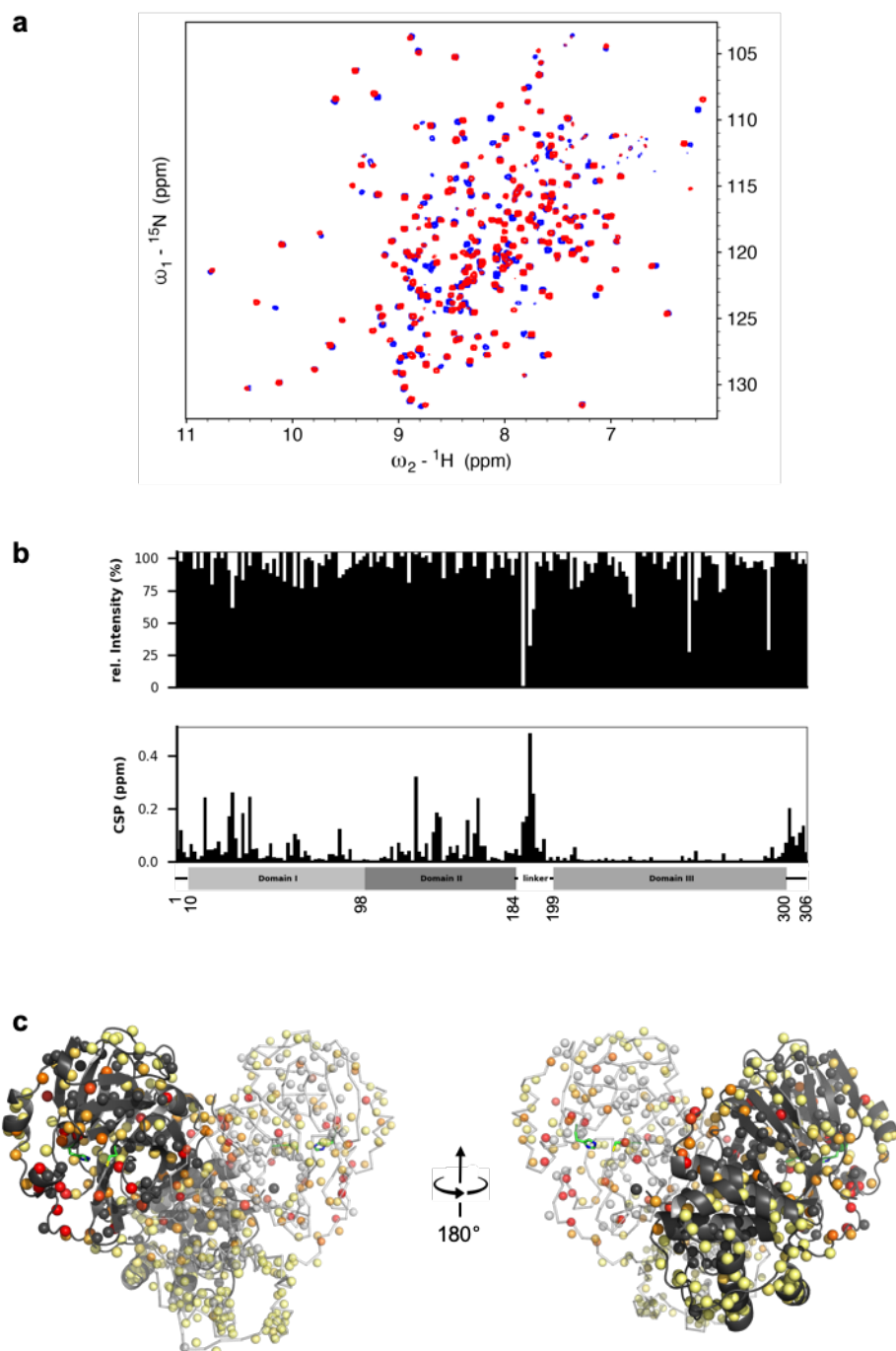
SUPPORTING INFORMATION

Scheme S2. Protein-based NMR secondary screening performed on ^2H , ^{15}N -3CLp. (a) Overlay of two 2D ^1H , ^{15}N -TROSY-HSQC spectra acquired on a ^2H , ^{15}N -3CLp sample (100 μM) in the absence (in blue) and in the presence (in red) of the individual fragment (2 mM). The DMSO- d_6 concentration was 3% in all experiments. The spectra were acquired, using 3mm tubes, at 305K on a 900 MHz spectrometer equipped with a ^1H cryogenic probe. (b) The spectral perturbations in the 3CLp spectrum induced upon fragment binding, corresponding both to the broadening of the resonances (rel. intens.) and to their ^1H and ^{15}N -combined chemical shift perturbations (CSP), are shown along the 3CLp sequence. The organization of 3CLp domains is shown at the bottom, with their boundaries numbered according to the amino-acid sequence. (c) Structure of the SARS-CoV-2 3CLp dimer (PDB: 7k3t), with protomers A and B shown in black and light grey, respectively. Each small ball represents a proton amide of the protein and thus should correspond to a resonance in the ^1H , ^{15}N 2D spectrum. The ^1H and ^{15}N -combined CSPs, from (a) and shown in (b, bottom), have been color coded (from light yellow to red) and are displayed on these balls. Unassigned residues were kept in the original color of the protomer. The side chains of the two catalytic residues, His41 and Cys145, are shown in green. Figures were prepared using PyMOL (PyMOL Molecular Graphics System, version 1.8, Schrodinger, LLC).

| | |
|-----|----|
| F01 | 62 |
| F18 | 63 |
| F35 | 64 |
| F04 | 65 |
| F23 | 66 |
| F12 | 67 |
| F13 | 68 |
| F10 | 69 |
| F24 | 70 |
| F03 | 71 |
| F11 | 72 |
| F05 | 73 |
| F17 | 74 |
| F28 | 75 |
| F21 | 76 |
| F06 | 77 |
| F02 | 78 |
| F08 | 79 |
| F29 | 80 |
| F33 | 81 |
| F07 | 82 |
| F19 | 83 |
| F36 | 84 |
| F26 | 85 |
| F15 | 86 |
| F31 | 87 |
| F20 | 88 |
| F38 | 89 |
| F25 | 90 |
| F30 | 91 |
| F22 | 92 |
| F09 | 93 |
| F36 | 94 |
| F14 | 95 |
| F34 | 96 |
| F32 | 97 |
| F16 | 98 |
| F27 | 99 |

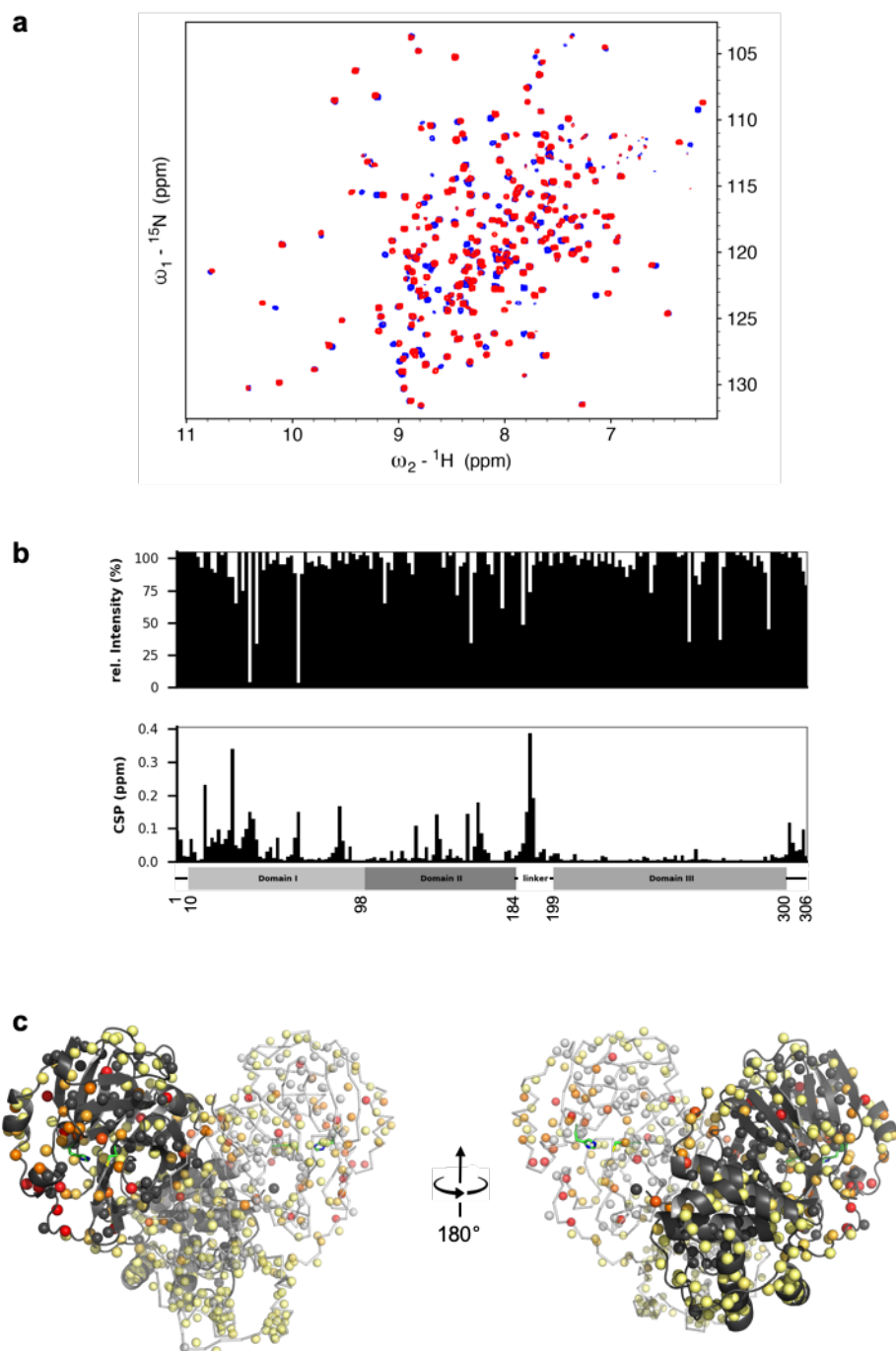
SUPPORTING INFORMATION

F01

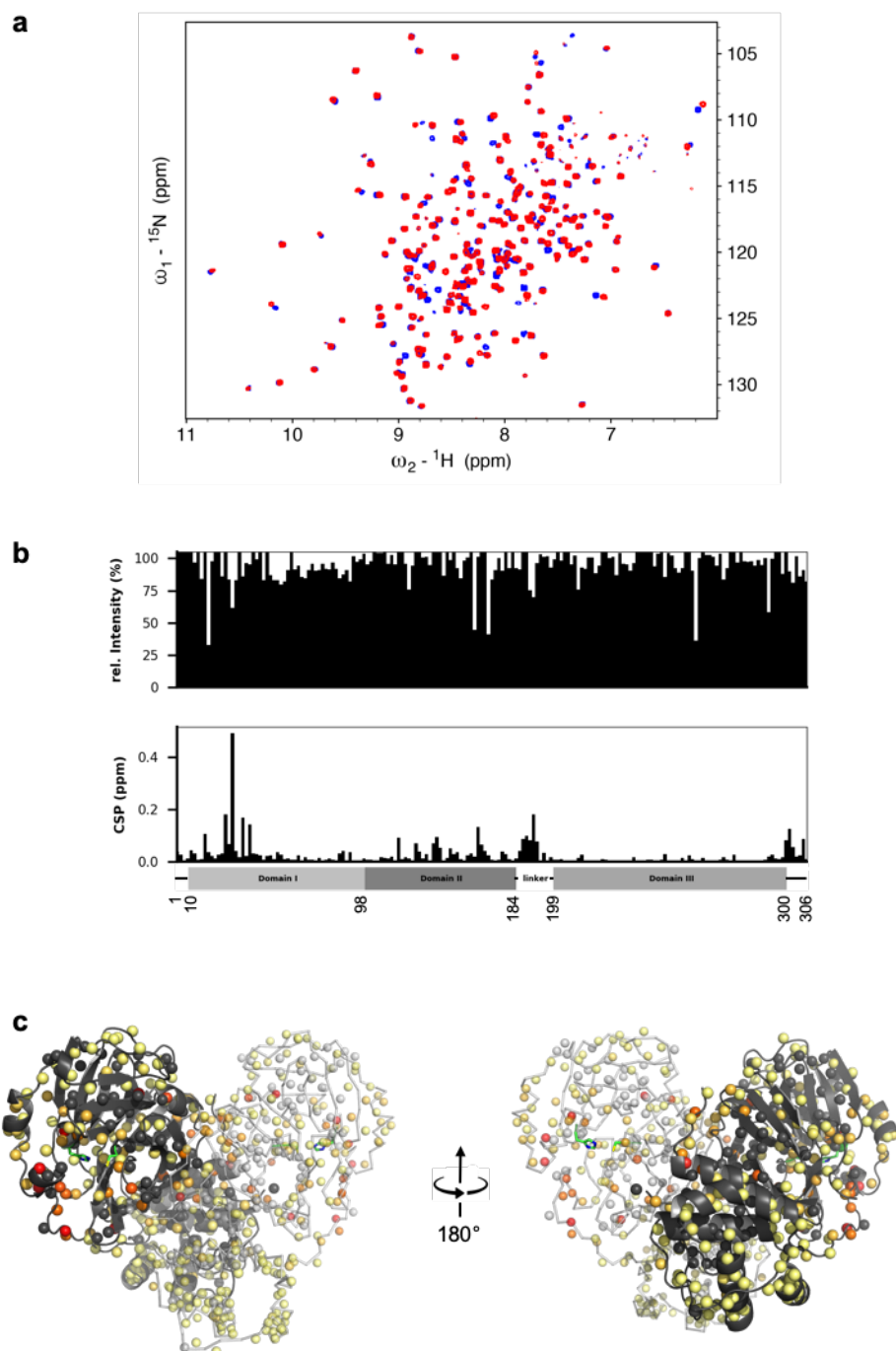


SUPPORTING INFORMATION

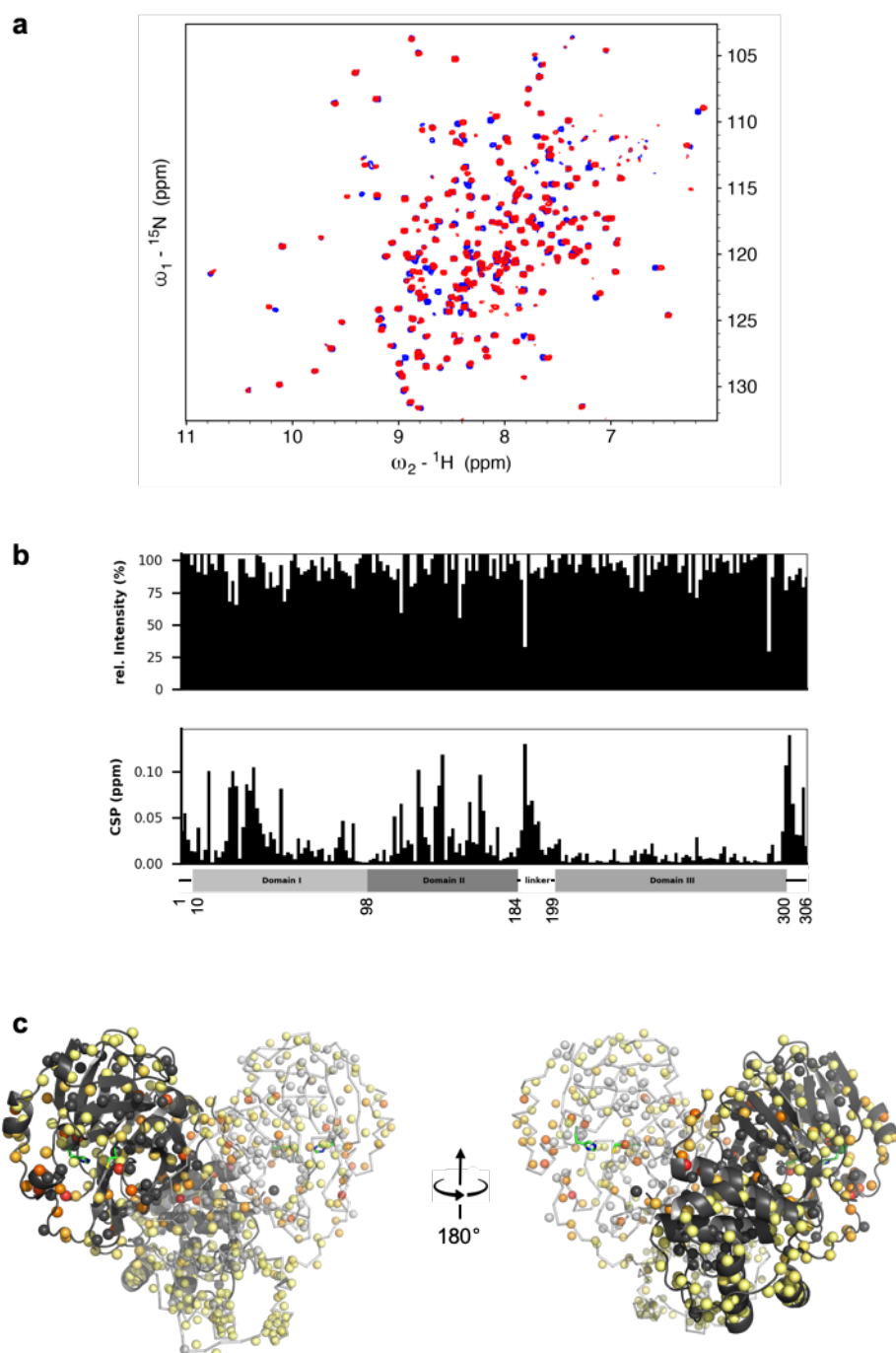
F18



F35

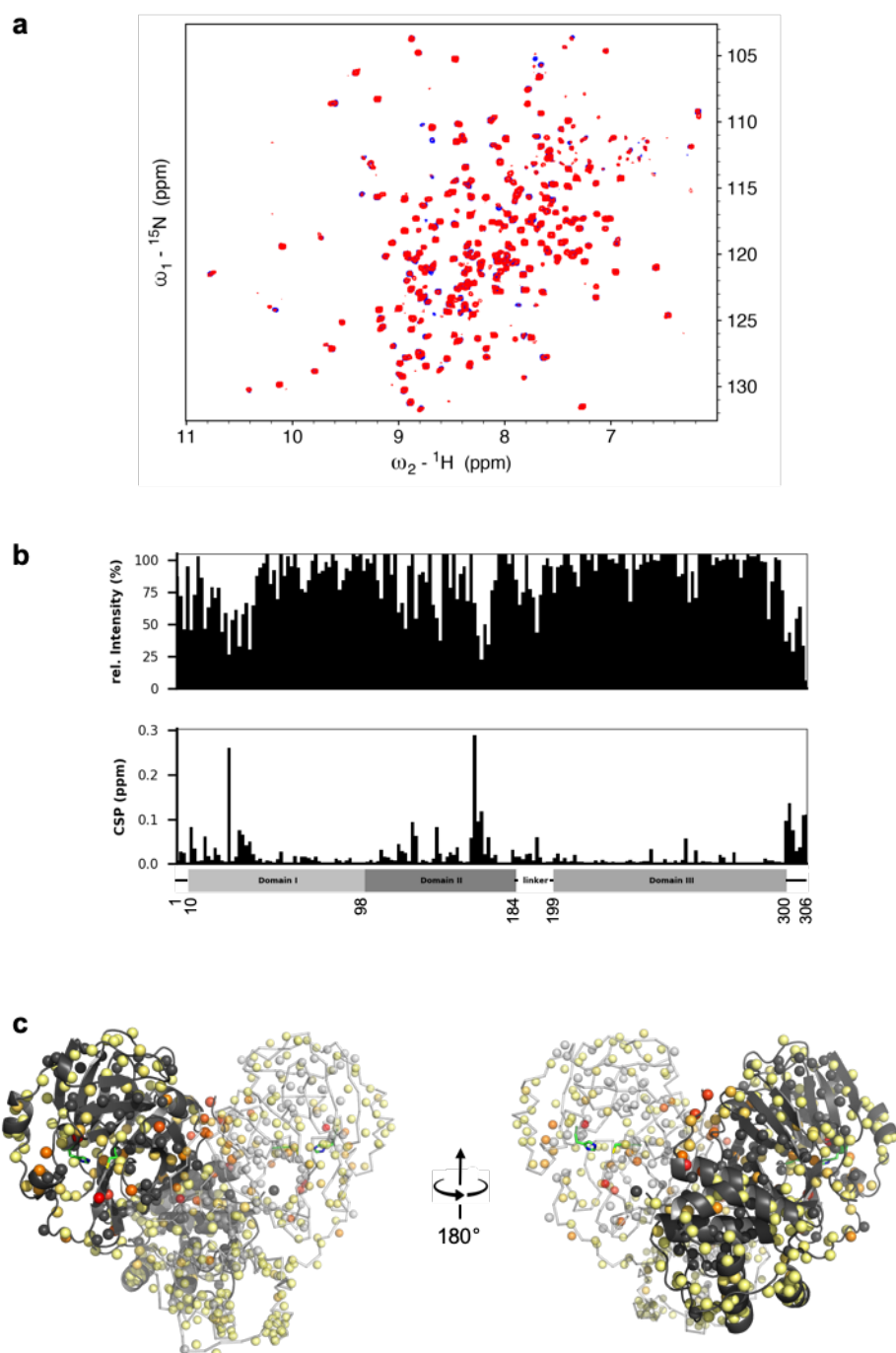


F04

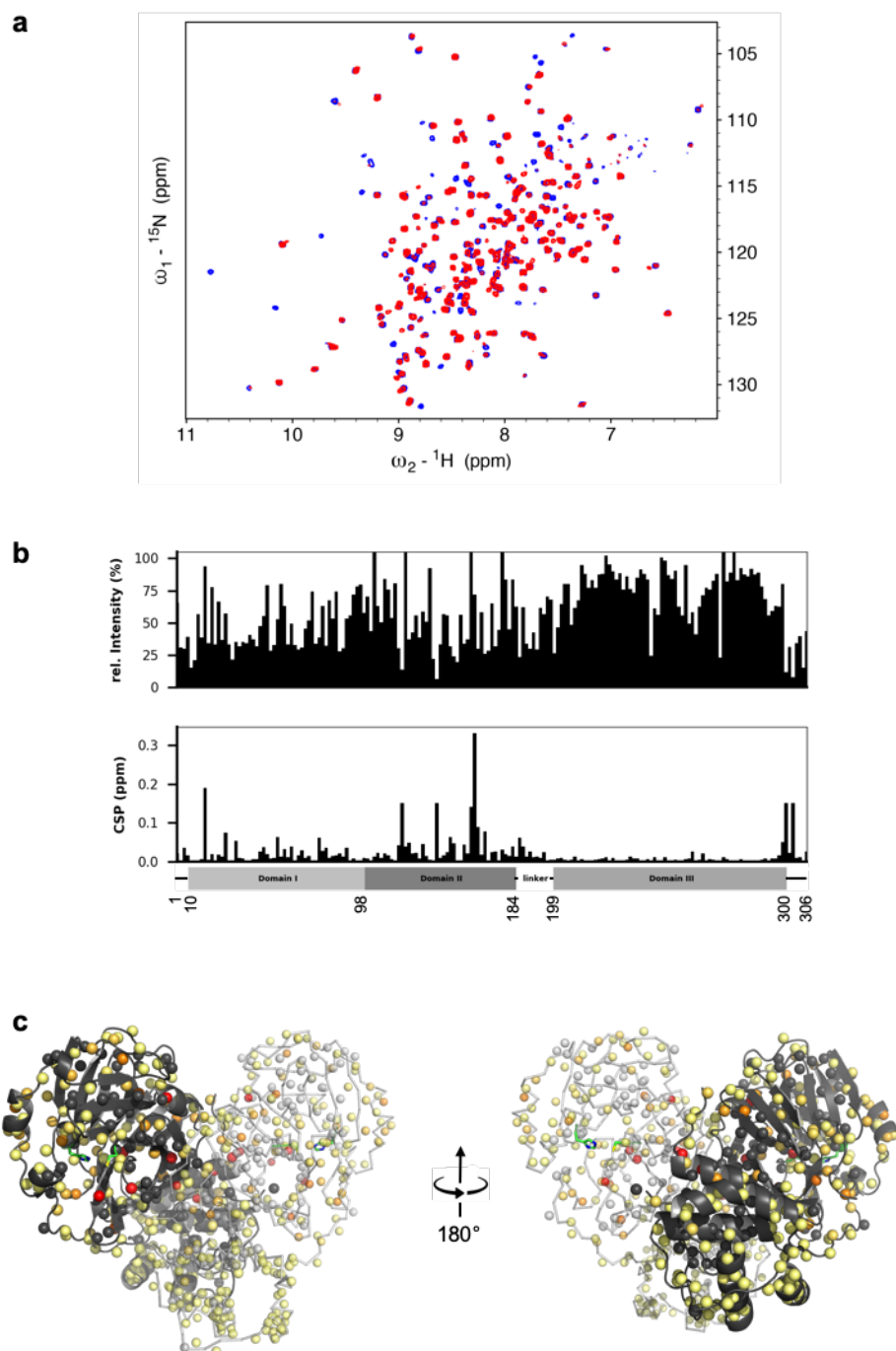


SUPPORTING INFORMATION

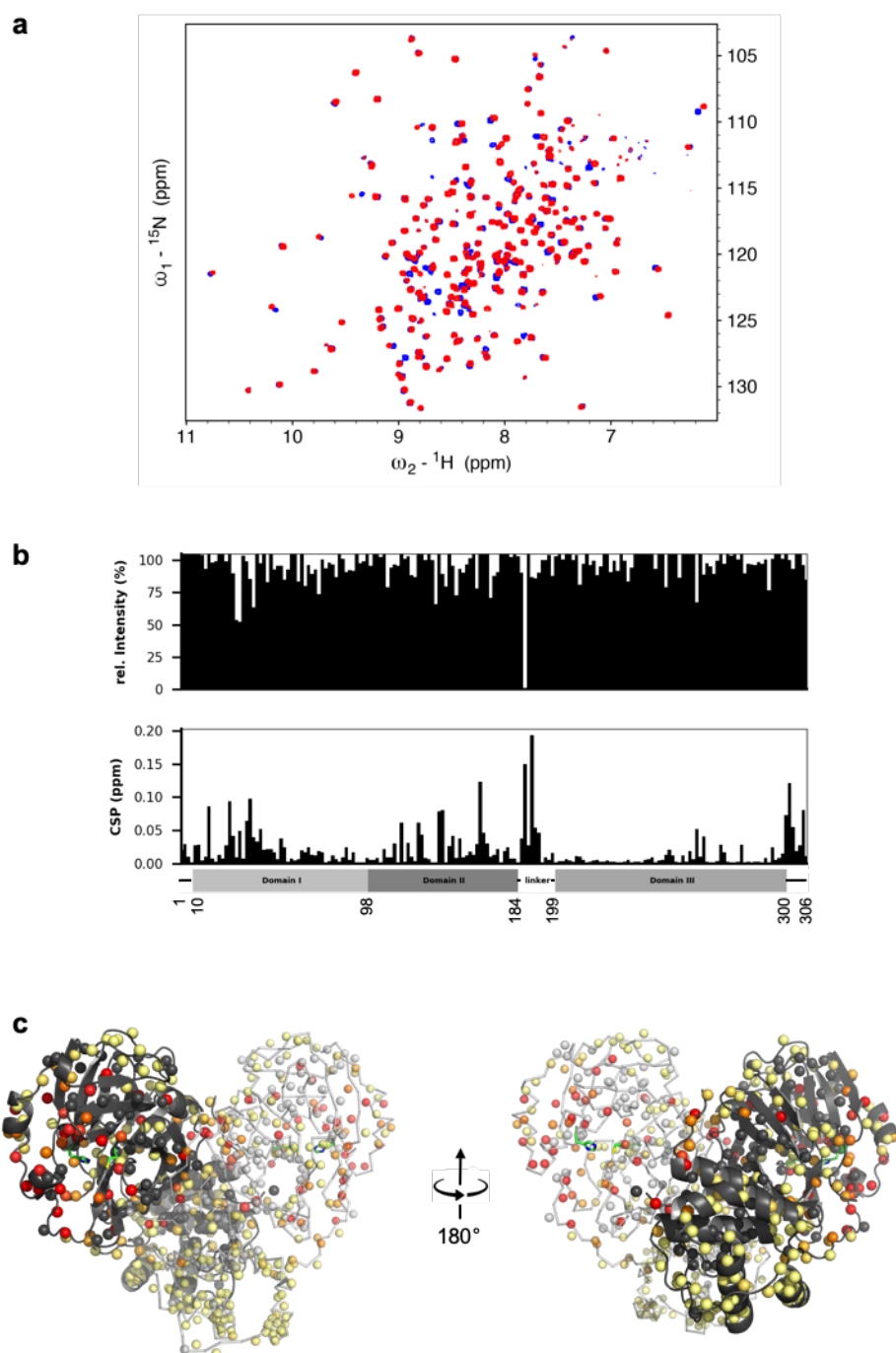
F23



F12

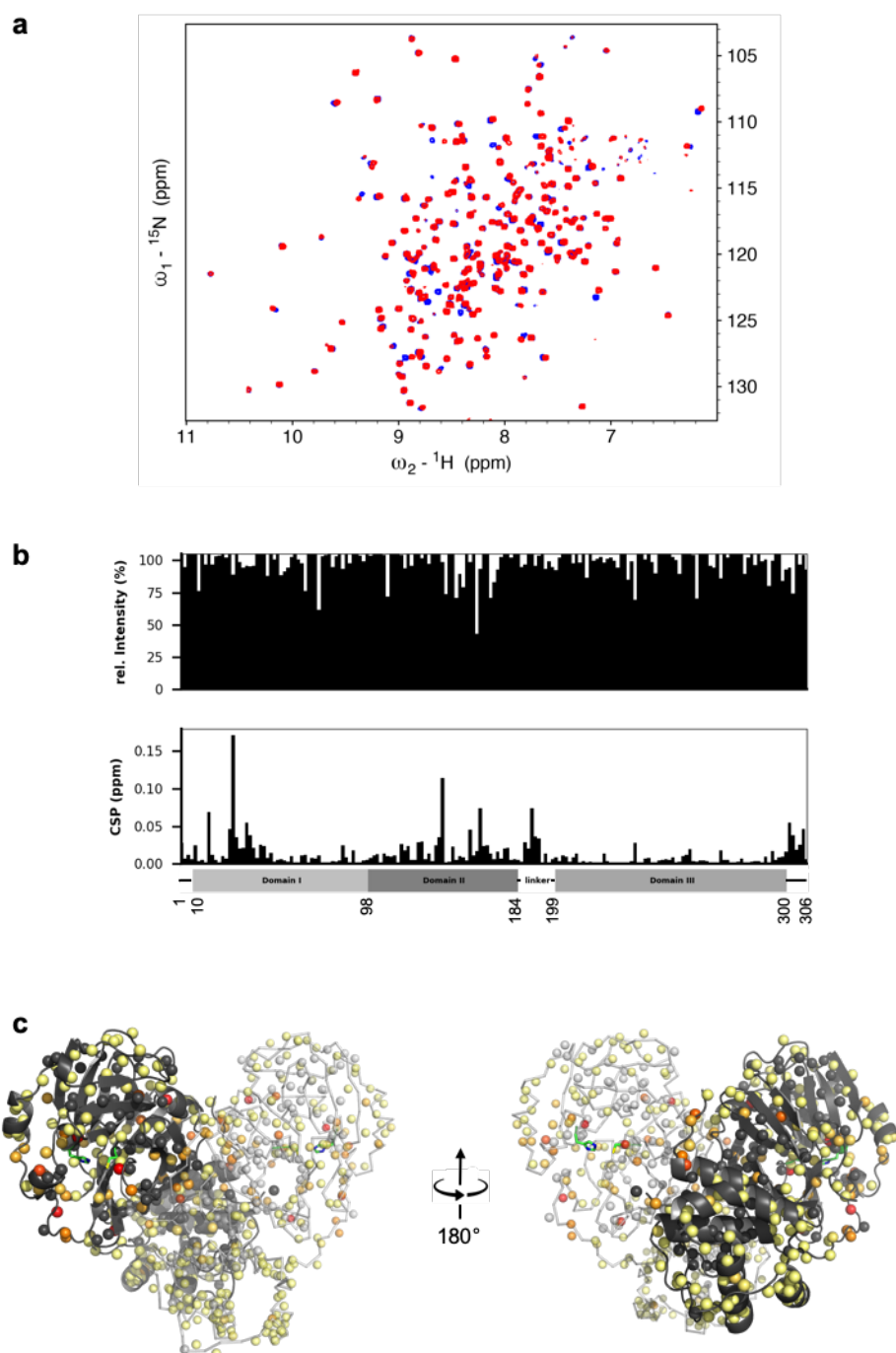


F13

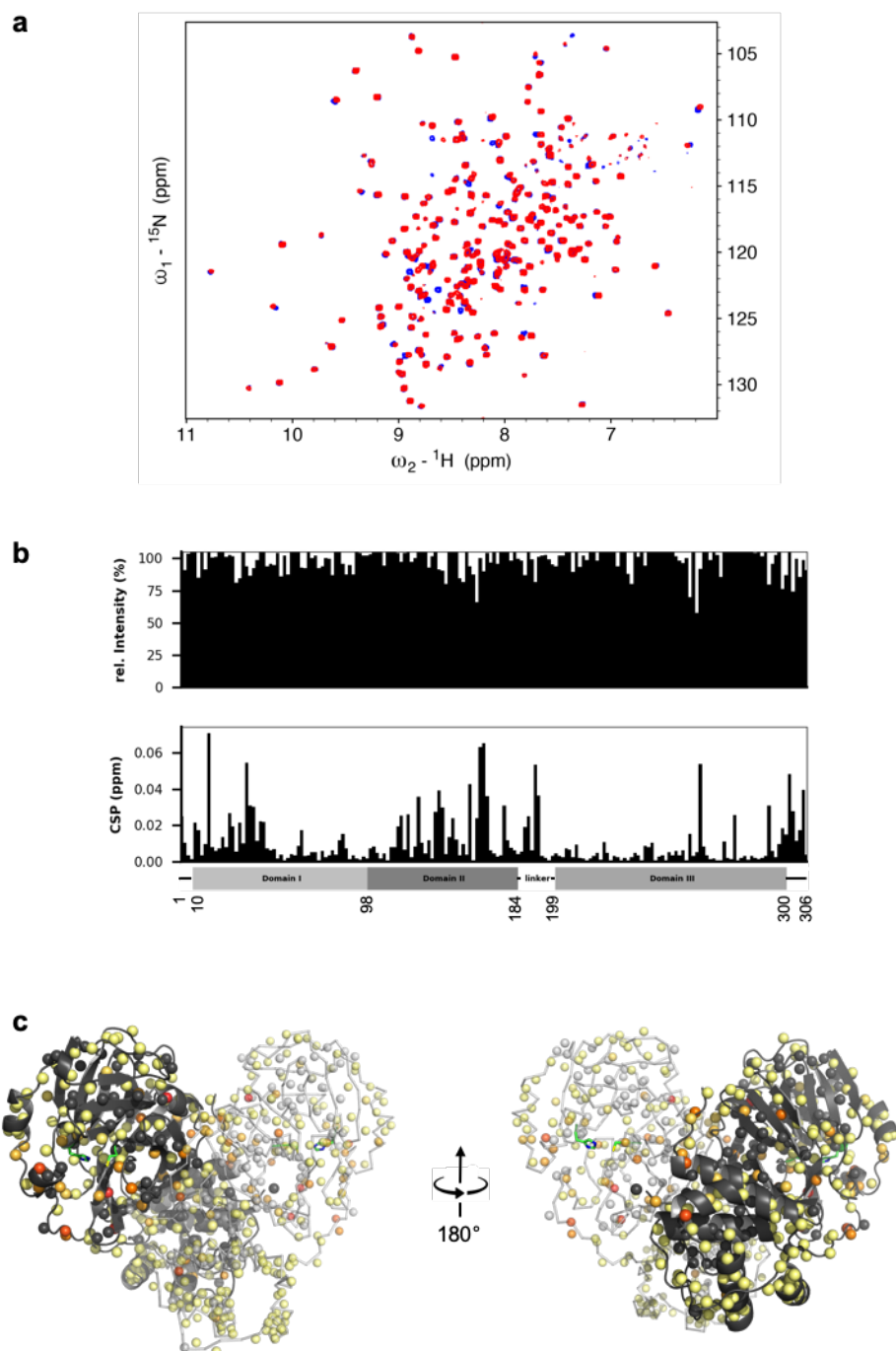


SUPPORTING INFORMATION

F10

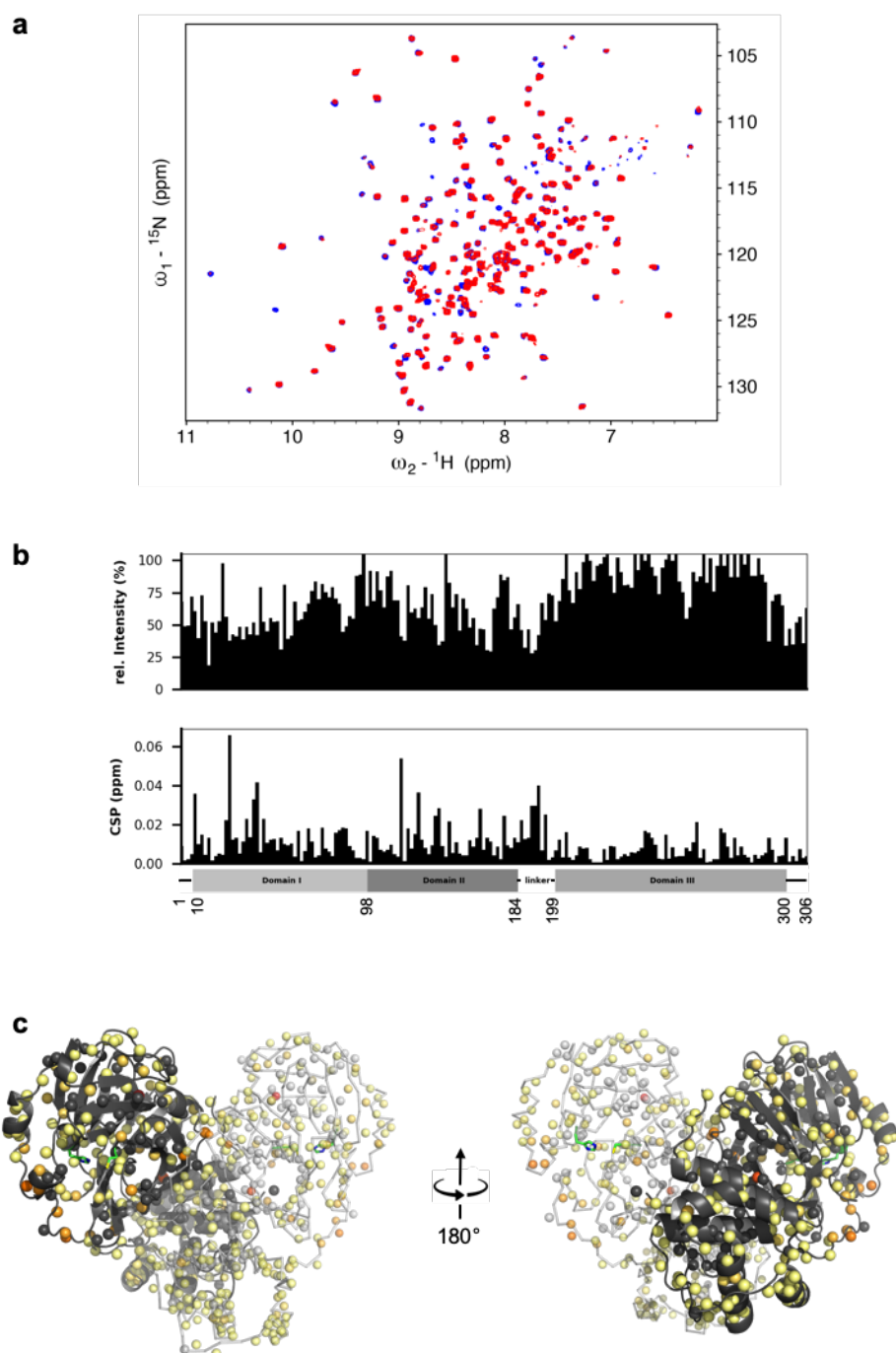


F24



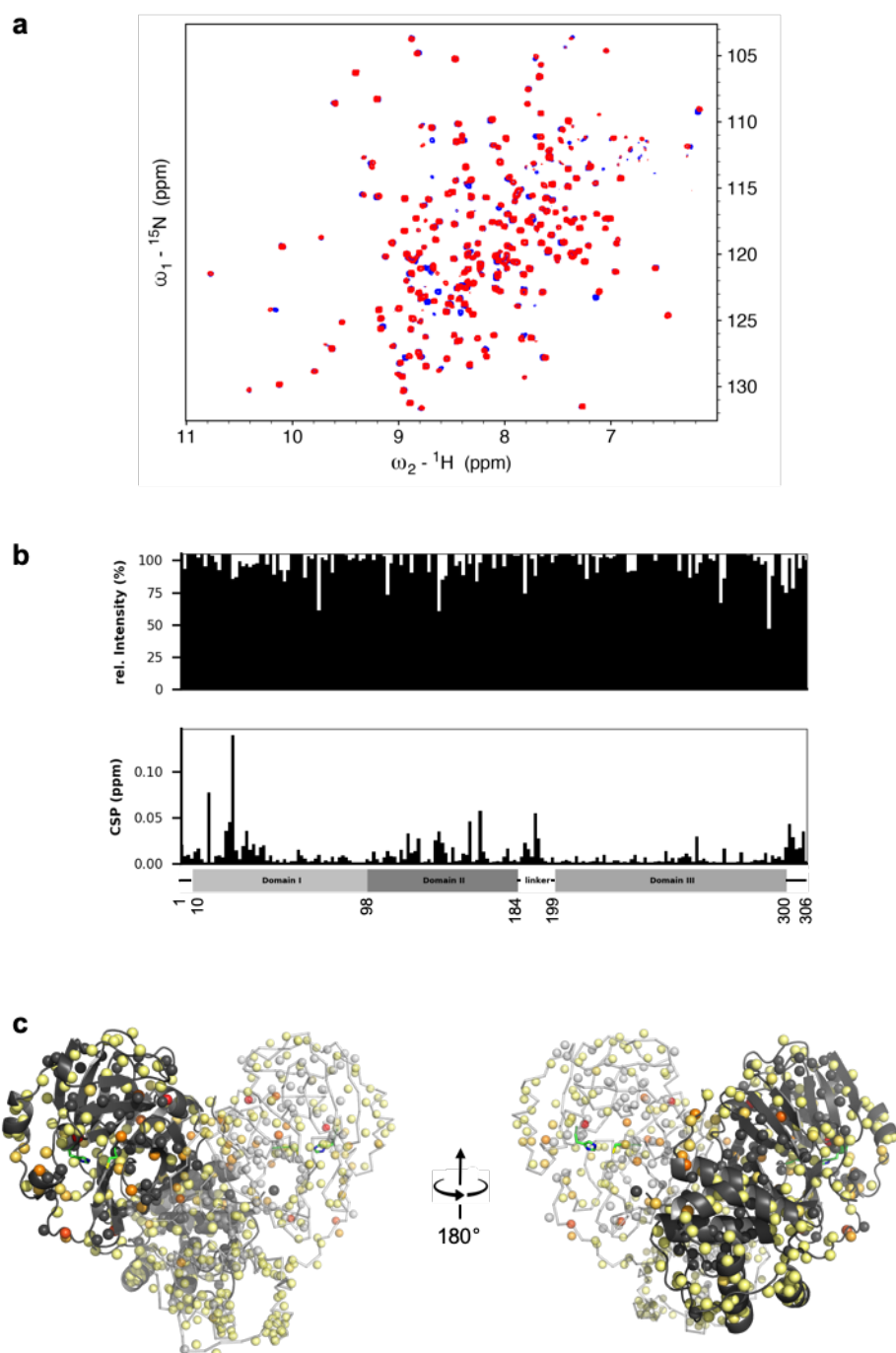
SUPPORTING INFORMATION

F03

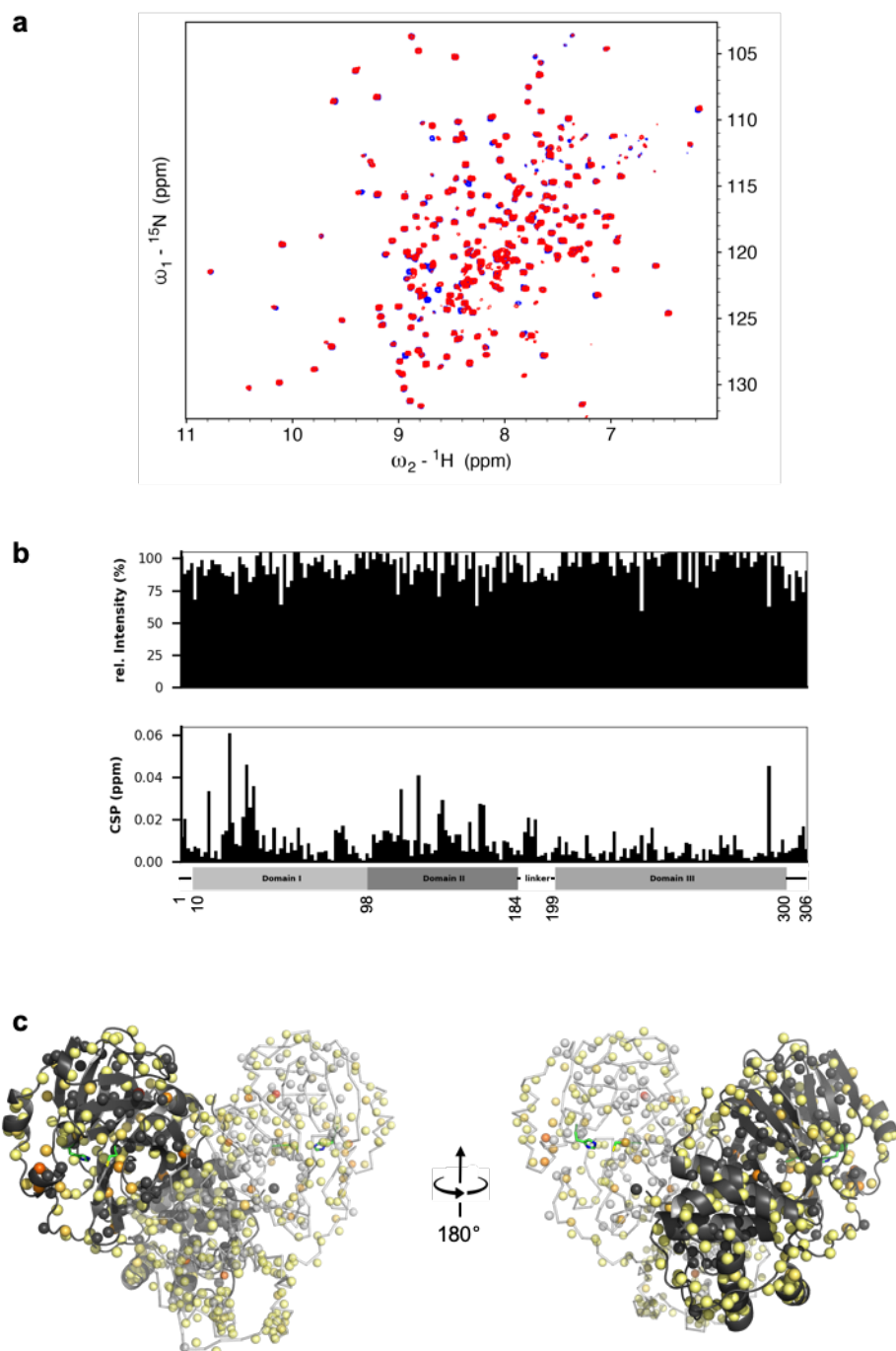


SUPPORTING INFORMATION

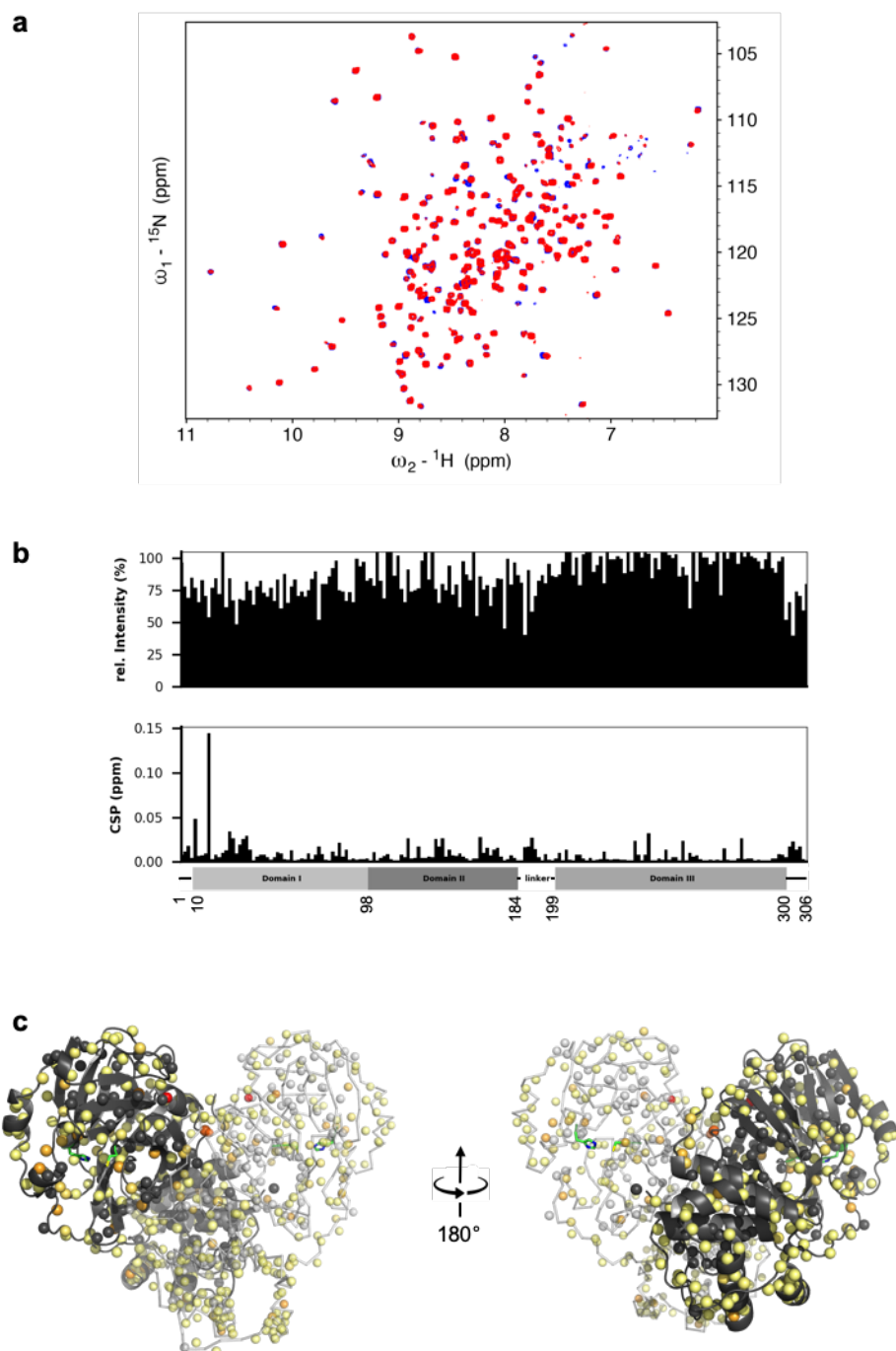
F11



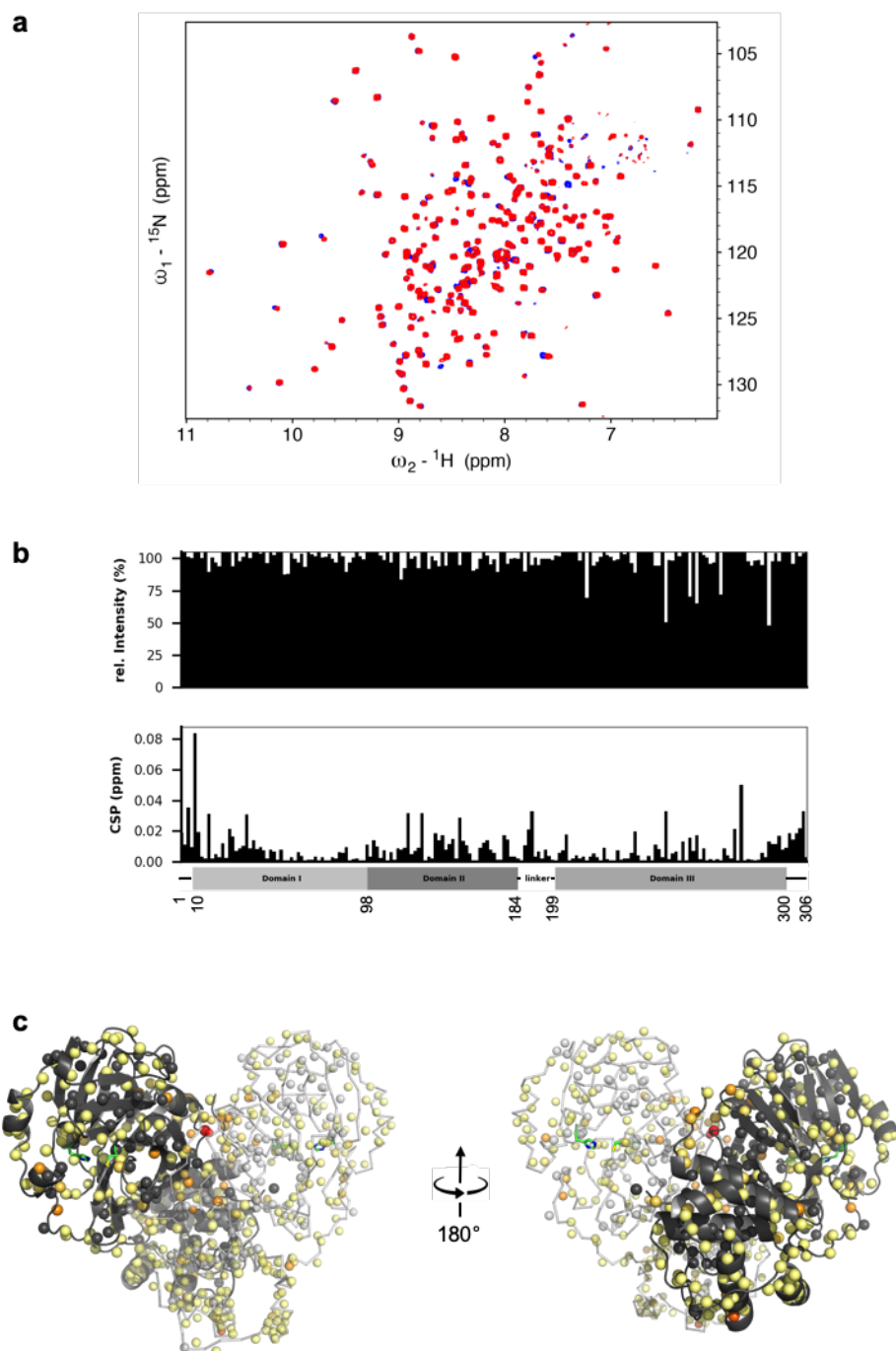
F05



F17

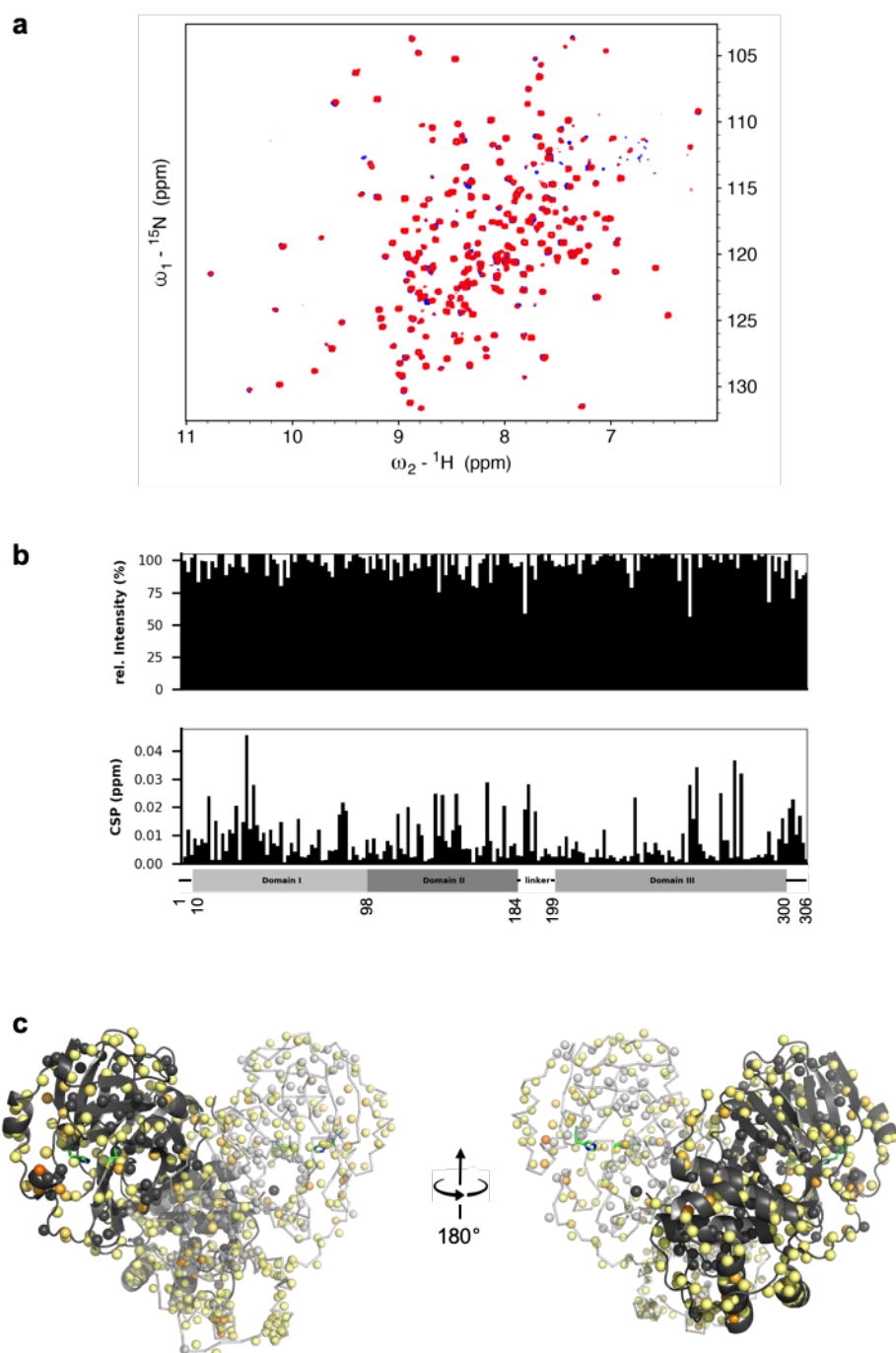


F28

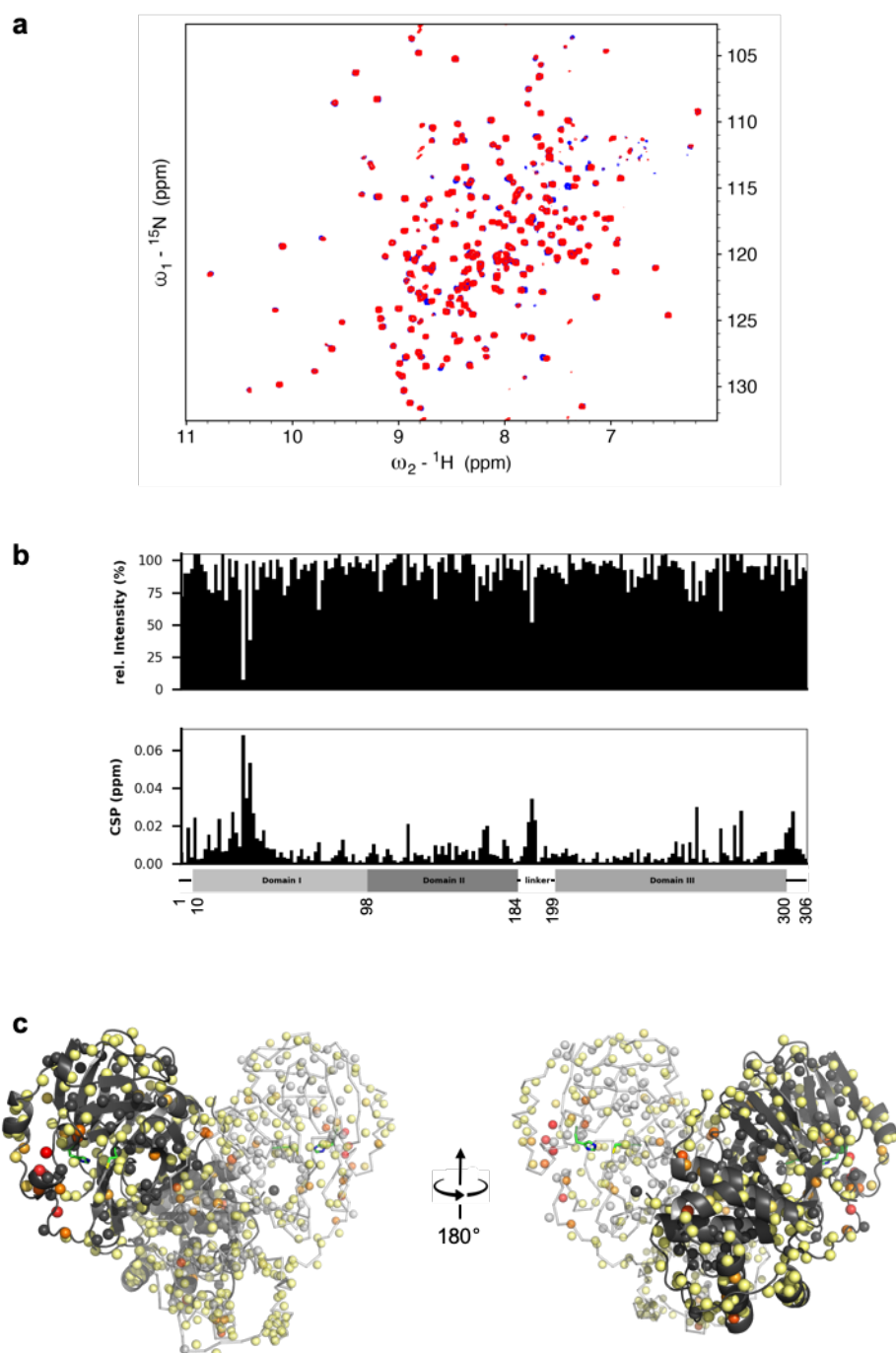


SUPPORTING INFORMATION

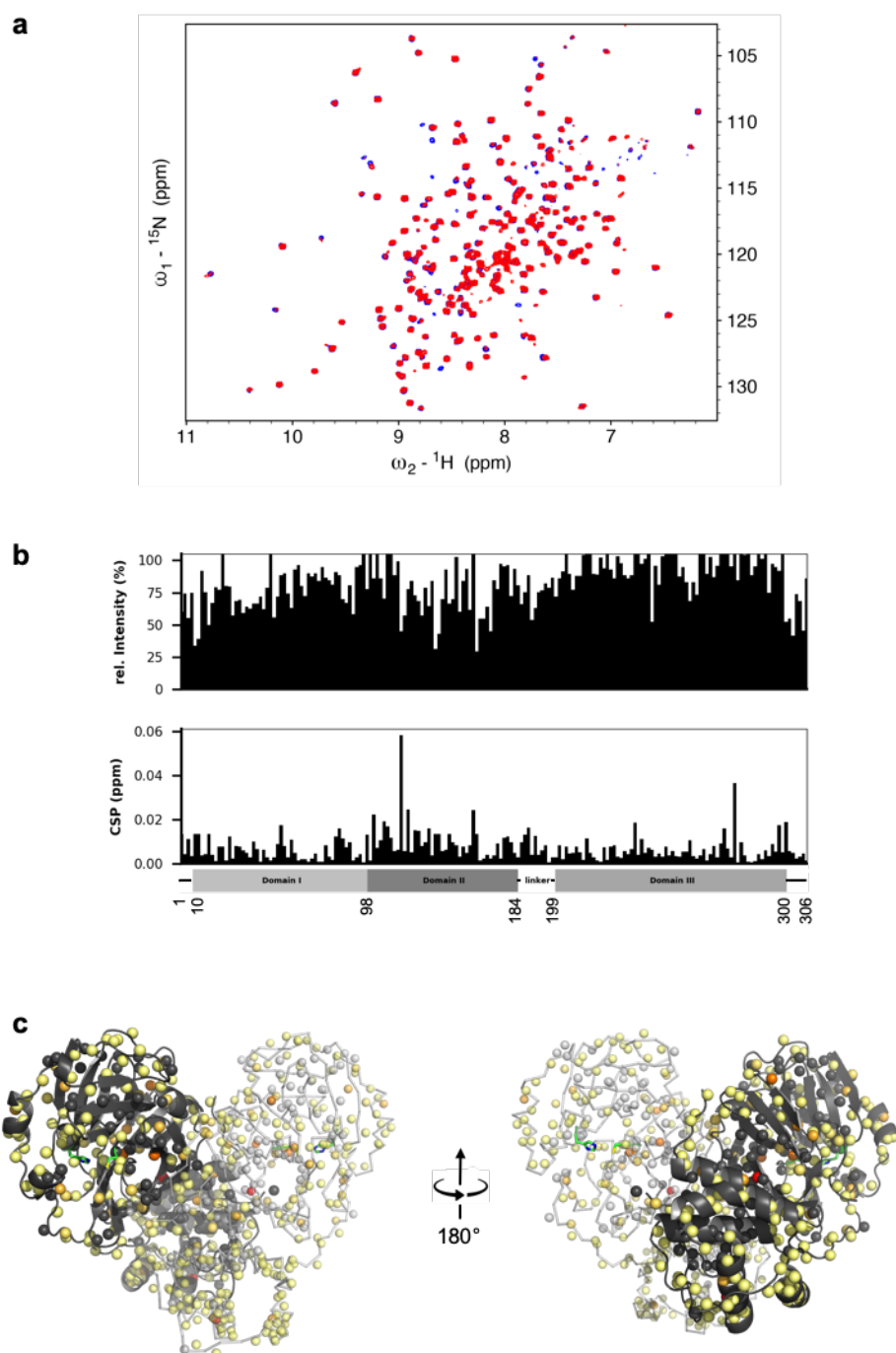
F21



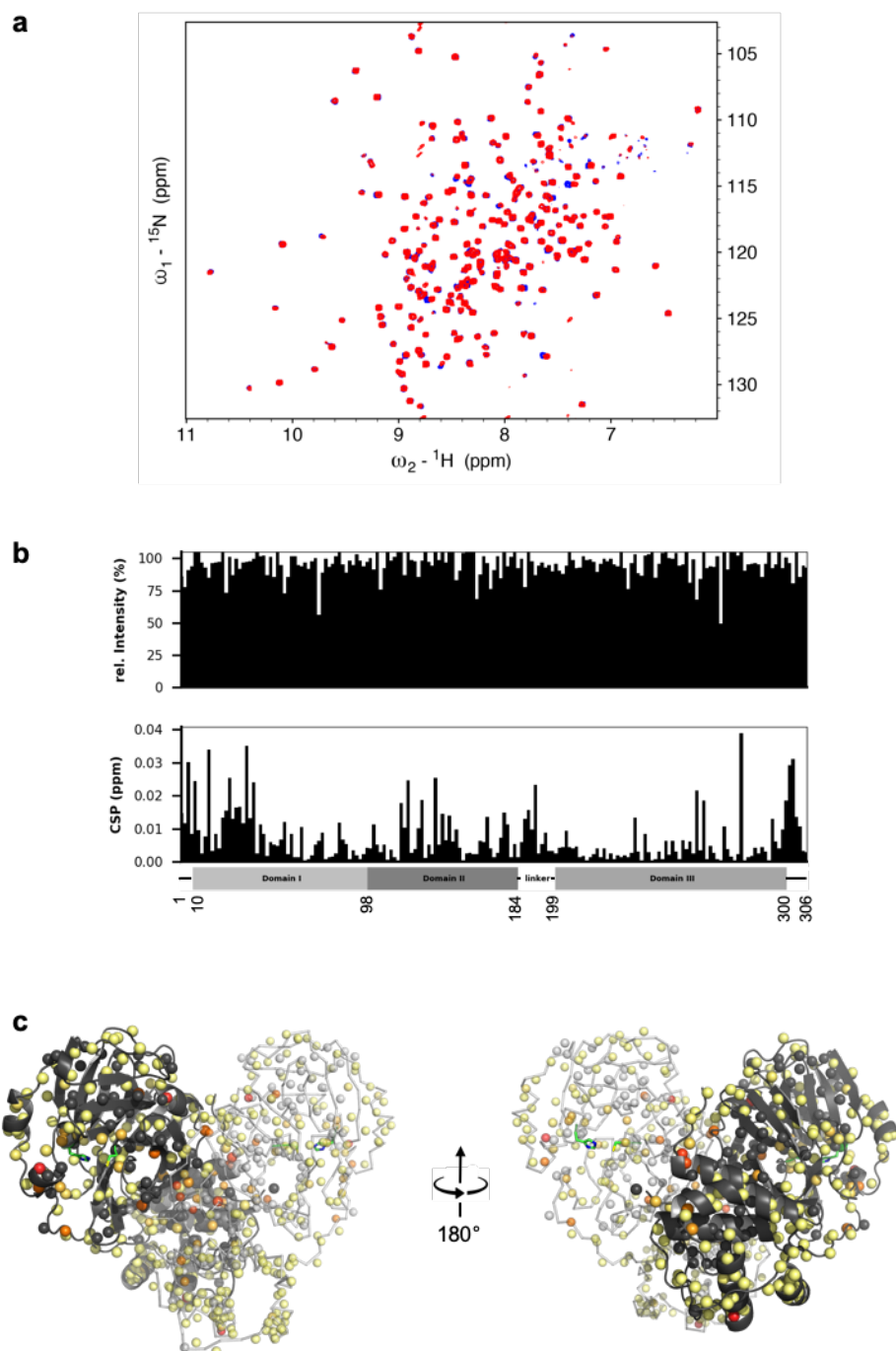
F06



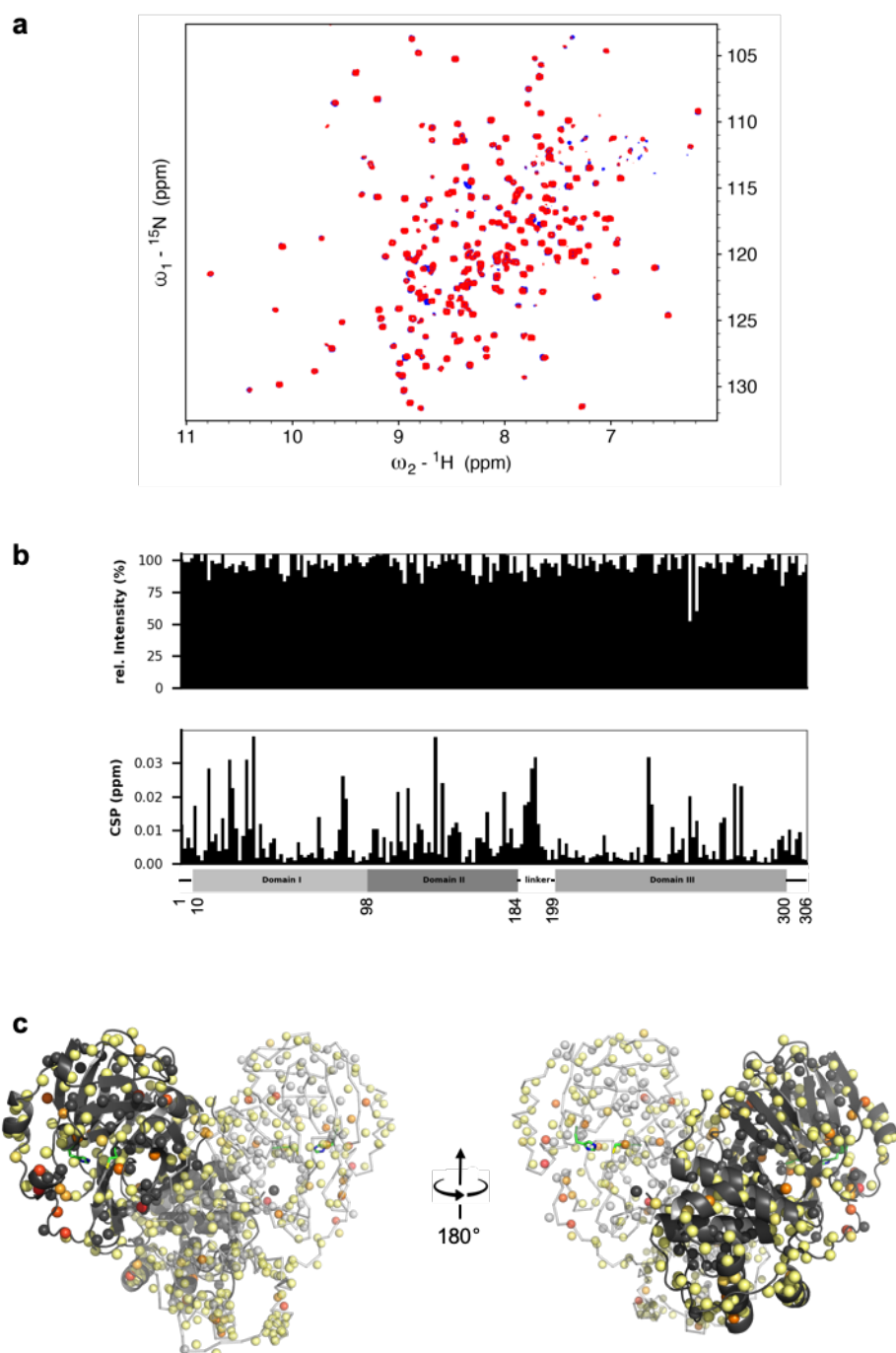
F02



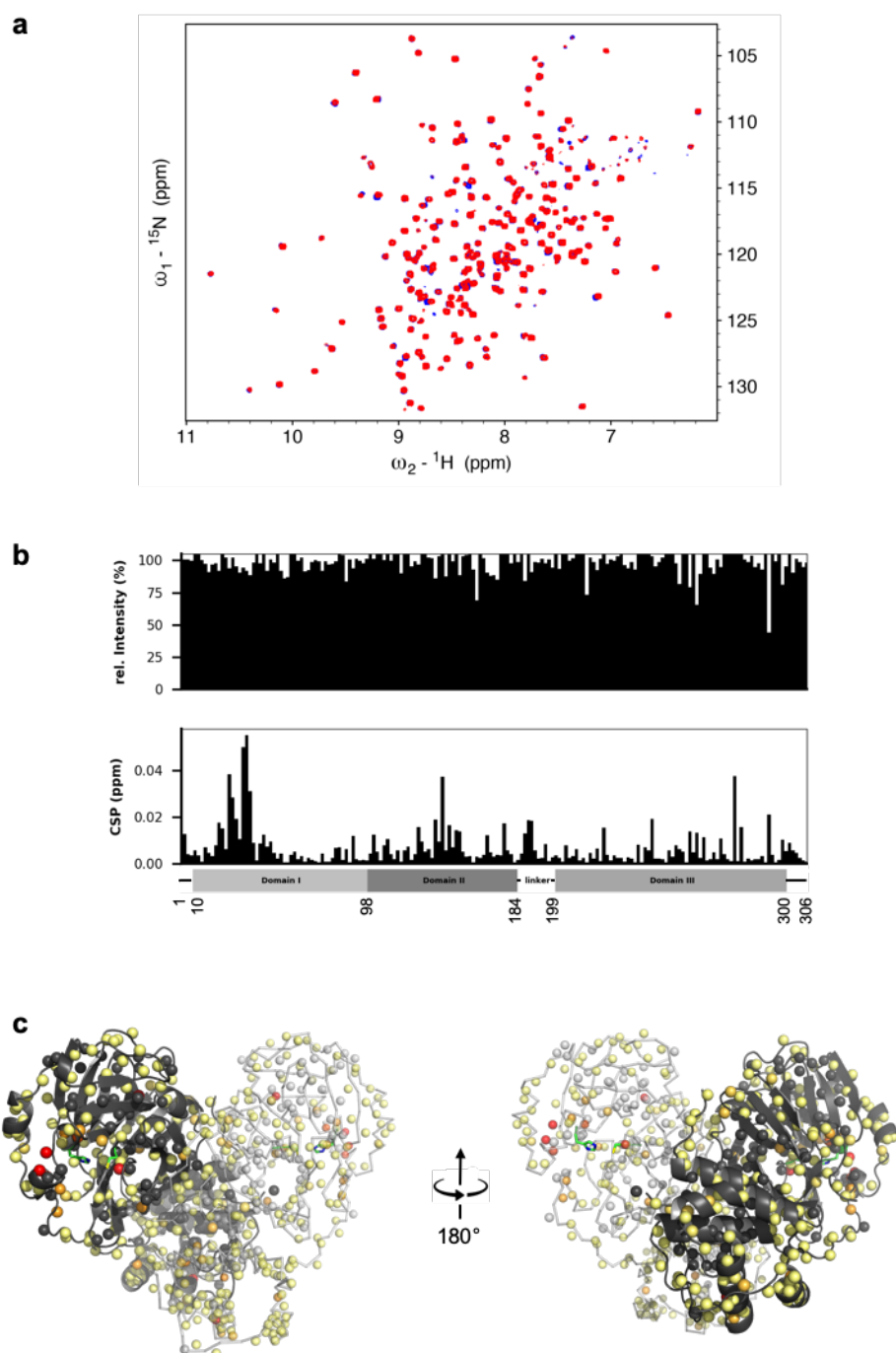
F08



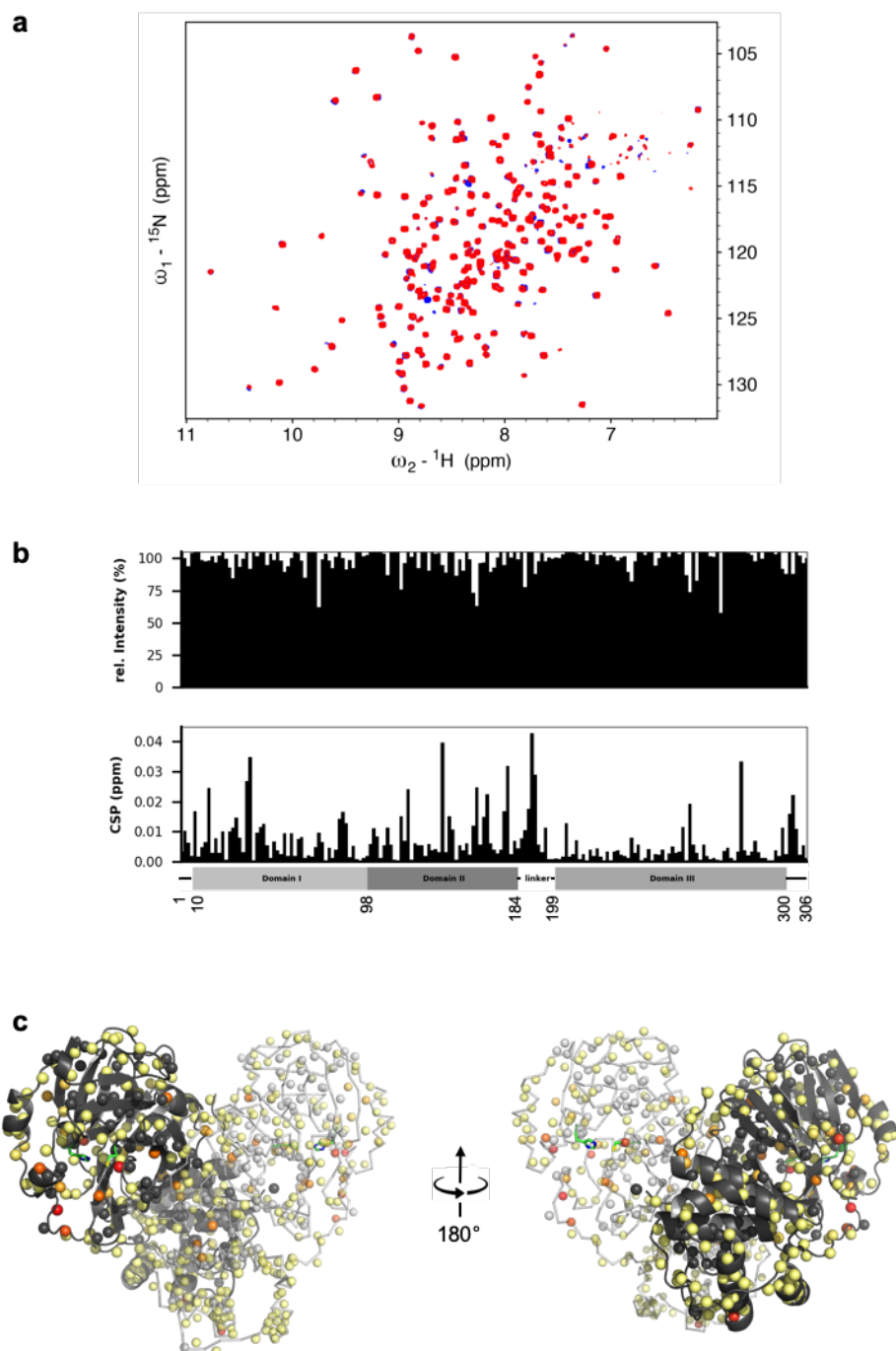
F29



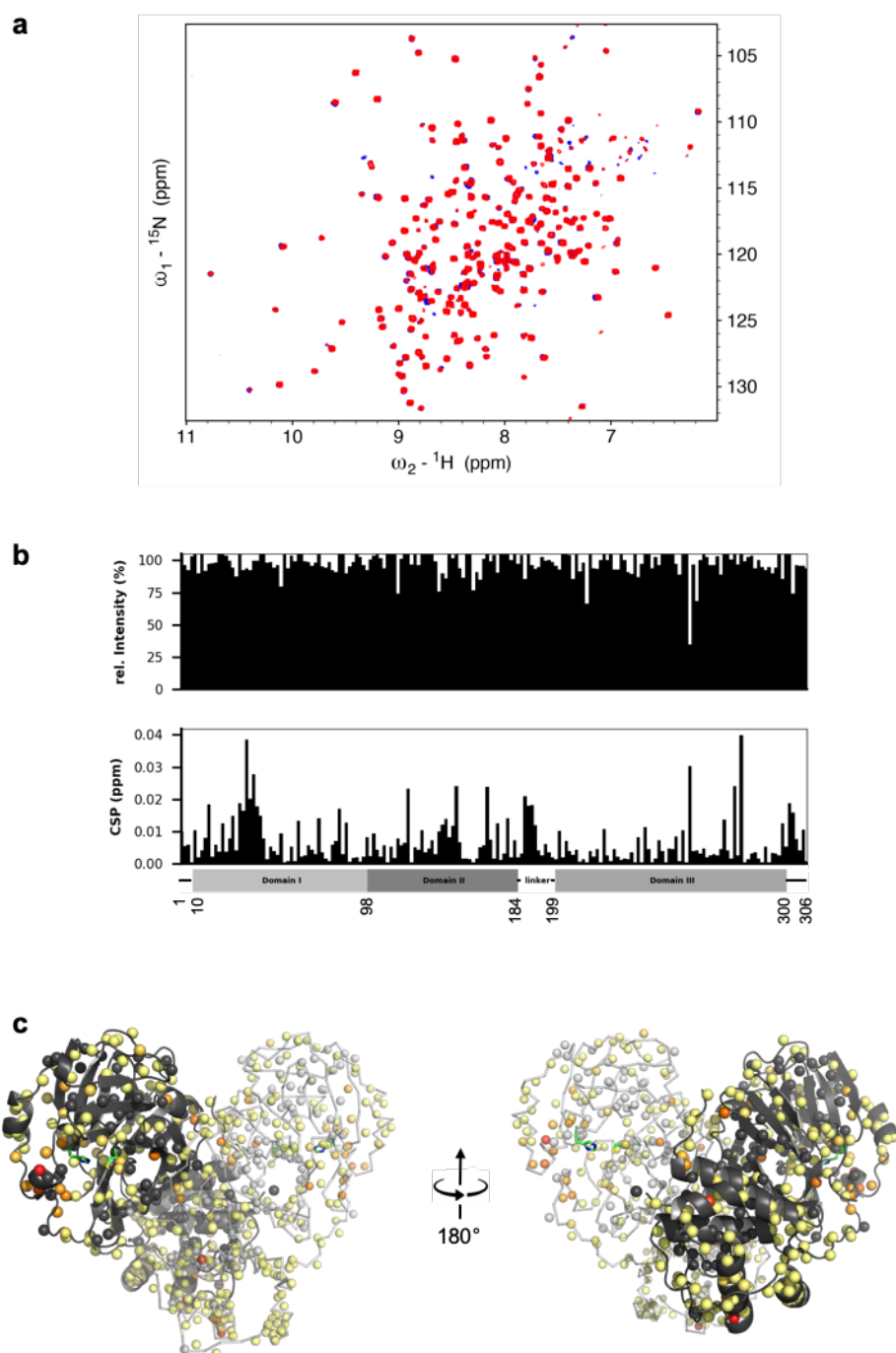
F33



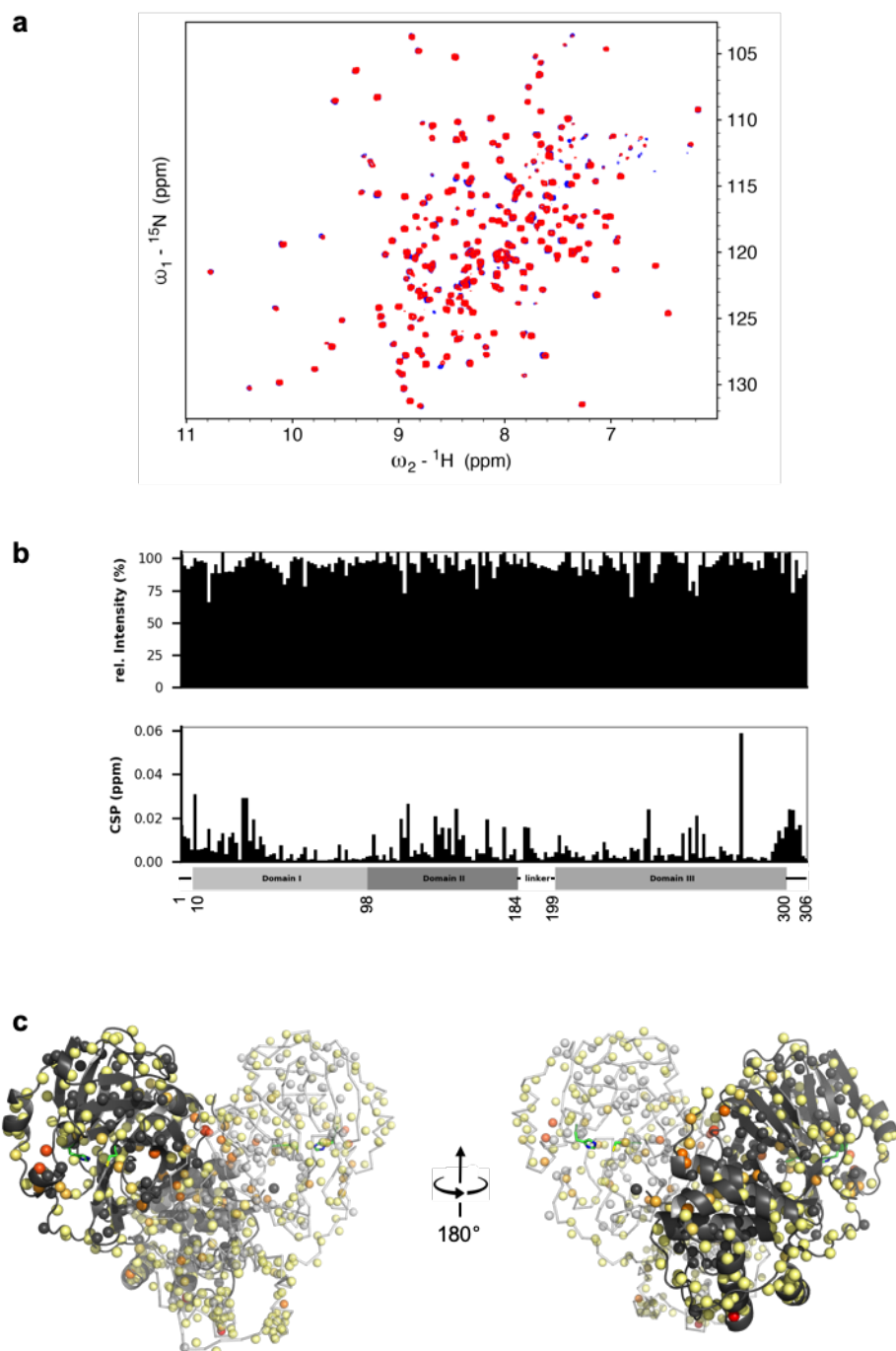
F07



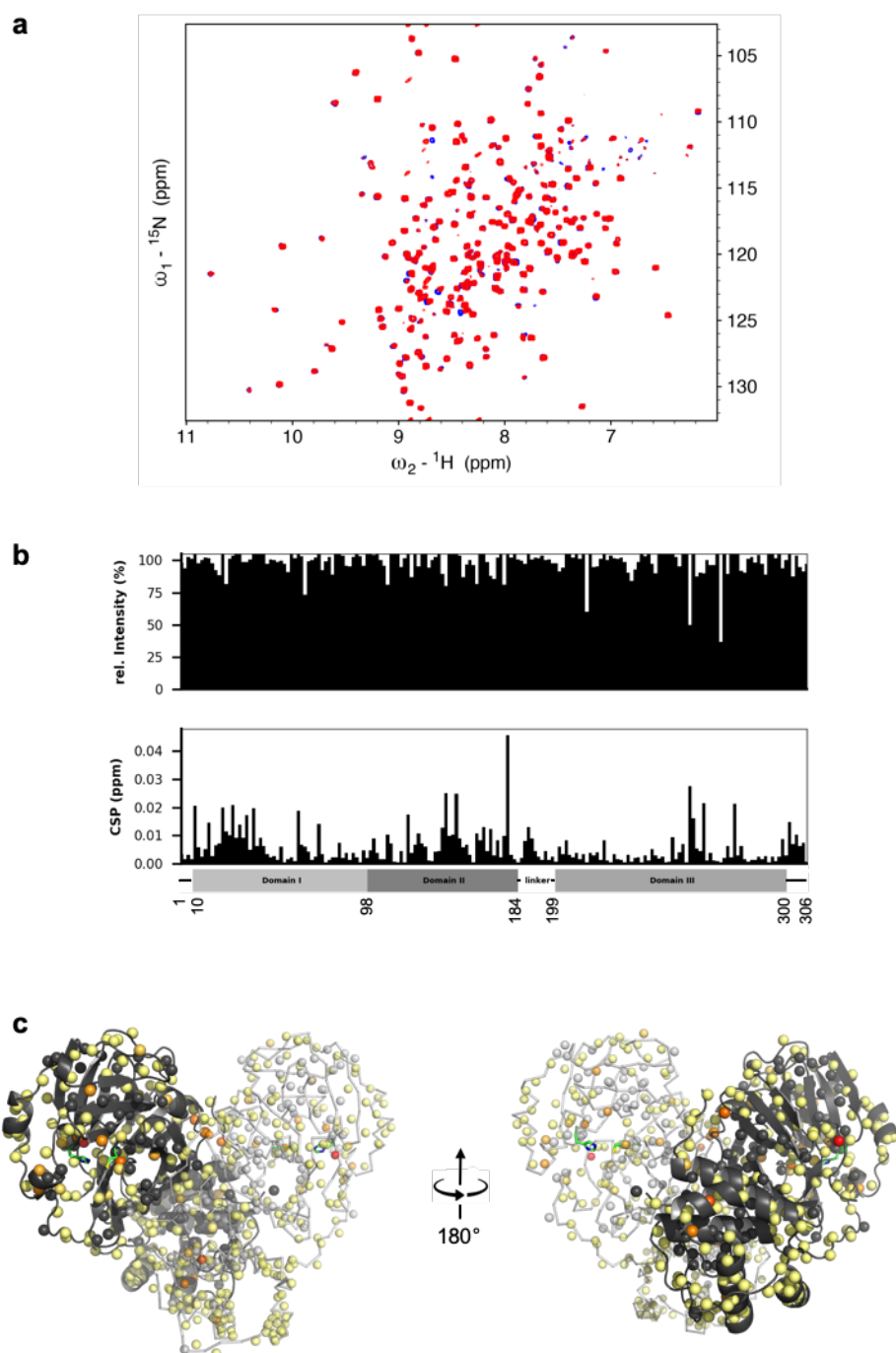
F19



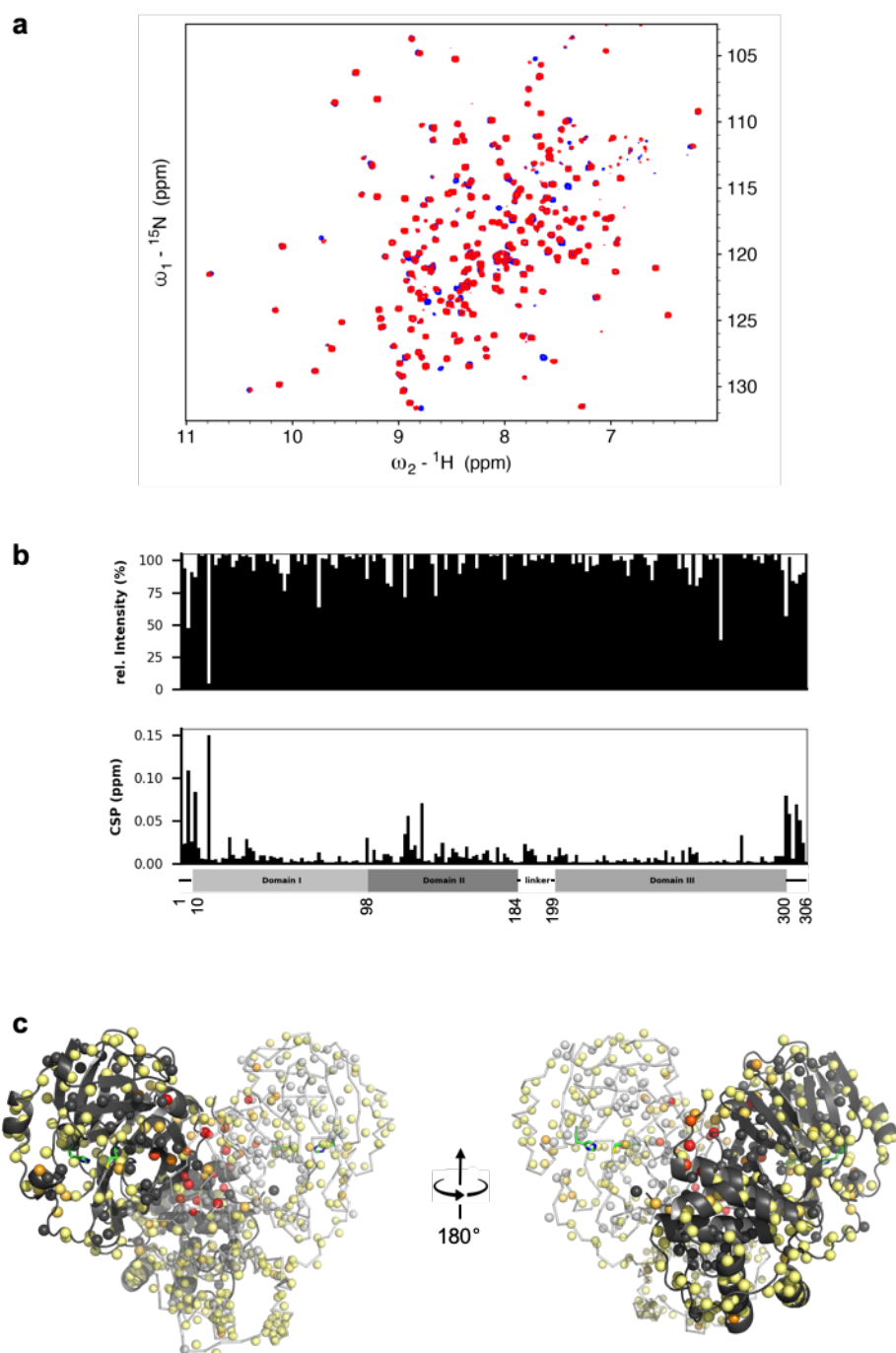
F36



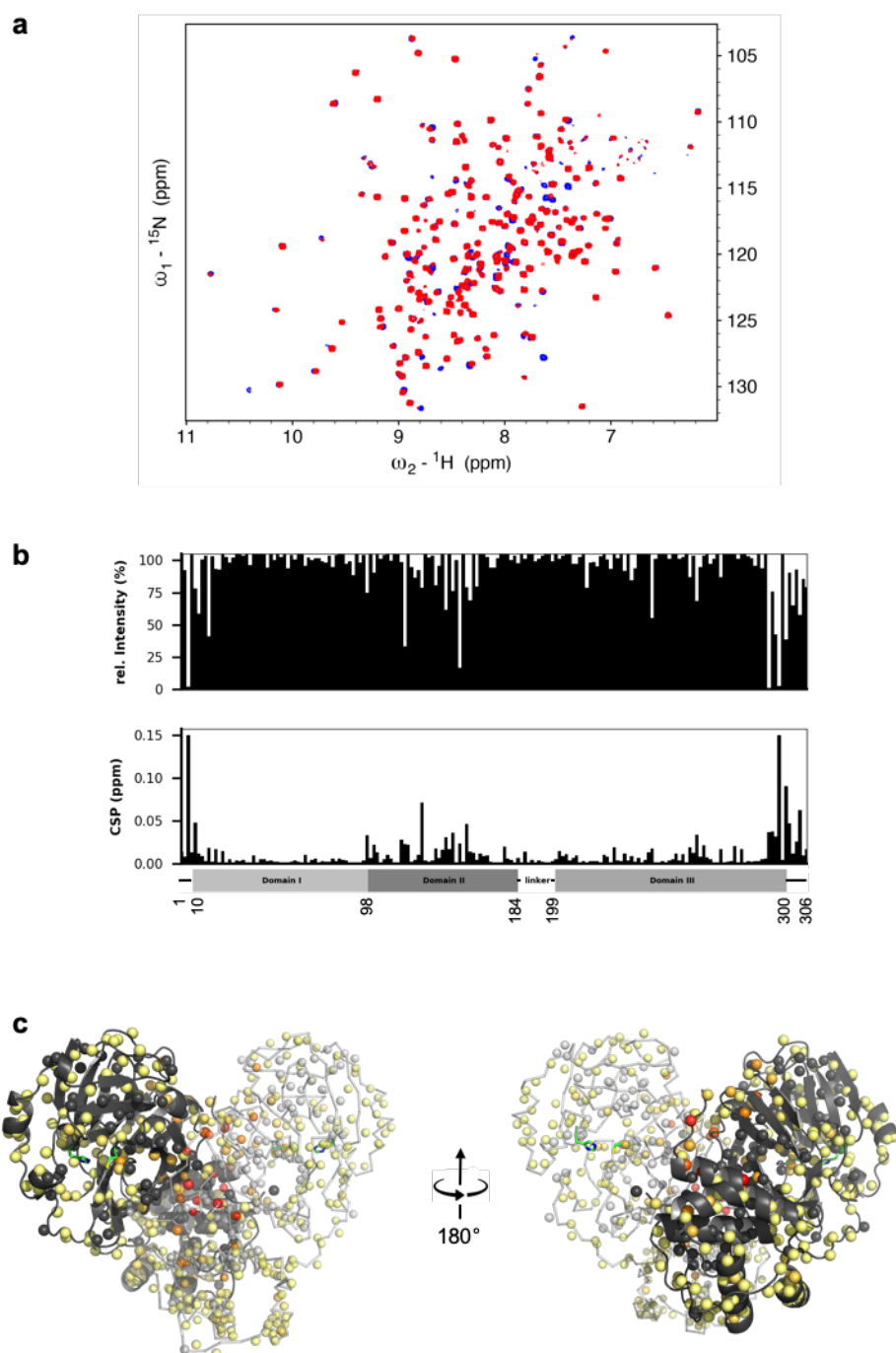
F26



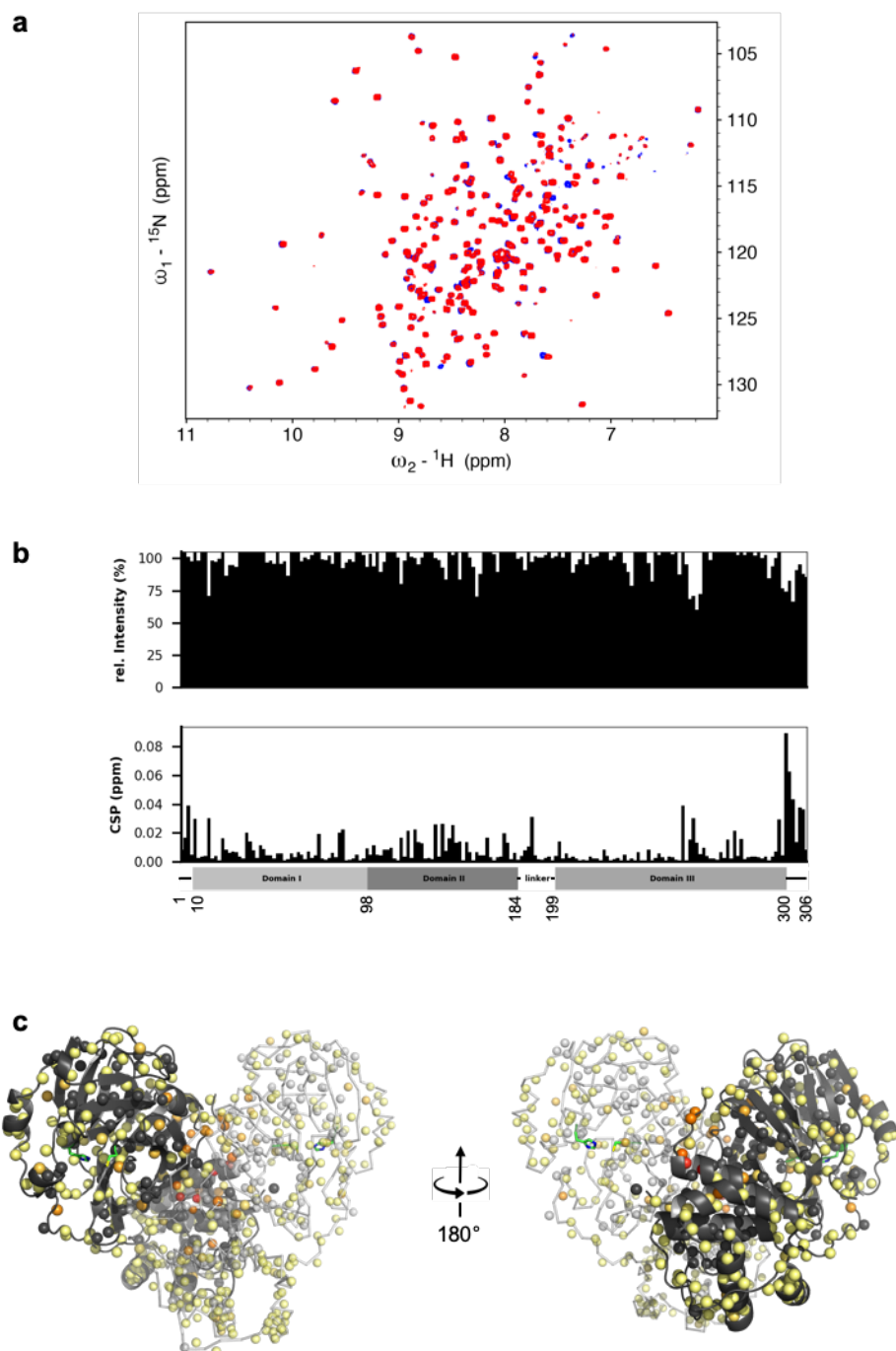
F15



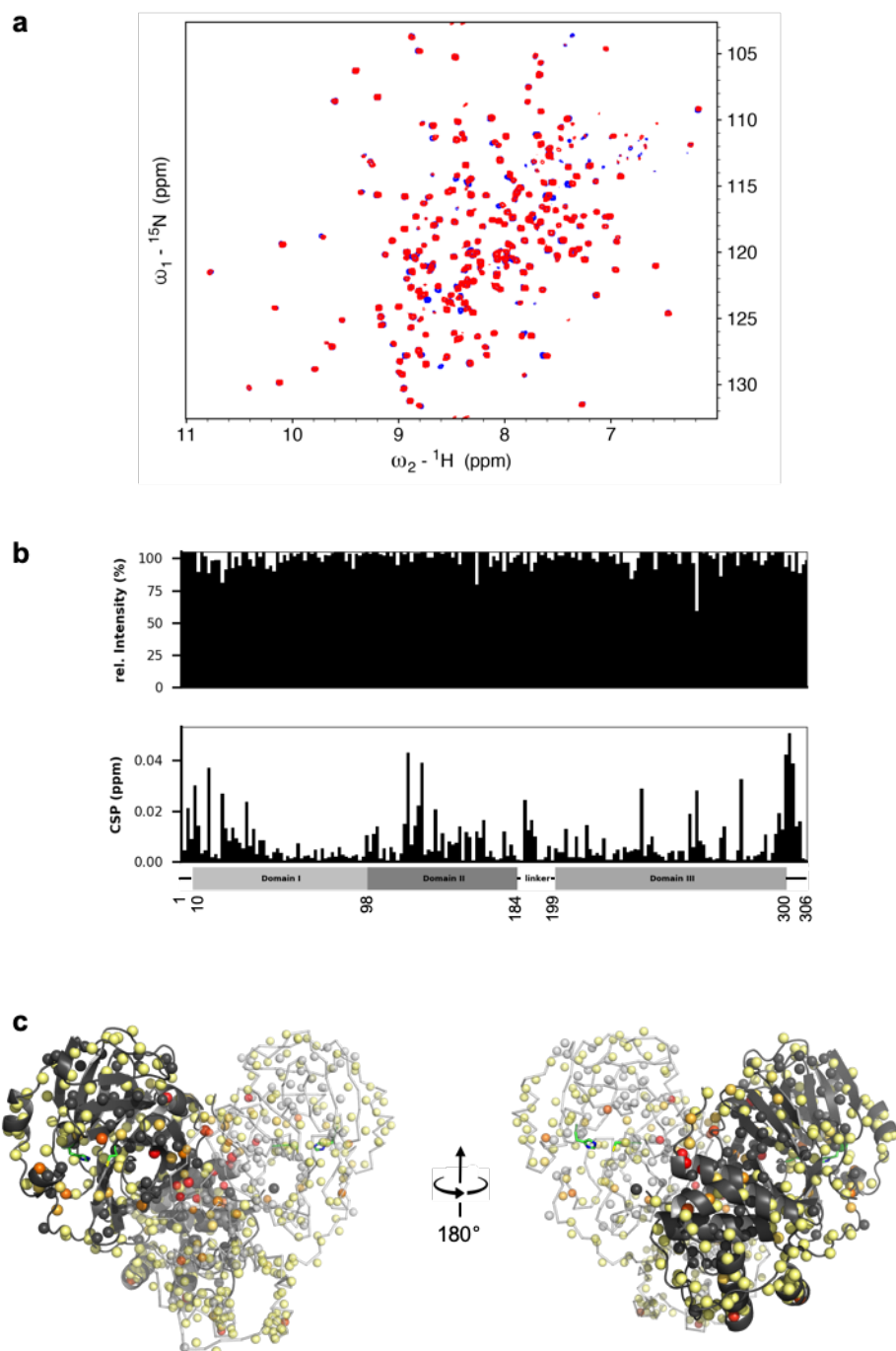
F31



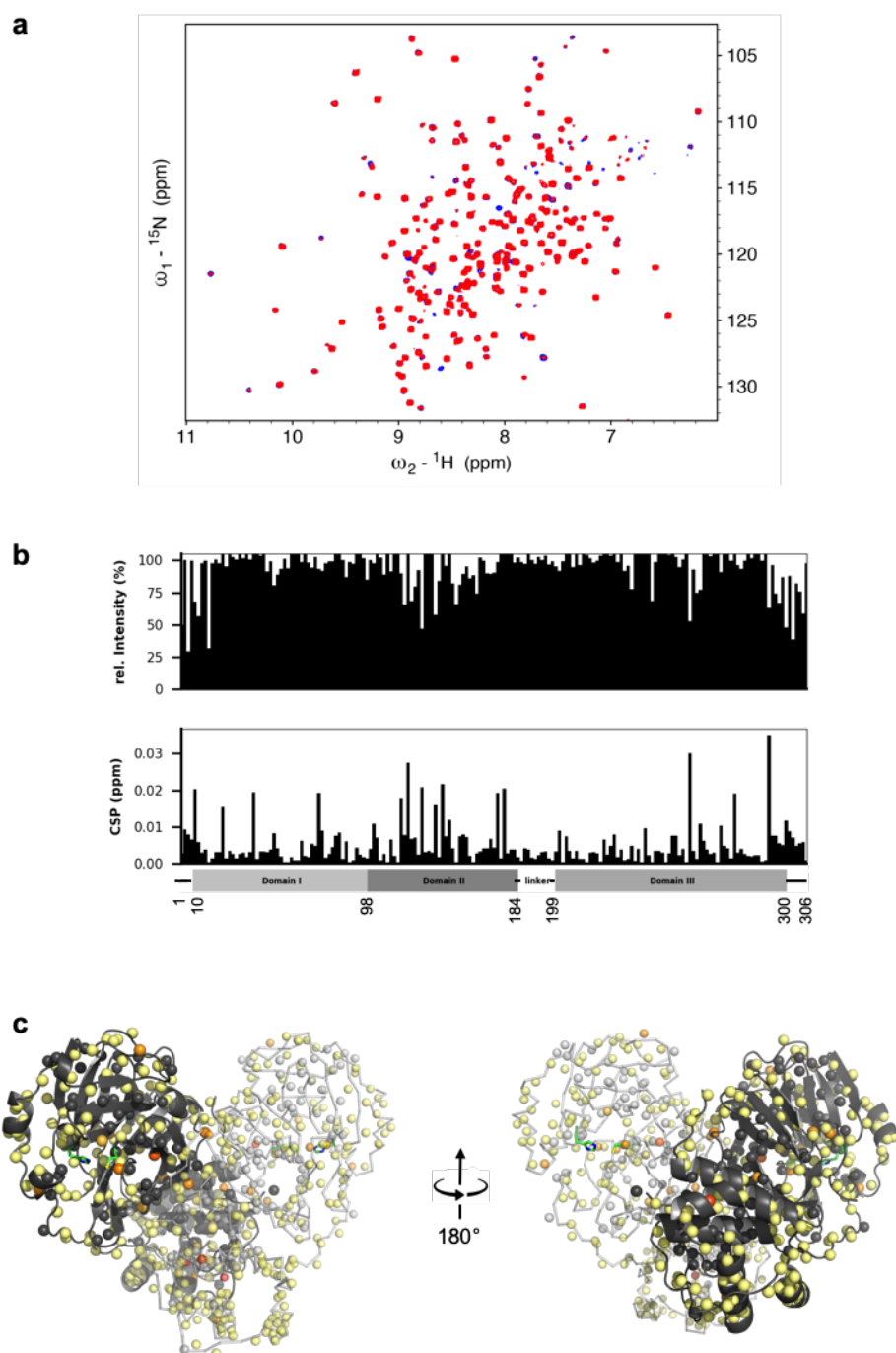
F20



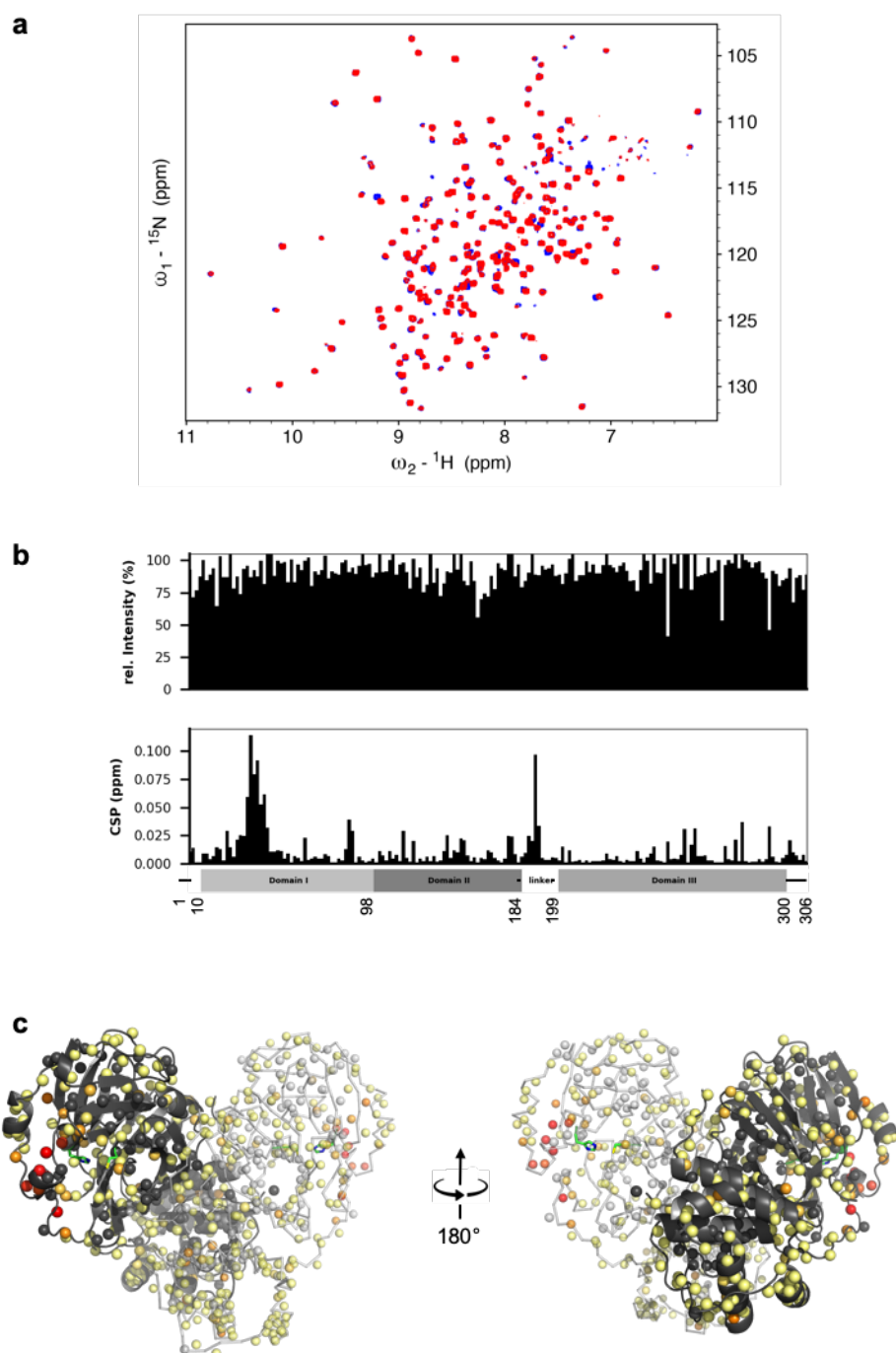
F38



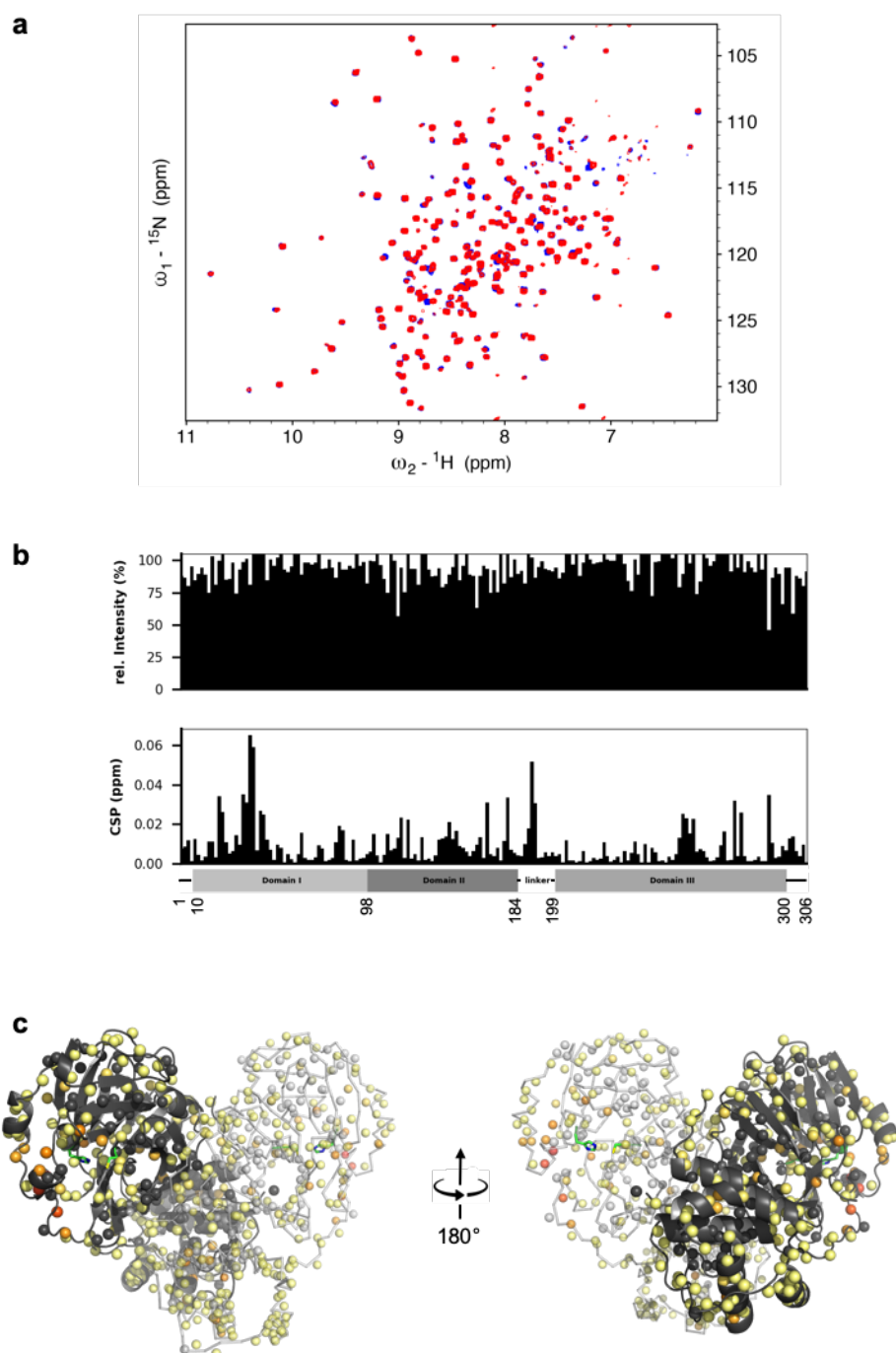
F25



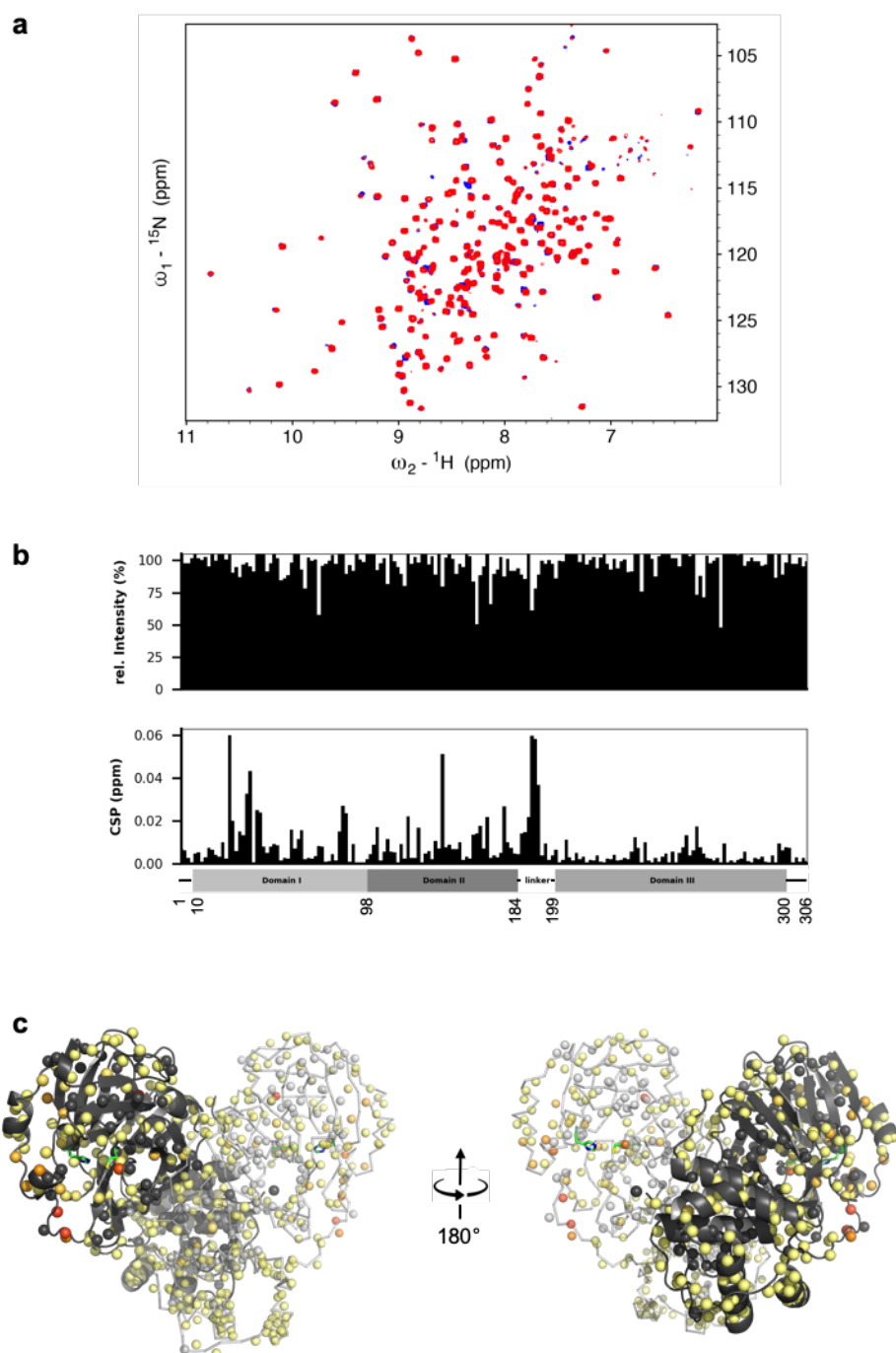
F30



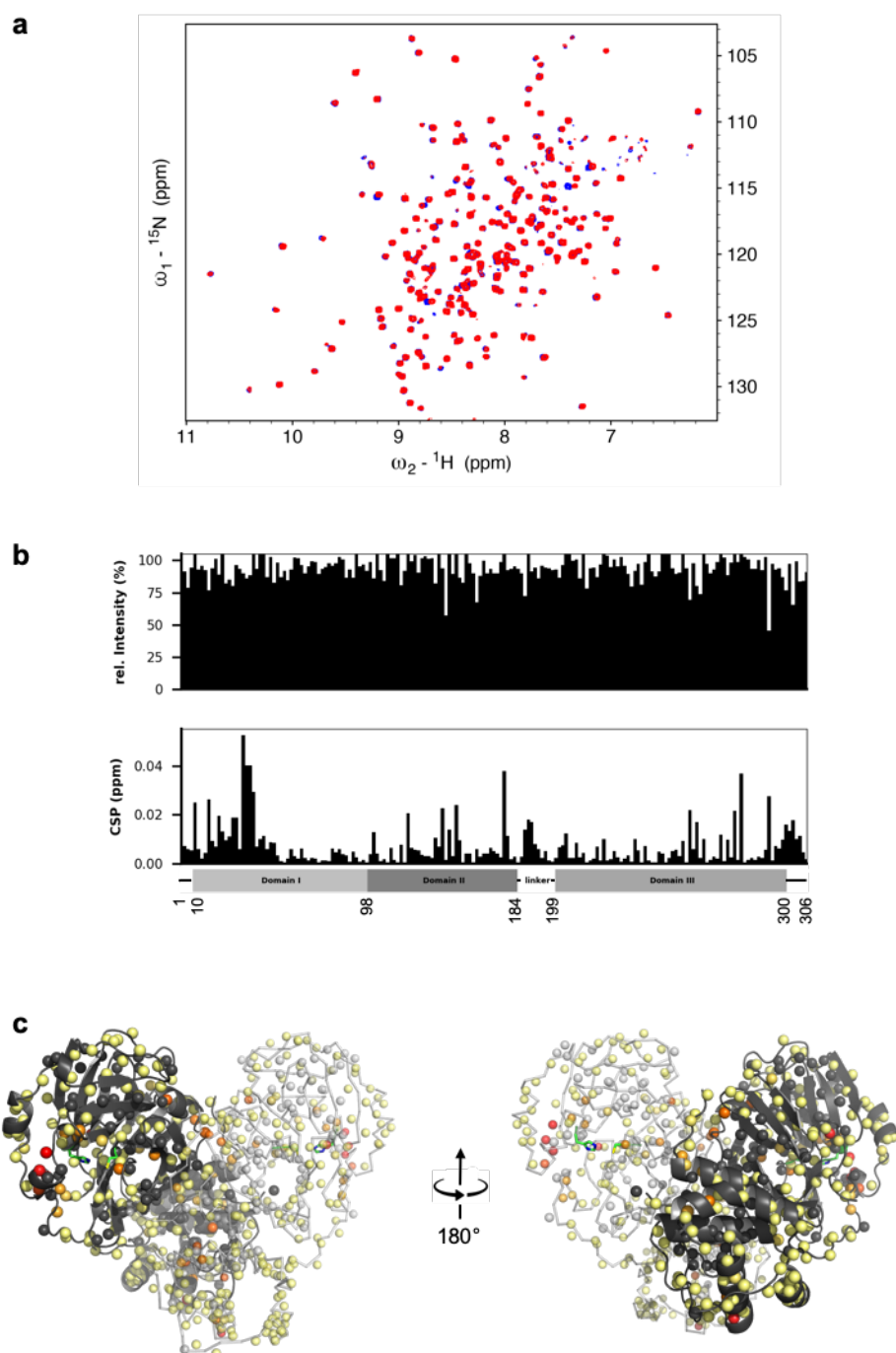
F22



F09

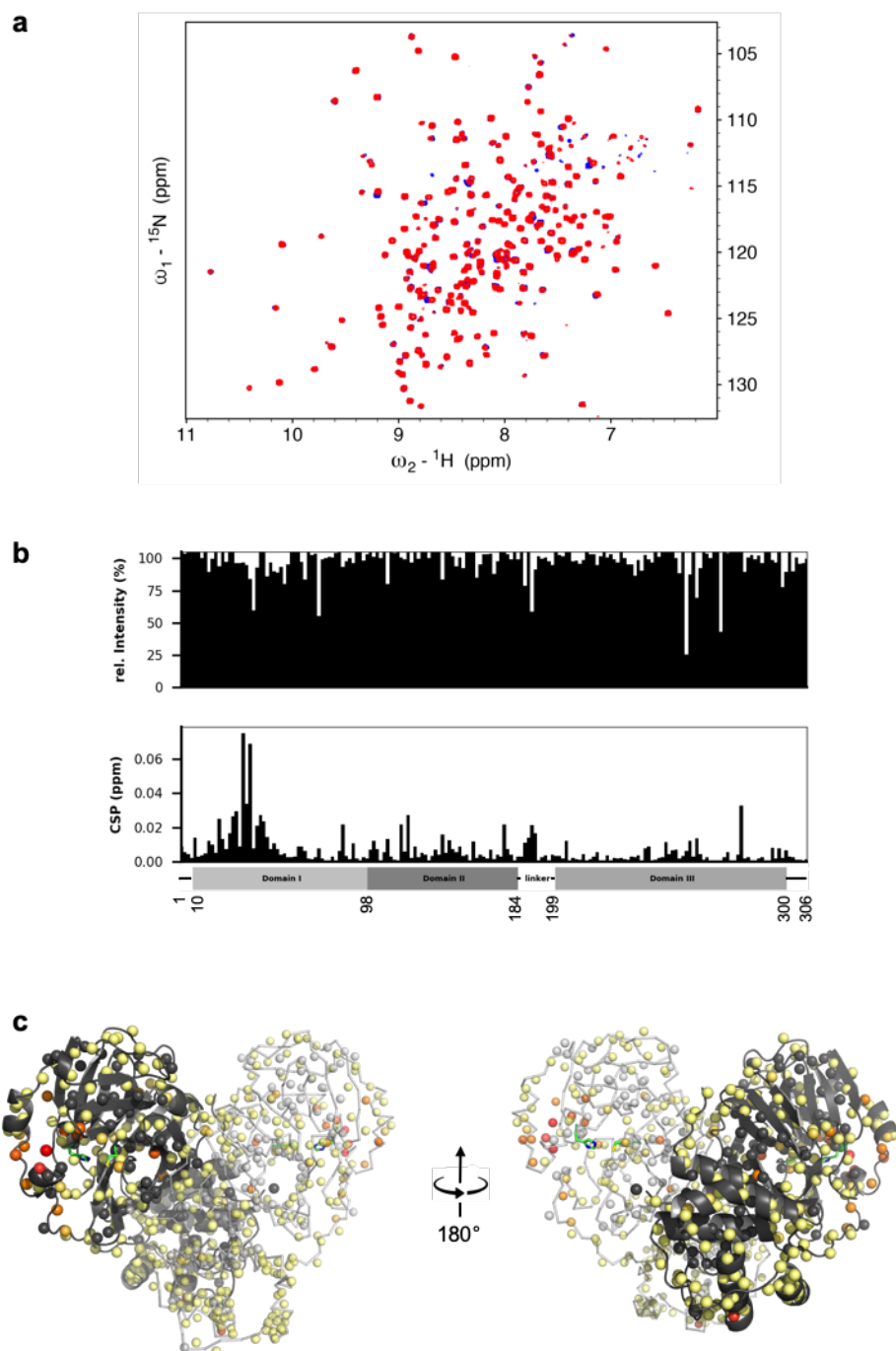


F36

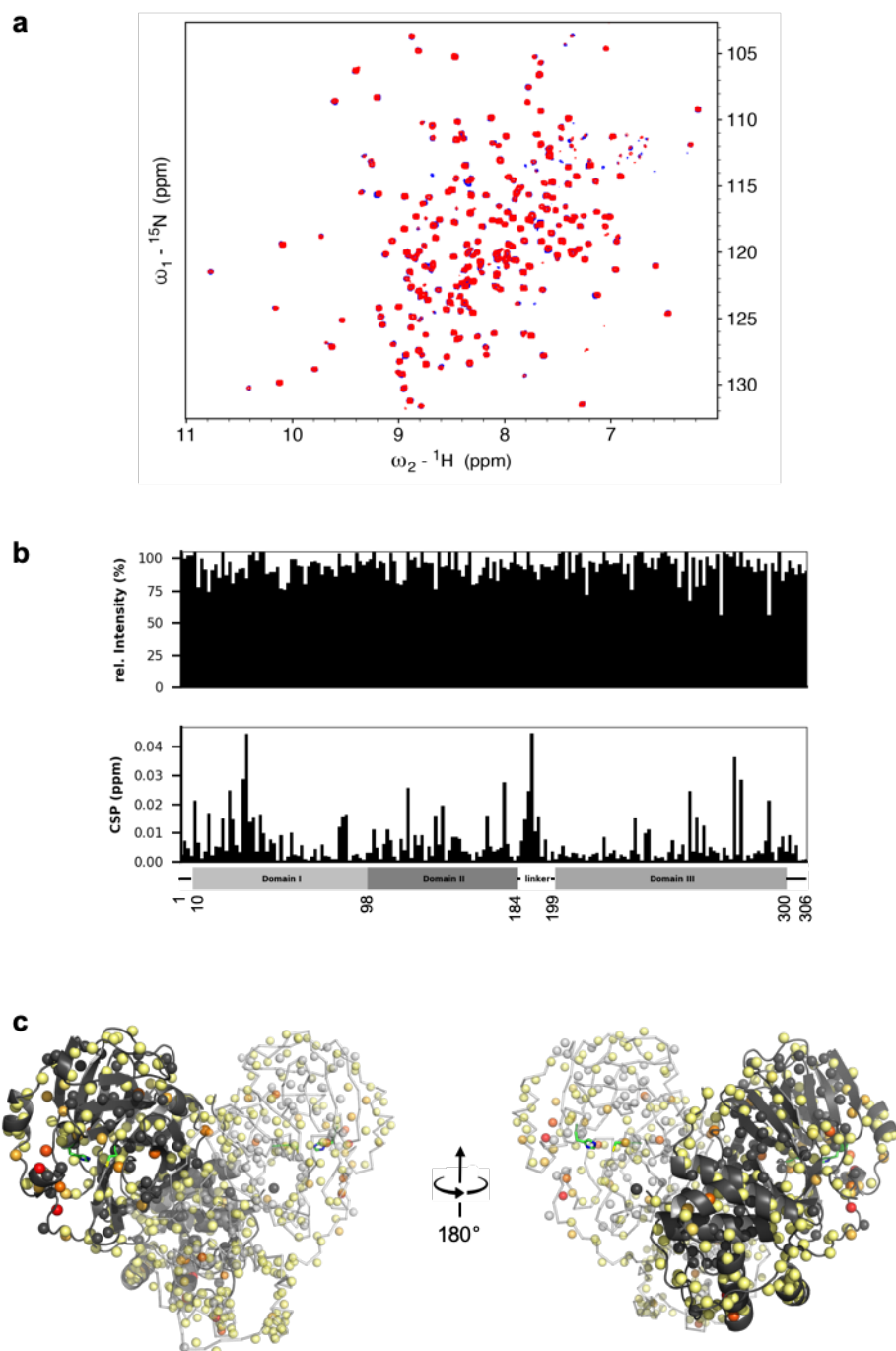


SUPPORTING INFORMATION

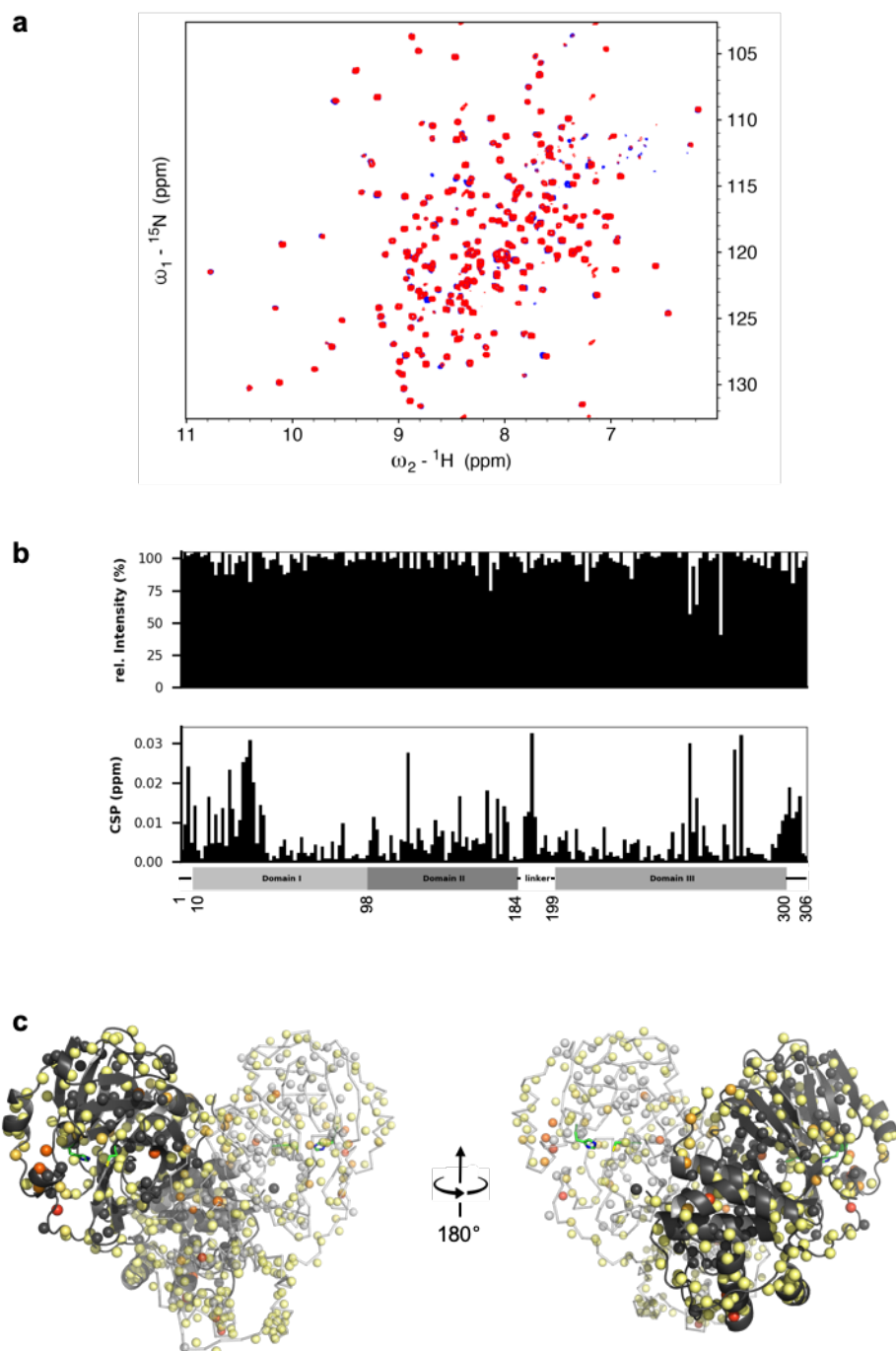
F14



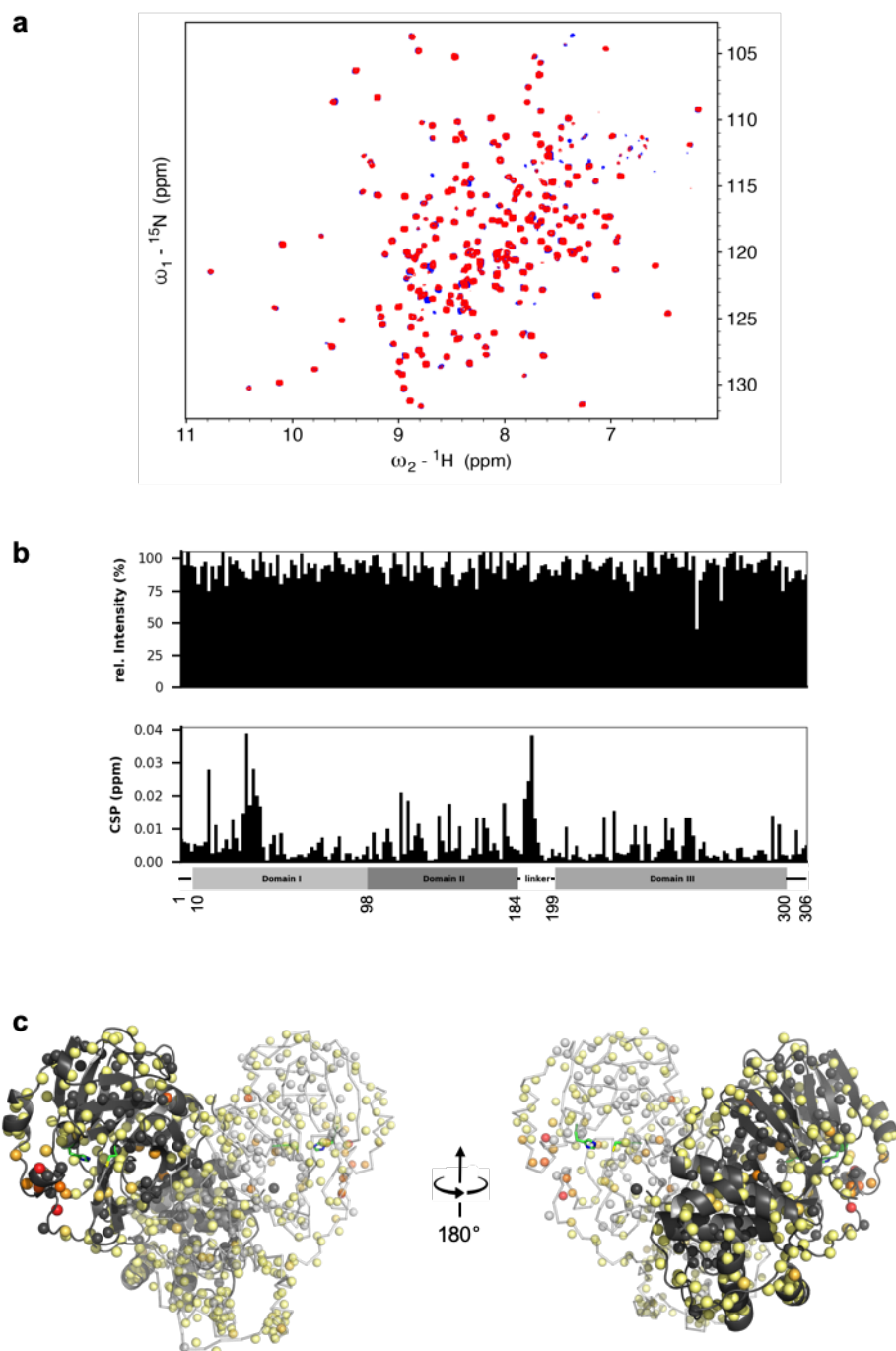
F34



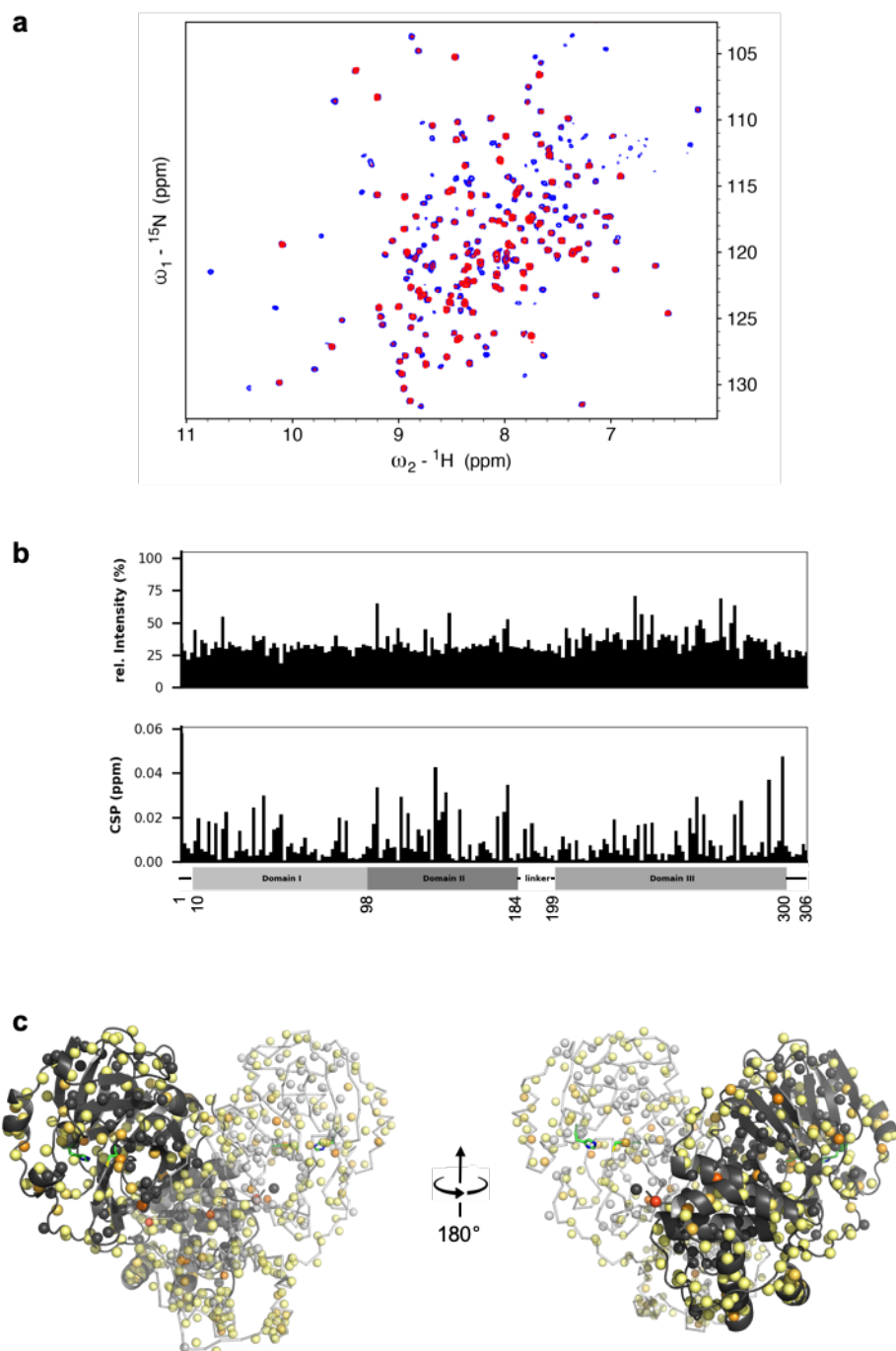
F32



F16



F27



SUPPORTING INFORMATION

Table S3. Crystallographic data and refinement statistics.

| Protein | 3CLp bound to F01 | 3CLp C145ox |
|--|---|---|
| PDB code | 7P51 | 7NTS |
| Data collection statistics | | |
| X-ray source | SOLEIL Proxima2 | SOLEIL Proxima1 |
| Wavelength (Å) | 0.980111 | 0.978564 |
| Solvent content (%) | 38.8 | 36.86 |
| Space group | C 1 2 1 | C 1 2 1 |
| Unit cell dimensions (Å) | a=115.17 b=53.56 c=44.88 | a=113.69 b=52.96 c=44.85 |
| Unit cell angles (°) | α =90.00 β =101.24 γ =90.00 | α =90.00 β =102.95 γ =90.00 |
| Resolution range (Å) ^[a] | 48.44 - 1.47 (1.56 - 1.47) | 47.83 - 1.48 (1.57 - 1.48) |
| N° observations ^[a] | 303693 (47067) | 300295 (46803) |
| N° unique reflections ^[a] | 45078 (7042) | 42473 (6567) |
| <i>R</i> _{meas} ^[a] | 0.074 (1.227) | 0.102 (1.194) |
| Completeness (%) ^[a] | 99.4 (96.9) | 97.1 (93.6) |
| Mean <i>I</i> / σ (<i>I</i>) ^[a] | 13.68 (1.52) | 11.7 (1.48) |
| Multiplicity ^[a] | 6.9 (6.7) | 7.1 (6.7) |
| CC1/2 ^{[a] [b]} | 0.999 (0.645) | 0.998 (0.700) |
| Wilson B-factor (Å ²) | 23 | 19.4 |
| Refinement statistics | | |
| <i>R</i> _{work} / <i>R</i> _{free} ^[c] | 0.181 / 0.216 | 0.166 / 0.207 |
| Average B, all atoms (Å ²) | 25.0 | 22.0 |
| Clashscore ^c | 3 | 3 |
| No. non-H atoms | 2689 | 2697 |
| Protein | 2404 | 2399 |
| Ligand/ion | 30 | 53 |
| Water | 255 | 245 |
| R.m.s. deviations | | |
| Bond lengths (Å) | 0.0121 | 0.0119 |
| Bond angles (°) | 1.856 | 1.700 |
| Ramachandran | | |
| favored | 99.01 | 97.72 |
| allowed | 0.98 | 1.95 |
| outliers | 0.33 | 0.32 |

SUPPORTING INFORMATION

[a] Values in parentheses are for highest-resolution shell. [b] percentage of correlation between intensities from random half-datasets^[32]. [c] calculated for a test set of reflections (5%) omitted from the refinement. defined as the number of clashes calculated for the model per 1000 atoms (including H) of the model. Hydrogens were added by MolProbity^[33].

References

- [1] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J. Biomol. NMR* **1995**, *6*, 277–293.
- [2] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins* **2005**, *59*, 687–696.
- [3] W. Lee, M. Tonelli, J. L. Markley, *Bioinformatics* **2015**, *31*, 1325–1327.
- [4] M. Mayzel, J. Rosenlöw, L. Isaksson, V. Y. Orekhov, *J. Biomol. NMR* **2014**, *58*, 129–139.
- [5] L. Fu, F. Ye, Y. Feng, F. Yu, Q. Wang, Y. Wu, C. Zhao, H. Sun, B. Huang, P. Niu, H. Song, Y. Shi, X. Li, W. Tan, J. Qi, G. F. Gao, *Nat Commun* **2020**, *11*, 4417.
- [6] B. Han, Y. Liu, S. W. Ginzinger, D. S. Wishart, *Journal of Biomolecular NMR* **2011**, *50*, 43–57.
- [7] S. Zhang, N. Zhong, X. Ren, C. Jin, B. Xia, *Biomolecular NMR Assignments* **2011**, *5*, 143–145.
- [8] C. Dalvit, G. Fogliatto, A. Stewart, M. Veronesi, B. Stockman, *Journal of biomolecular NMR* **2001**, *21*, 349–359.
- [9] A. Lingel, A. Vulpetti, T. Reinsperger, A. Proudfoot, R. Denay, A. Frommlet, C. Henry, U. Hommel, A. D. Gossert, B. Luy, A. O. Frank, *Angewandte Chemie International Edition* **2020**, *59*, 14809–14817.
- [10] M. P. Williamson, *Progress in Nuclear Magnetic Resonance Spectroscopy* **2013**, *73*, 1–16.
- [11] S. Eymieux, Y. Rouillé, O. Terrier, K. Seron, E. Blanchard, M. Rosa-Calatrava, J. Dubuisson, S. Belouzard, P. Roingard, *Cell Mol Life Sci* **2021**, *78*, 3565–3576.
- [12] A. L. Hopkins, G. M. Keserü, P. D. Leeson, D. C. Rees, C. H. Reynolds, *Nat Rev Drug Discov* **2014**, *13*, 105–121.
- [13] P. Kirsch, A. M. Hartman, A. K. H. Hirsch, M. Empting, *Molecules* **2019**, *24*, 4309.
- [14] A. Coati, L. M. G. Chavas, P. Fontaine, N. Foos, B. Guimaraes, P. Gourhant, P. Legrand, J.-P. Itie, P. Fertey, W. Shepard, T. Isabet, S. Sirigu, P.-L. Solari, D. Thiaudiere, A. Thompson, *Eur. Phys. J. Plus* **2017**, *132*, 174.
- [15] M. Oscarsson, A. Beteva, D. Flot, E. Gordon, M. Gujjarro, G. Leonard, S. McSweeney, S. Monaco, C. Mueller-Dieckmann, M. Nanao, D. Nurizzo, A. Popov, D. von Stetten, O. Svensson, V. Rey-Bakaikoa, I. Chado, L. Chavas, L. Gadea, P. Gourhant, T. Isabet, P. Legrand, M. Savko, S. Sirigu, W. Shepard, A. Thompson, U. Mueller, J. Nan, M. Eguiraun, F. Bolmsten, A. Nardella, A. Milan-Otero, M. Thunnissen, M. Hellmig, A. Kastner, L. Schmuckermaier, M. Gerlach, C. Feiler, M. S. Weiss, M. W. Bowler, A. Gobbo, G. Papp, J. Sinoir, A. McCarthy, I. Karpics, M. Nikolova, G. Bourenkov, T. Schneider, J. Andreu, G. Cuní, J. Juanhuix, R. Boer, R. Fogh, P. Keller, C. Flensburg, W. Paciorek, C. Vonrhein, G. Bricogne, D. de Sanctis, *J Synchrotron Rad* **2019**, *26*, 393–405.
- [16] W. Kabsch, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 125–132.
- [17] L. Potterton, J. Agirre, C. Ballard, K. Cowtan, E. Dodson, P. R. Evans, H. T. Jenkins, R. Keegan, E. Krissinel, K. Stevenson, A. Lebedev, S. J. McNicholas, R. A. Nicholls, M. Noble, N. S. Pannu, C. Roth, G. Sheldrick, P. Skubak, J. Turkenburg, V. Uski, F. von Delft, D. Waterman, K. Wilson, M. Winn, M. Wojdyr, *Acta Crystallogr D Struct Biol* **2018**, *74*, 68–84.
- [18] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr D Biol Crystallogr* **2011**, *67*, 235–242.
- [19] J. Shi, J. Song, *FEBS Journal* **2006**, *273*, 1035–1045.
- [20] A. L. Kantsadi, E. Cattermole, M.-T. Matsoukas, G. A. Spyroulias, I. Vakonakis, *J. Biomol. NMR* **2021**, *75*, 167–178.
- [21] T. J. El-Baba, C. A. Lutowski, A. L. Kantsadi, T. R. Malla, T. John, V. Mikhailov, J. R. Bolla, C. J. Schofield, N. Zitzmann, I. Vakonakis, C. V. Robinson, *Angewandte Chemie International Edition* **2020**, *59*, 23544–23548.
- [22] B. Goyal, D. Goyal, *ACS Combinatorial Science* **2020**, *22*, 297–305.
- [23] S. Günther, P. Y. A. Reinke, Y. Fernández-García, J. Lieske, T. J. Lane, H. M. Ginn, F. H. M. Koua, C. Ehrt, W. Ewert, D. Oberthuer, O. Yefanov, S. Meier, K. Lorenzen, B. Krichel, J.-D. Kopiccki, L. Gelisio, W. Brehm, I. Dunkel, B. Seychell, H. Gieseler, B. Norton-Baker, B. Escudero-Pérez, M. Domaracky, S. Saouane, A. Tolstikova, T. A. White, A. Hänle, M. Gressler, H. Fleckenstein, F. Trost, M. Galchenkova, Y. Gevorkov, C. Li, S. Awel, A. Peck, M. Barthelmeß, F. Schluenzen, P. L. Xavier, N. Werner, H. Andaleeb, N. Ullah, S. Falke, V. Srinivasan, B. A. França, M. Schwinzer, H. Brognaro, C. Rogers, D. Melo, J. J. Zaitseva-Doyle, J. Knoska, G. E. Peña-Murillo, A. R. Mashhour, V. Henricke, P. Fischer, J. Hakanpää, J. Meyer, P. Gribbon, B. Ellinger, M. Kuzikov, M. Wolf, A. R. Beccari, G. Bourenkov, D. von Stetten, G. Pompidor, I. Bento, S. Panneerselvam, I. Karpics, T. R. Schneider, M. M. Garcia-Alai, S. Niebling, C. Günther, C. Schmidt, R. Schubert, H. Han, J. Boger, D. C. F. Monteiro, L. Zhang, X. Sun, J. Pletzer-Zelgert, J. Wollenhaupt, C. G. Feiler, M. S. Weiss, E.-C. Schulz, P. Mehrabi, K. Karničar, A. Usenik, J. Loboda, H. Tidow, A. Chari, R. Hilgenfeld, C. Uetrecht, R. Cox, A. Zaliani, T. Beck, M. Rarey, S. Günther, D. Turk, W. Hinrichs, H. N. Chapman, A. R. Pearson, C. Betzel, A. Meents, *Science* **2021**, DOI 10.1126/science.abf7945.
- [24] A. Douangamath, D. Fearon, P. Gehrtz, T. Krojer, P. Lukacik, C. D. Owen, E. Resnick, C. Strain-Damerell, A. Aimon, P. Ábrányi-Balogh, J. Brandão-Neto, A. Carbery, G. Davison, A. Dias, T. D. Downes, L. Dunnett, M. Fairhead, J. D. Firth, S. P. Jones, A. Keeley, G. M. Keserü, H. F. Klein, M. P. Martin, M. E. M. Noble, P. O'Brien, A. Powell, R. N. Reddi, R. Skyner, M. Snee, M. J. Waring, C. Wild, N. London, F. von Delft, M. A. Walsh, *Nature Communications* **2020**, *11*, DOI 10.1038/s41467-020-18709-w.
- [25] J. Shi, J. Sivaraman, J. Song, *Journal of Virology* **2008**, *82*, 4620–4629.
- [26] D. W. Kneller, G. Phillips, H. M. O'Neill, R. Jedrzejczak, L. Stols, P. Langan, A. Joachimiak, L. Coates, A. Kovalevsky, *Nature Communications* **2020**, *11*, DOI 10.1038/s41467-020-16954-7.
- [27] D. W. Kneller, G. Phillips, K. L. Weiss, S. Pant, Q. Zhang, H. M. O'Neill, L. Coates, A. Kovalevsky, *Journal of Biological Chemistry* **2020**, *295*, 17365–17373.
- [28] D. Suárez, N. Díaz, *J. Chem. Inf. Model.* **2020**, *60*, 5815–5831.
- [29] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox, R. Hilgenfeld, *Science* **2020**, *368*, 409–412.
- [30] Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L. W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao, H. Yang, *Nature* **2020**, *582*, 289–293.
- [31] R. A. Laskowski, M. B. Swindells, *J. Chem. Inf. Model.* **2011**, *51*, 2778–2786.
- [32] P. A. Karplus, K. Diederichs, *Science* **2012**, *336*, 1030–1033.
- [33] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 12–21.