Aromatic SOFAST-HMBC for proteins at natural ^{13}C abundance

Alicia Vallet, Adrien Favier, Bernhard Brutscher

Aromatic SOFAST-HMBC for proteins at natural ^{13}C abundance

Alicia Vallet, Adrien Favier, and Bernhard Brutscher

Univ. Grenoble Alpes, CEA, CNRS, IBS, 38000 Grenoble, France

6 May 2020

To whom correspondence should be addressed:

E-mail: bernhard.brutscher@ibs.fr
Phone: (33) 4 57 42 85 62

Abstract

We propose here SOFAST-HMBC as a new complementary NMR tool for aromatic side chain assignment of protein samples at natural ^{13}C abundance. The characteristic peak patterns detected in SOFAST-HMBC for each aromatic side chain allow straightforward assignment of all protons and carbons (including quaternary ones) of the aromatic ring, and for tyrosine and phenylalanine, connection to the CB of the aliphatic chain. The performance of SOFAST-HMBC is demonstrated for three small proteins (7 – 14 kDa) at millimolar sample concentration using modern high-field NMR instruments equipped with cryogenically cooled probes. Despite the low amount of NMR-active ^{13}C nuclei in these samples, ^{1}H-^{13}C multiple-bond correlation spectra of good quality were obtained in reasonable experimental times of typically less than 24 hours.

Key words

SOFAST, protein, aromatics, resonance assignment, longitudinal relaxation optimization, cryogenic probe, high field NMR

Introduction

Heteronuclear ^{1}H-^{13}C correlation experiments are routinely used as important NMR tools for the chemical and structural investigation of small organic compounds at natural ^{13}C abundance. Typically, these molecules can be prepared as high concentrated (tens to hundreds of mM) NMR samples, yielding a few 100 µM up to mM amounts of NMR-active ^{13}C nuclei. Natural abundance one-bond ^{1}H-^{13}C correlation experiments (HSBC) have also found some application in the field of protein research, especially at times where uniform ^{13}C isotope-labeling of proteins in bacterial expression systems was not yet commonly available [1,2]. Also, multiple-bond ^{1}H-^{13}C correlation experiments (HMBC) have been proposed in the past for assigning ^{13}C spectra of small proteins at natural abundance. However, at that time, the experiment was performed on a 11 mM sample of BPTI at 63°C [3]. With the advent of high-magnetic field NMR instruments, helium-cooled cryogenic probes, and sensitivity-optimized pulse schemes, the requirements on the minimal amount of NMR-active spins in the NMR detection volume has been dramatically reduced, making today ^{1}H-^{13}C one-bond correlation experiments of proteins at µM concentration feasible within a reasonable experimental time of hours. Here we demonstrate that aromatic ^{1}H-^{13}C long-range correlation experiments at natural ^{13}C abundance may present a valuable additional tool for the NMR investigation of small proteins, using modern NMR instrumentation, and provided that NMR samples at millimolar concentration are available.

Aromatic amino-acid side chains are particularly important for the protein's structure and function. They are generally buried within the hydrophobic protein core, stabilizing the protein fold via hydrophobic contacts, hydrogen bonding, and π-stacking interactions [4]. Aromatic side chains also play an important role in molecular interactions [5], enzyme activity [6,7], and ion coordination [8]. Nuclear spins (^{1}H, ^{13}C, ^{15}N) in the aromatic side chains are therefore essential probes for the protein's structure, dynamics, and functional properties. A number of heteronuclear NMR pulse experiments, exploiting the strong one-bond scalar couplings between adjacent nuclear spins, have been proposed in the past to correlate and assign aromatic nuclei in uniformly ^{13}C or $^{13}C/^{15}N$ labelled proteins [9–11]. However, these experiments suffer from the presence of branched ^{13}C-^{13}C topologies in the aromatic rings, as well as strong coupling effects due to the occurrence of almost identical chemical shift values for neighboring ^{13}C spins [12]. These problems may be avoided if a particular ^{13}C is only bound to ^{12}C, as it is the case (to good approximation) in a protein at natural ^{13}C abundance, where the ^{13}C spins are diluted by a factor ~100 with respect to ^{12}C. This has prompted us to explore the potential of

long-range ^{1}H-^{13}C correlation experiments in aromatic spin systems of proteins at natural ^{13}C abundance.

^{1}H-^{13}C SOFAST-HMQC: one-bond (HSBC) and multiple-bond (HMBC) correlation experiments

The pulse sequence of the SOFAST ^{1}H-^{13}C correlation experiment, used in this work, is depicted in figure 1b. This sequence differs from the originally proposed SOFAST HMQC [13,14] only by two additional ^{1}H 180° pulses that are applied with a BIP shape [15] in order to ensure broadband ^{1}H decoupling during the ^{13}C chemical shift editing period t_1. The SOFAST HMQC experiment employs band-selective shaped ^{1}H pulses for selective manipulation of aromatic (and amide) protons, typically centered at 7.5 ppm and covering a band width of 3 ppm. The selective manipulation of a small subset of proton spins, while leaving all others unperturbed during the experiment yields shortened longitudinal ^{1}H relaxation times, and thus allows for faster repetition rates of the pulse sequence, resulting in higher sensitivity and reduced overall experimental time requirements, as demonstrated for amide [16,17] and aromatic protons [18,19] in proteins, as well as imino protons [20] in nucleic acids. In addition, a further enhancement of the ^{1}H steady-state polarization is achieved by adjusting the nutation angle of the ^{1}H excitation pulse to the Ernst angle (typically 120°). These 2 features make SOFAST the most sensitive heteronuclear correlation experiment available today for protein studies. For applications to protein samples at natural ^{13}C abundance, it is important to keep in mind that the suppression of the hundred-times stronger ^{12}C-bound proton signals is achieved by phase cycling the ^{13}C 90° excitation pulse prior to the ^{13}C (t_1) frequency editing period. For a proper "filtering" performance, the overall stability of the NMR instrument is therefore of utmost importance in order to avoid, or more exactly minimize, spectral t_1-noise arising from a time-modulation of the residual (^{12}C)-^{1}H signals.

We demonstrate the performance of this experiment on a natural ^{13}C abundance sample of lysozyme, a 14.3 kDa protein, purified from hen egg white, that can be purchased (Sigma) at low costs, and is soluble up to a few mM concentrations. The high content (10%) of aromatic residues (1 His, 3 Tyr, 3 Phe, and 6 Trp) in the lysozyme sequence makes this protein a good test case for our experiments. The aromatic one-bond ^{1}H-^{13}C SOFAST-HMQC spectrum (HSBC), recorded with the pulse sequence of figure 1b and the transfer delay Δ set to $\Delta \approx 1/(2J_{CH}) = 2.7$ ms is shown in figure 1c. The spectrum was recorded on a sample of 2 mM

lysozyme in 95% H_2O, 5% D_2O (pH 7.5) at 45°C on a Bruker Avance IIIHD 850 MHz spectrometer equipped with a cryogenically cooled HCN 5 mm probe in an experimental time of 20 min. The spectrum is well resolved allowing to identify all expected ^{1}H-^{13}C spin pairs in the 13 aromatic rings of this protein.

The lysozyme SOFAST-HMBC spectrum recorded with the pulse sequence of figure 1b, but setting the ^{1}H-^{13}C transfer delay to $\Delta = 50$ ms, is shown in figure 1d. The transfer delay has been optimized experimentally by recording a series of spectra with transfer delays of 30, 40, 50, 60, and 70 ms. The total acquisition time for the spectrum shown in figure 1d was 12 hours. The correlation peaks between aromatic ring protons and carbons are spread over ^{13}C chemical shifts ranging from 36 to 160 ppm, and they can have either positive or negative intensities, as will be explained later on. Note that in this spectrum, we also observe cross peaks between amide protons, resonating within the ^{1}H excitation window, and carbonyl and aliphatic carbons, although typically with significantly lower signal-to-noise ratios (data not shown). The question how to analyze such a "crowded" spectrum and assign the detected cross peaks to individual aromatic side chains will be discussed in the following section.

Identifying aromatic spin systems in SOFAST-HMBC

The correlation peaks detected in the SOFAST-HMBC spectrum are due to sizeable ^{1}H-^{13}C scalar couplings over two or three chemical bonds ($^{2}J_{HC}$, $^{3}J_{HC}$) in the aromatic side chains of the protein. Although some information on the magnitude of these coupling constants is available from NMR measurements on small compounds, such as benzene derivatives [21], only sparse data are available for aromatic protein side chains. Therefore, we relied on the identification of characteristic peak patterns for the different aromatic residues, using the information available in the BMRB on the ^{13}C chemical shift ranges for the different ^{13}C sites, for assigning the proton and carbon resonances of individual aromatic side chains.

The most straightforward case is tyrosine (figure 2a) that can be easily distinguished by the characteristic frequency range of the ^{13}C at the CZ position (155 – 160 ppm). The CZ gives a strong cross peak with HD and a weaker one with HE. The HD proton can also be correlated to the CB, while the HE gives an additional peak with CG. The SOFAST-HMBC spectrum thus allows to unambiguously connect and assign all ^{1}H and ^{13}C of the tyrosine ring, as well as the side chain CB which connects the ring system with the aliphatic part of the protein chain. This allows sequential assignment if the backbone assignment has been achieved with a set of triple-

resonance experiments as typically performed on a uniformly $^{13}C/^{15}N$ labeled sample of the protein.

The HD of phenylalanine also gives rise to a cross peak with the side chain CB. Therefore, the remaining (non-tyrosine) peaks detected in the upper part of the SOFAST-HMBC spectrum can be safely assigned to the three Phe spin systems of lysozyme (figure 2b). The HD also gives a cross peak with CZ. The HZ can then be connected to the spin system via a cross peak with CD. Unfortunately, the HE, CE, and CG are disconnected from HD, HZ, CB, CD, CZ, and can thus not be unambiguously assigned from this experiment alone. Assignment of the Phe side chain is also challenging because of similar, overlapping ^{13}C frequency ranges of the ring carbons (CG, CD, CE, CZ) that may result in degenerate frequencies.

Histidine assignment is again straightforward, as both ring protons (HD2 and HE1) correlate with all three ring carbons (CD2, CG, CE1), and the HE1 proton is downfield shifted, thus providing a convenient starting point for spin-system identification (figure 2c). Unfortunately, no cross peak is detected between one of the 2 aromatic protons and the side chain CB. The measured CD2 and CG chemical shifts also provide a direct and straightforward measure of the tautomeric state of the imidazole ring [22]. In the present case, the large chemical shift difference observed between CD2 and CG indicates the presence of a τ tautomer with the imidazole ring protonated at the NE2 position.

Finally, as expected, tryptophan shows the most complex peak pattern due to its indole ring structure with a total of 8 carbons and 5 protons, giving rise to up to 17 cross peaks detected in the SOFAST-HMBC spectrum. Despite this complexity, it was possible to identify all 6 Trp spin systems. For 5 of them, all expected 17 correlation peaks could be assigned, while the sixth one gives rise to correlation peaks with lower signal to noise ratio, most likely due to intermediate time scale dynamics of this particular side chain at 45°C. Therefore, no correlation peaks to the CG were detected for this spin system, and also correlations to the HD1 and HH2 are missing. Two examples of Trp spin systems, identified in the SOFAST-HMBC spectrum of lysozyme, are shown in figure 3. For Trp, no cross peak is detected between aromatic protons and the side chain CB, and thus sequential assignment to a particular residue in the protein sequence is not possible from this experiment alone.

Schematic representations summarizing the peak patterns observed in the SOFAST-HMBC spectrum for the four aromatic residue types are shown in figure 4, together with the

expected ^{13}C chemical shift ranges. Interestingly, for tyrosine, histidine, and tryptophan, these peak patterns form correlated networks that allow the interconnection of all ^{1}H and ^{13}C within the same side chain, at least as long as peak degeneracy is not a limiting issue. For phenylalanine, as pointed out above, the (HE, CE, CG) and (HD, HZ, CB, CD, CZ) spins are unconnected and form separate "spin systems" that can be identified from the SOFAST-HMBC spectrum.

Aromatic ^{2}J and ^{3}J ^{1}H-^{13}C coupling constants and transfer efficiencies

The transfer efficiency of the SOFAST HMBC sequence (figure 1b), and thus the corresponding cross peak intensity for a particular ^{1}H-^{13}C correlation, is given by the following relation:

$$S(\Delta) \propto sin^2(\pi\, J_{CH}\Delta)\, cos^n\left(\pi\ {}^{3}J_{HH}\, 2\Delta\right) \exp\left(-2\Delta/T_2^{eff}\right) \qquad (1)$$

with J_{CH} the one-bond or multiple-bond ^{1}H-^{13}C coupling constant, $^{3}J_{HH}$ the vicinal ^{1}H-^{1}H coupling constant, n the number of vicinal protons, and T_2^{eff} an effective relaxation time constant, accounting for relaxation-induced signal loss during the coherence-transfer delays. Equation (1) neglects the presence of protons separated by more than 3 chemical bonds, as the corresponding coupling constants are typically less than 2-3 Hz, while the vicinal ^{1}H-^{1}H scalar couplings vary between 7 and 10 Hz. Interestingly, the ^{1}H-^{1}H coupling evolution is not only a source of undesirable signal loss, but it also explains the sign of the detected cross peak intensities in the SOFAST-HMBC spectrum, recorded with an experimentally optimized transfer delay Δ of 50 ms. A proton coupled to only one other proton gives rise to negative cross peaks, while a proton with zero or two neighboring (vicinal) protons can be identified from correlation peaks of positive intensity. For example, the peak pattern detected for phenylalanine (figure 4b) shows negative cross peaks for HD that is only coupled to HE, while HE (coupled to HD and HZ) and HZ (coupled to HE1 and HE2) show up with positive intensities. This topologic information, encoded in the sign of the HMBC correlation peaks, greatly helps in peak pattern identification and aromatic resonance assignment. It is also worth noting that one-

bond ^{1}H-^{13}C correlations in SOFAST-HMBC are absent or of low intensity, independent of the exact value chosen for the transfer delay, because of the significantly reduced transverse relaxation time of a ^{13}C-bound ^{1}H compared to a ^{12}C-bound ^{1}H. These correlations are therefore best obtained from the SOFAST-HSBC using short transfer delays adjusted to the strong one-bond ^{1}H-^{13}C couplings (figure 1c).

Finally, we attempted to get some more quantitative information about the strengths of the various 2-bond and 3-bond ^{1}H-^{13}C coupling constants in the aromatic protein side chains. NMR experiments performed on benzene [21] in the 1960s yielded coupling constants of 7.4 Hz for $^{3}J_{CH}$, and 1.0 Hz for $^{2}J_{CH}$. Here, we performed quantitative J_{CH} measurements that are based on the pulse sequence of figure 1b, inserting additional ^{13}C 180°C broadband inversion pulses during the transfer delays Δ. This allows tuning of the effective J_{CH}-coupling evolution from 0 to Δ (set to 60 ms), while ensuring identical signal loss due to spin relaxation and ^{1}H-^{1}H coupling evolution for each measurement point. A pseudo-3D data set is recorded with 2 frequency dimensions, and the third dimension corresponding to the J_{CH} coupling evolution time T. The extracted intensity buildups for each cross peak, detected in the spectrum of figure 1d, were then fitted to the function $I(T) = A \sin^2(\pi J_{CH} T)$, with A and J_{CH} the fitting variables. The results of these J-coupling measurements obtained for a highly concentrated (~ 8 mM) sample of lysozyme are provided in table 1. Two data sets were recorded at a sample temperature of 45°C, and magnetic field strengths corresponding to ^{1}H frequencies of 850 MHz and 600 MHz, respectively, in order to estimate the reproducibility of the quantitative J coupling measurements.

Application to small ^{15}N-labeled proteins

In this last section, we will show additional applications of SOFAST-HMBC to two small ^{15}N-labeled proteins, produced in an *E. coli* expression system using standard isotope-labeling protocols. Such samples are often available for NMR studies of the structure, function, and interaction of a protein, used for example for initial sample characterization and optimization, for ^{15}N relaxation experiments, or for chemical shift mapping experiments with an interacting molecule. We first performed aromatic SOFAST-HSBC and SOFAST-HMBC experiments on a 1.5 mM ^{15}N-labeled sample of ubiquitin, a 76-residue protein that is widely used as a NMR

standard. An overlay of the recorded spectra is shown in figure 5a. Ubiquitin is relatively poor in aromatic residues, with only 1 histidine, 2 phenylalanine, and 1 tyrosine. Therefore, the spectra in figure 5a are sparse with only a small number of observed correlation peaks (compared to lysozyme), making it easy to identify and assign the different aromatic spin systems. All cross peaks expected according to the peak pattern graphs of figure 4 are visible in these spectra, except for F4 where no peak is detected for the correlation HE-CG. Also, the correlation peak HD-CB is very weak for this spin system, indicating some NMR-unfavorable ring flip motion of this particular aromatic side chain.

As a third test case, we have chosen the small bacterial cold-shock protein CspA (69 residues). CspA functions as a RNA chaperone that transiently binds single-stranded nucleic acids (RNA or DNA) via interactions that are mainly mediated by aromatic side chains. CspA has 10 aromatic residues (1 His, 6 Phe, 1 Trp, 1 Tyr, and 1 His), and half of these aromatic side chains (W11, F18, F20, F31, and H33) are located at the protein surface, and are involved in RNA binding. An overlay of the SOFAST-HSBC and HMBC spectra, recorded on a moderately concentrated sample (0.7 mM) of ^{15}N-labeled CspA at 30°C, is shown in Figure 5b. For this sample, only parts of the expected peak patterns are detected. While Y42 gives rise to strong correlation peaks, and the complete peak pattern is detected, for other aromatic side chains, some cross peaks are of low intensity or even completely absent. This observation holds especially true for the surface exposed side chains that require some local mobility in order to facilitate accommodation of the binding interface to different nucleic acids. [23] We therefore conclude that the relative intensity of peak patterns detected for different aromatic side chains in SOFAST-HMBC is a direct indication of the presence/absence of conformational exchange effects (ring flips, side chain motions, …) affecting the transverse relaxation of the involved spin coherences.

Conclusions

We have shown herein that SOFAST-HMBC presents an attractive additional NMR pulse tool for the study of aromatic side chains in small proteins at natural ^{13}C abundance, provided that protein samples at mM concentration are available. For each aromatic residue type a particular peak pattern is detected (expected) due to sizeable 2-bond and 3-bond ^{1}H-^{13}C scalar couplings. In addition, the topological information on the number of vicinal protons is encoded in the sign of the detected cross peaks. This allows correlation and assignment of all protons and carbons of the aromatic ring, as well as for tyrosine and phenylalanine, the extension to the

CB of the aliphatic chain. We are confident that this experiment will prove useful for aromatic resonance assignment of proteins that cannot be isotope labeled, or as a complement to other NMR data recorded on uniformly ^{13}C labeled protein samples.

Acknowledgments

We thank Isabel Ayala and Karine Giandoreggio for protein sample preparation. This work used the NMR and isotope labeling platforms of the Grenoble Instruct center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB).

Figure captions

Figure 1. (a) Schematic drawing of the 4 aromatic protein side chains: phenylalanine (F), tyrosine (Y), tryptophan (W), and histidine (H). Semi-bold letters indicate proton-bound carbons. (b) SOFAST-HMQC pulse sequence for single-bond (HSBC) and multiple-bond (HMBC) correlation experiments of aromatic side chains in proteins at natural ^{13}C abundance. Filled and open pulse symbols indicate 90° and 180° rf pulses, respectively. Unless indicated, all pulses are applied with phase x. The following pulse shapes are used for band-selective ^{1}H pulses: PC9 for excitation [24], REBURP for refocusing [25]. They are typically centered at 7.5 ppm, covering a bandwidth of 3.0 ppm (aromatic region). In addition, a pair of broadband ^{1}H inversion pulses (BIP) [15] ensures broadband heteronuclear decoupling during t_1. For optimal sensitivity of the experiment, the inter-scan delay and nutation angle (PC9 pulse) are typically set to 300 ms and 120°. The ^{1}H-^{13}C transfer delay is set to Δ = 2.5 - 3.0 ms for one-bond correlation experiments, and to Δ = 40 - 50 ms for multiple-bond correlation experiments. A 2-step phase cycle is applied with φ_1 = x, -x and φ_{rec} = x, -x. Quadrature detection in t_1 is achieved by time-proportional phase incrementation of φ_1 according to STATES-TPPI. ^{13}C decoupling during detection makes use of a WALTZ65 sequence covering a bandwidth of 30 ppm. A pair of pulsed field gradients (PFGz) ensures good water suppression. (c) SOFAST-HSBC spectrum recorded on a 2 mM sample of hen egg white lysozyme at 45°C on a Bruker Avance IIIHD 850 MHz spectrometer equipped with a cryogenically cooled HCN probe. All observed peaks are annotated by their atom type, as identified from the SOFAST-HMBC spectrum, using a color code for the amino-acid type and the nomenclature introduced in (a). Note that the His HE1-CE1peak is outside the plotted spectral region. The SOFAST-HMBC spectrum recorded on the same lysozyme sample and NMR spectrometer is plotted in (d).

Figure 2. Spectral overlay of SOFAST-HSBC (red peak contours) and SOFAST-HMBC (blue and green peak contours for positive and negative intensities, respectively) of hen egg white lysozyme. The peak patterns observed for (a) tyrosine, (b) phenylalanine, and (c) histidine side chains are highlighted by dashed lines and circles. Only the relevant spectral regions are plotted. For tyrosine and phenylalanine, the peak patterns can be assigned to a specific residue via the known CB frequency (BMRB entry 4562).

Figure 3. Spectral overlay of SOFAST-HSBC (red peak contours) and SOFAST-HMBC (blue and green peak contours for positive and negative intensities, respectively) of hen egg

white lysozyme, and identification of peak patterns for two tryptophan side chains, highlighted by dashed lines and circles.

Figure 4. Peak pattern topologies detected in the SOFAST-HSBC and SOFAST-HMBC spectra for (a) tyrosine, (b) phenylalanine, (c) histidine, and (d) tryptophan. A red colored peak indicates a one-bond correlation that is detected in SOFAST-HSBC, while blue and green colors are indicative of positive and negative peaks observed in SOFAST-HMBC. The number highlights whether the transfer is mediated by a one-bond (1), two-bond (2), or three-bond ^{1}H-^{13}C coupling. In addition, ^{13}C chemical shift ranges (average and standard deviation) as extracted from the BMRB data bank are indicated on the right.

Figure 5. Spectral overlay of SOFAST-HSBC (red peak contours) and SOFAST-HMBC (blue and green peak contours) of (a) ubiquitin and (b) CspA. Spectra were recorded on a 1.5 mM ^{15}N-labled ubiquitin sample at 850 MHz and 30°C, and on a 0.7 mM ^{15}N-labled CspA sample at 700 MHz. Overall data acquisition times were 30 min (HSBC) - 6 hours (HMBC) for ubiquitin, and 2 hours (HSBC) - 48 hours (HMBC) for CspA. The different peak patterns identified for the aromatic side chains of the 2 proteins are annotated and color-coded.

Table

Tyrosine			
^{2}J (CZ-HE)	**^{3}J (CG-HE)**	**^{3}J (CZ-HD)**	**^{3}J (CB-HD)**
< 3.0	7.0 ± 0.3	9.2 ± 0.3	4.5 ± 1.0
Phenylalanine			
^{3}J (CD-HZ)	**^{3}J (CG-HE)**	**^{3}J (CZ-HD)**	**^{3}J (CB-HD)**
6.1 ± 0.7	6.5 ± 0.5	6.4 ± 0.8	4.5 ± 1.0
Histidine			
^{2}J (CG-HD2)	**^{3}J (CE1-HD2)**	**^{3}J (CD2-HE1)**	**^{3}J (CG-HE1)**
11.5 ± 0.5	3.5 ± 0.5	4.4 ± 0.5	5.8 ± 0.5
Tryptophane			
^{2}J (CG-HD1)	**^{3}J (CZ3-HZ2)**	**^{3}J (CD2-HZ2)**	**^{3}J (CD2-HD1)**
7.1 ± 0.5	6.6 ± 0.5	6.5 ± 0.5	6.7 ± 0.5
^{3}J (CE2-HD1)	**^{3}J (CG-HE3)**	**^{3}J (CH2-HE3)**	**^{3}J (CE2-HE3)**
7.2 ± 0.5	< 3.0	6.7 ± 0.5	7.4 ± 0.5
^{3}J (CZ2-HZ3)	**^{3}J (CD2-HZ3)**	**^{3}J (CE3-HH2)**	**^{3}J (CE2-HH2)**
6.9 ± 0.5	7.5 ± 0.5	8.1 ± 0.3	8.8 ± 1.0

Table 1. Scalar two-bond (^{2}J) and three-bond (^{3}J) ^{1}H-^{13}C coupling constants (in Hz) measured for aromatic side chains in the protein lysozyme. Note that ^{3}J(CDHD) and ^{3}J(CEHE) couplings in Phe and Tyr were not extracted due to the potential overlap with cross peaks resulting from transfer via one-bond ^{1}H-^{13}C couplings.

References

[1] H. Kessler, P. Schmieder, W. Bermel, Complete Assignment of the Noncarbonylic Carbon- 13 Resonances of Tendamistatt, Biopolymers. 30 (1990) 465–475.

[2] S. Medvedeva, J.-P. Simorre, B. Brutscher, F. Guerlesquin, D. Marion, Extensive ^{1}H NMR resonance assignment of proteins using natural abundance gradient-enhanced ^{13}C-^{1}H correlation spectroscopy, FEBS Lett. 333 (1993).

[3] P.E. Hansen, Assignment of the Natural Abundance C-13 Spectrum of Proteins Using C-13 H-1-Detected Heteronuclear Multiple-Bond Correlation Nmr-Spectroscopy - Structural Information and Stereospecific Assignments from 2-Bond and 3-Bond Carbon Hydrogen Coupling-Constant, Biochemistry. 30 (1991) 10457–10466.

[4] A.A. Bogan, K.S. Thorn, Anatomy of hot spots in protein interfaces, J. Mol. Biol. 280 (1998) 1–9.

[5] C.M. Baker, G.H. Grant, Role of Aromatic Amino Acids in Protein–Nucleic Acid Recognition Role of Aromatic Amino Acids in Protein–Nucleic Acid Recognition Christopher, Biopolymers. 85 (2007) 456–470.

[6] G.J. Bartlett, C.T. Porter, N. Borkakoti, J.M. Thornton, Analysis of catalytic residues in enzyme active sites, J. Mol. Biol. 324 (2002) 105–121.

[7] G.L. Holliday, J.B.O. Mitchell, J.M. Thornton, Understanding the Functional Roles of Amino Acid Residues in Enzyme Catalysis, J. Mol. Biol. 390 (2009) 560–577.

[8] M.M. Yamashita, L. Wesson, G. Eisenman, D. Eisenberg, Where metal ions bind in proteins., Proc. Natl. Acad. Sci. 87 (1990) 5648–5652.

[9] T. Yamazaki, J.D. Forman-Kay, L.E. Kay, Two-Dimensional NMR Experiments for Correlating 13CB and 1HD Chemical Shifts of Aromatic Residues in 13C-Labeled Proteins via Scalar Couplings, J. Am. Chem. Soc. 115 (1993) 11054–11055.

[10] J.J. Prompers, A. Groenewegen, C.W. Hilbers, H.A.M. Pepermans, Two-Dimensional NMR Experiments for the Assignment of Aromatic Side Chains in 13 C-labeled Proteins, J Biomol NMR. 75 (1998) 68–75.

[11] F. Löhr, H. Rüterjans, Novel pulse sequences for the resonance assignment of aromatic

side chains in 13C-labeled proteins, J. Magn. Reson. - Ser. B. 112 (1996) 259–268.

[12] J. Jacob, J.M. Louis, I. Nesheiwat, D.A. Torchia, Biosynthetically directed fractional13C labeling facilitates identification of Phe and Tyr aromatic signals in proteins, J. Biomol. NMR. 24 (2002) 231–235.

[13] P. Schanda, B. Brutscher, Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds., J. Am. Chem. Soc. 127 (2005) 8014–8015.

[14] P. Schanda, E. Kupce, B. Brutscher, SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds., J. Biomol. NMR. 33 (2005) 199–211.

[15] M.A. Smith, H. Hu, A.J. Shaka, Improved Broadband Inversion Performance for NMR in Liquids, J. Magn. Reson. 151 (2001) 269–283.

[16] B. Brutscher, Z. Solyom, Polarization-enhanced Fast-pulsing Techniques, in: Fast NMR Data Acquis., 2017: pp. 1–48.

[17] P. Schanda, Fast-pulsing longitudinal relaxation optimized techniques: Enriching the toolbox of fast biomolecular NMR spectroscopy, Prog. Nucl. Magn. Reson. Spectrosc. 55 (2009) 238–265.

[18] A. Eletsky, H.S. Atreya, G. Liu, T. Szyperski, Probing Structure and Functional Dynamics of (Large) Proteins with Aromatic Rings : L-GFT-TROSY (4 , 3) D HCCH NMR Spectroscopy, J. Am. Chem. Soc. 127 (2005) 14578–14579.

[19] N.E. Christou, B. Brutscher, BEST and SOFAST experiments for resonance assignment of histidine and tyrosine side chains in 13C/15N labeled proteins, J. Biomol. NMR. 72 (2018) 115–124.

[20] J. Farjon, J.R. Boisbouvier, P. Schanda, A. Pardi, J.-P. Simorre, B. Brutscher, Longitudinal-relaxation-enhanced NMR experiments for the study of nucleic acids in solution, J. Am. Chem. Soc. 131 (2009).

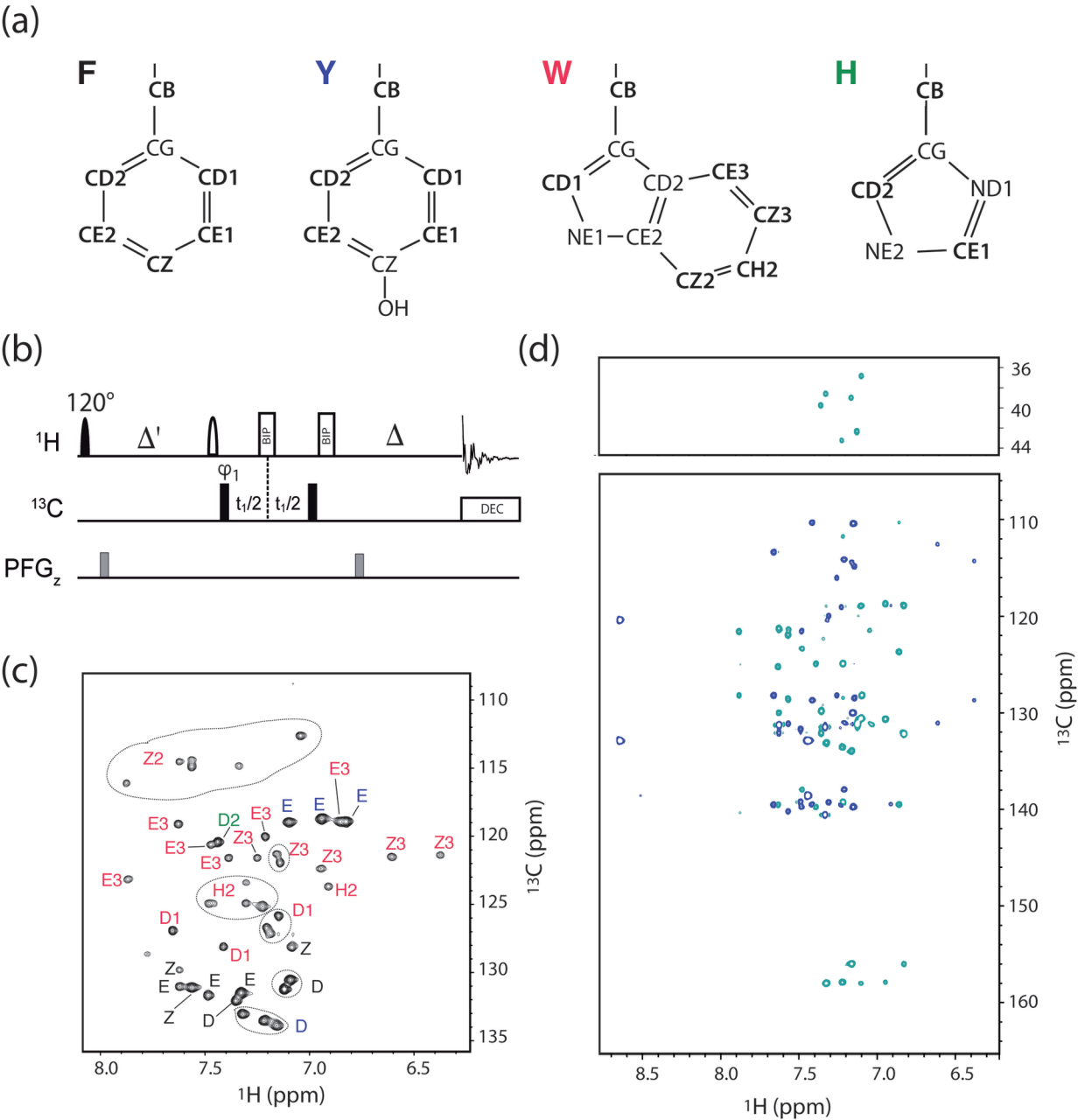
[21] F.J. Weigert, J.D. Roberts, Nuclear Magnetic Resonance Spectroscopy. Long-Range Spin-Spin Couplings Involving Carbon-13, J. Am. Chem. Soc. 91 (1969) 4940–4941.

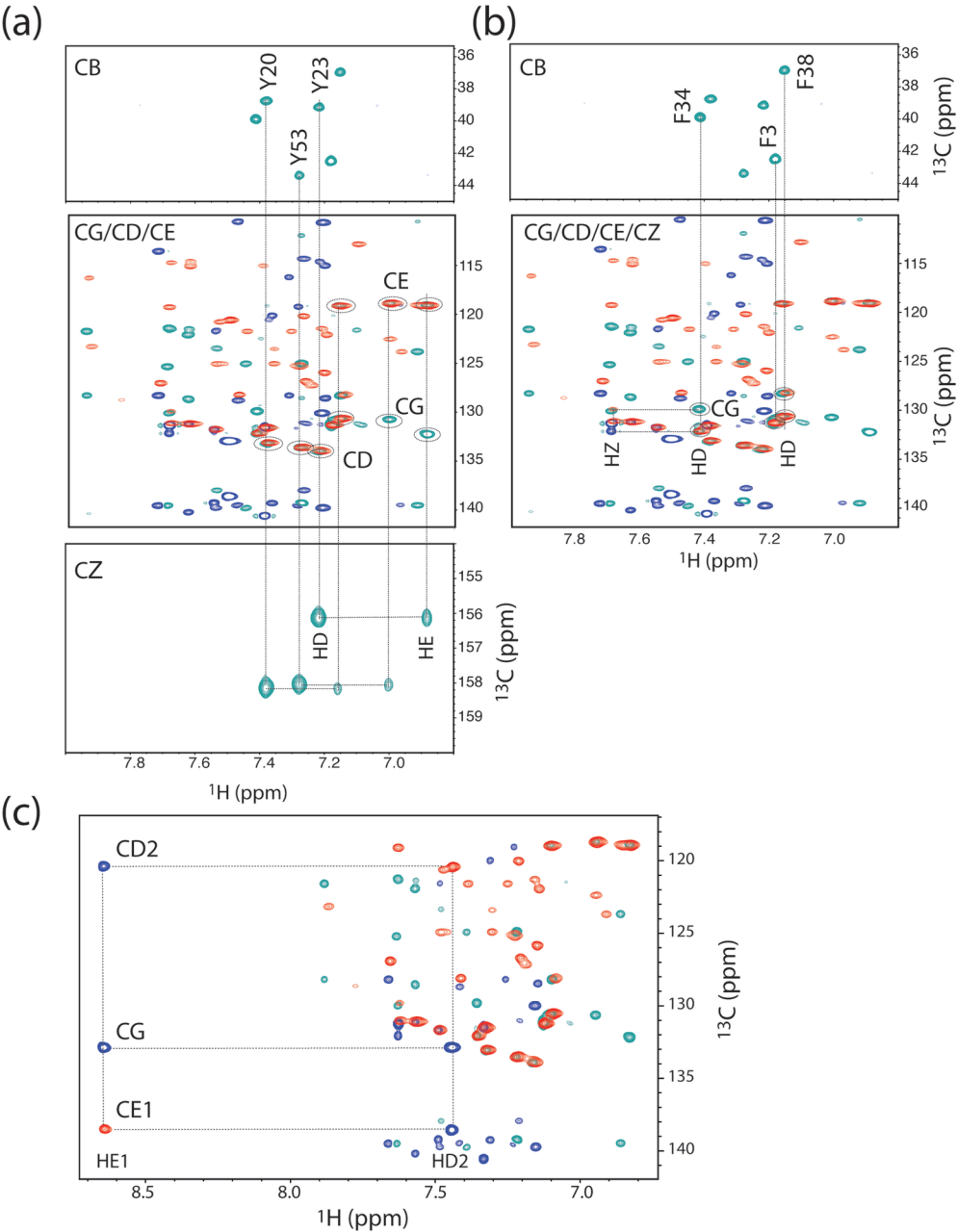
[22] A.L. Hansen, L.E. Kay, Measurement of histidine pKa values and tautomer populations in invisible protein states, Proc. Natl. Acad. Sci. 111 (2014) E1705–E1712.

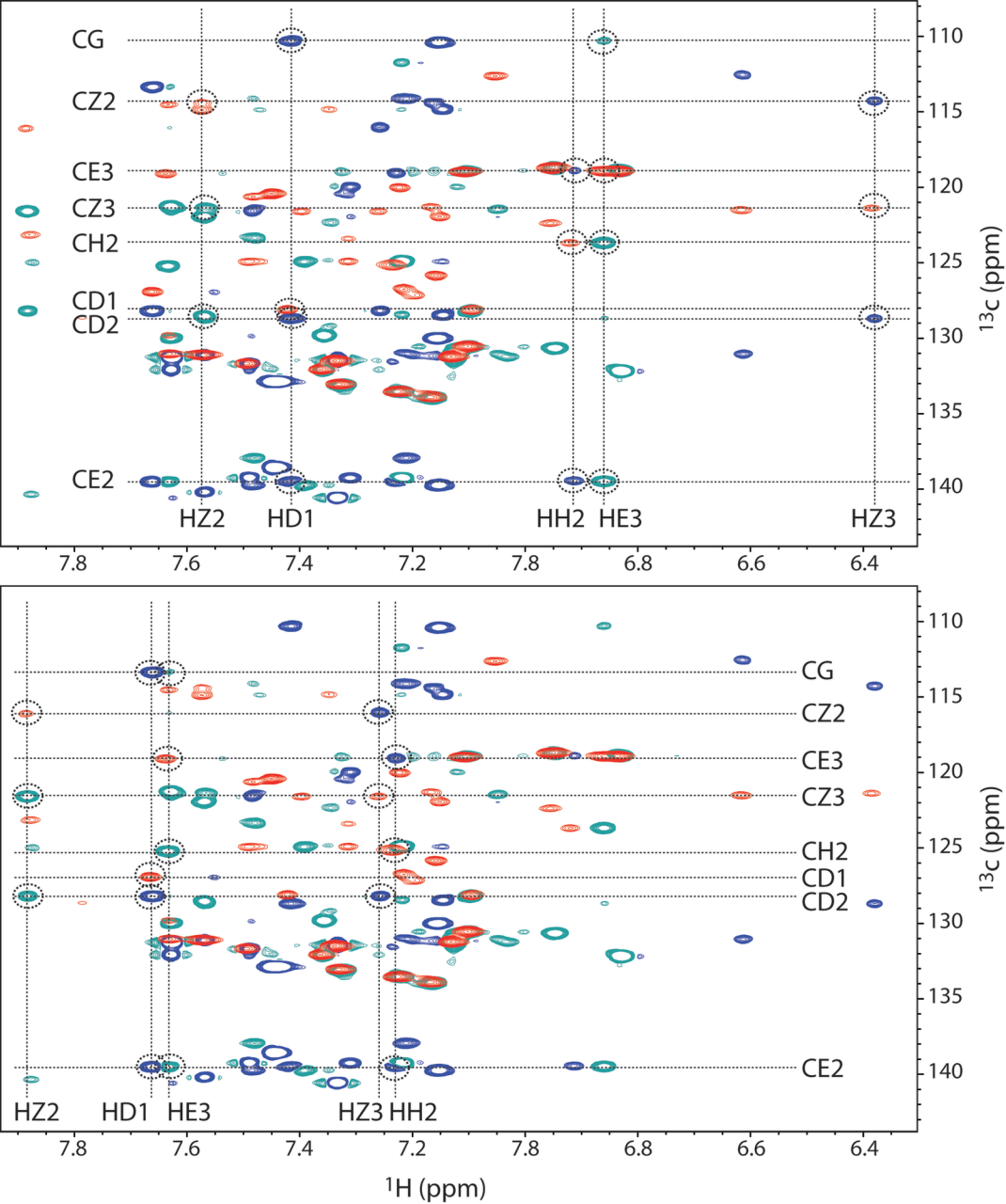
[23] E. Rennella, T.S. Sára, M. Juen, C. Wunderlich, L. Imbert, Z. Solyom, A. Favier, I. Ayala, K. Weinhäupl, P. Schanda, R. Konrat, C. Kreutz, B. Brutscher, RNA binding and chaperone activity of the E. coli cold-shock protein CspA, Nucleic Acids Res. (2017) 1–14.

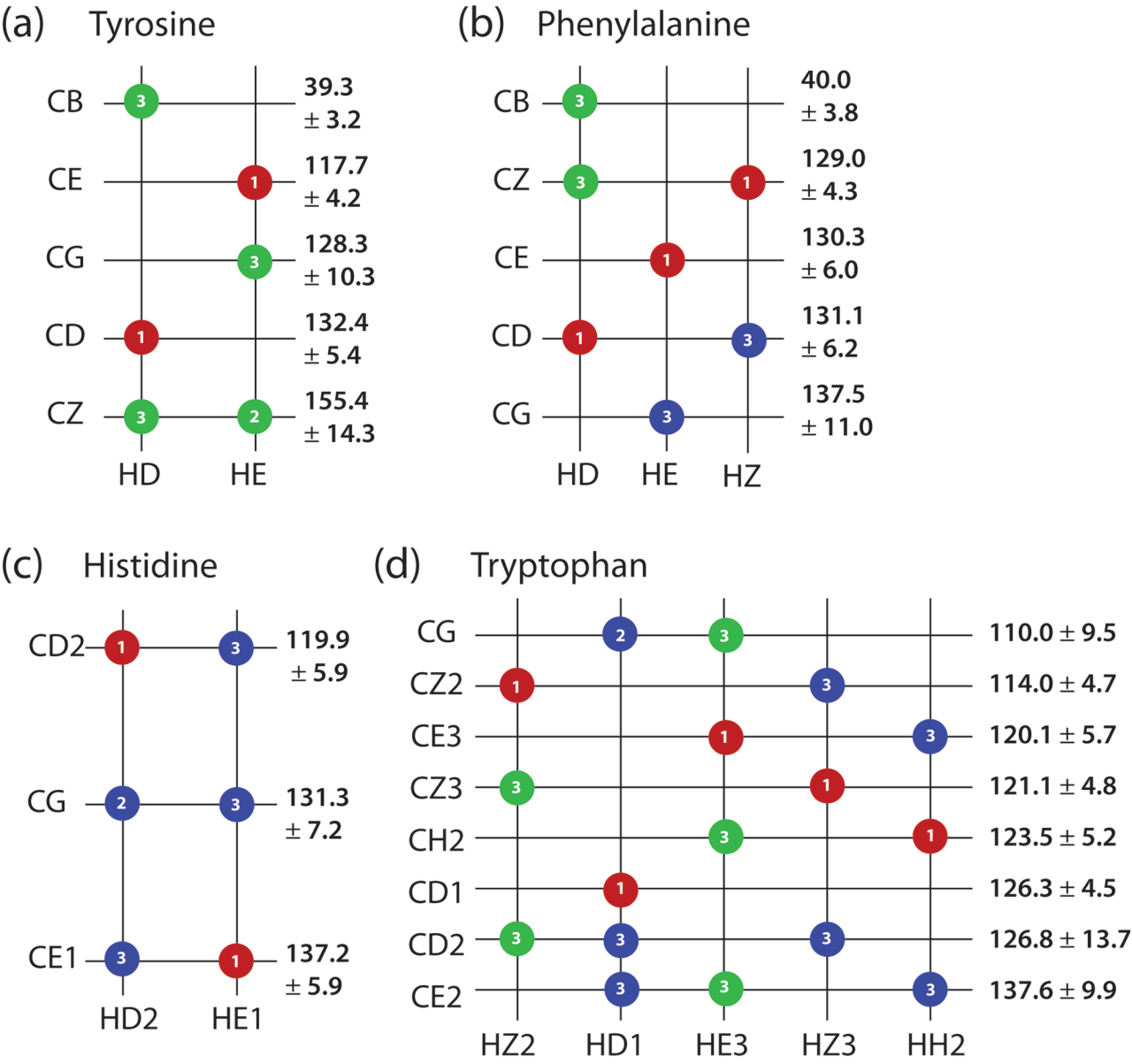
[24] E. Kupce, R. Freeman, Polychromatic Selective Pulses, J. Magn. Reson. A. 102 (1993) 122–126.

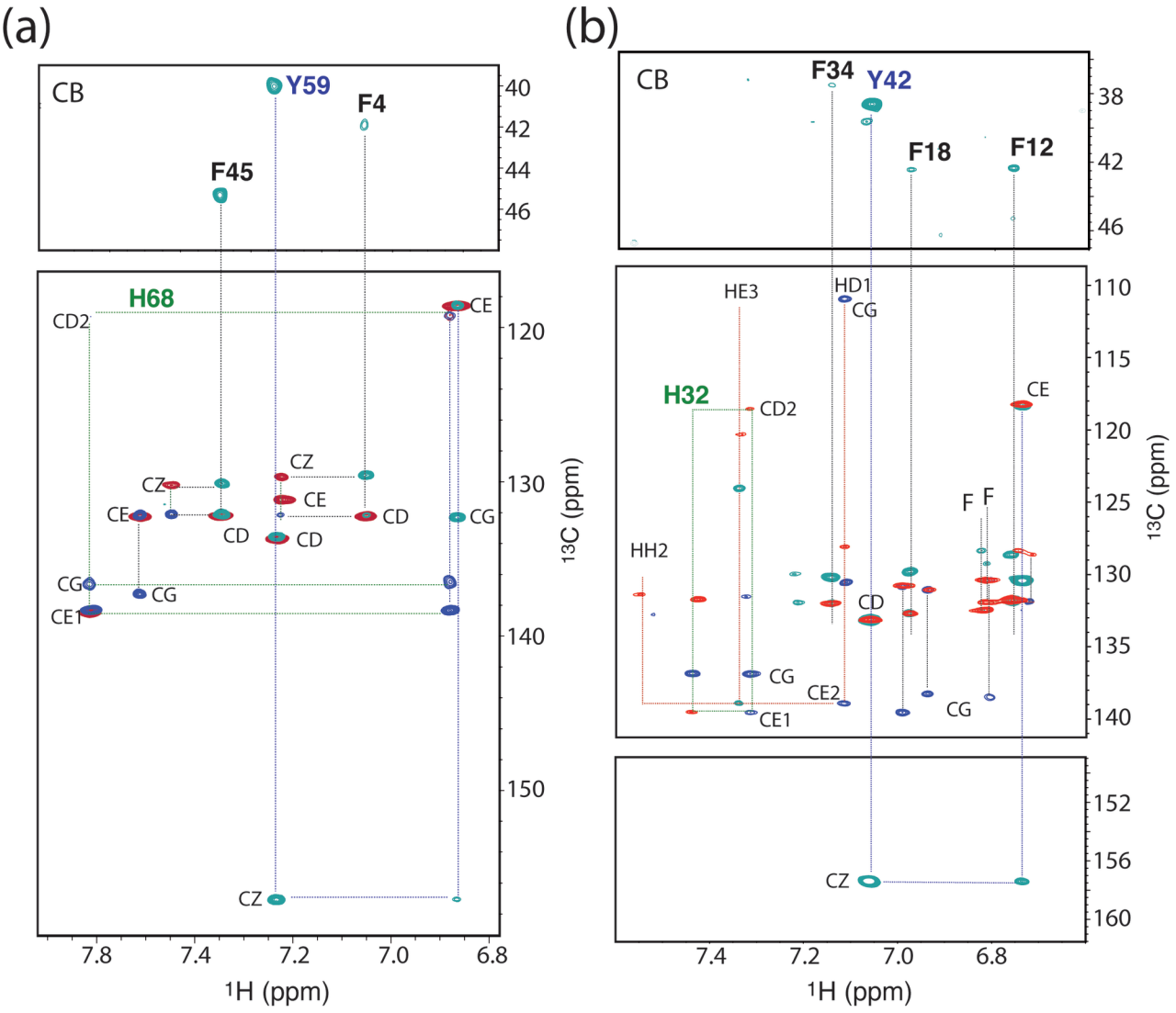
[25] H. Geen, R. Freeman, Band-Selective Radiofrequency Pulses, J. Magn. Reson. 93 (1991) 93–141.


(a)
F
CB
CG
CD2
CD1
CE2
CE1
CZ
Y
CB
CG
CD2
CD1
CE2
CE1
CZ
OH
W
CB
CG
CD1
CD2
CE3
CZ3
NE1
CE2
CZ2
CH2
H
CB
CG
CD2
ND1
NE2
CE1
(b)
120°
1H
Δ'
BIP
BIP
Δ
φ1
13C
t1/2
t1/2
DEC
PFGz
(c)
Z2
E3
D2
E
Z3
H2
D1
Z
D
13C (ppm)
1H (ppm)
(d)
36
40
44
110
120
130
140
150
160
13C (ppm)
8.5
8.0
7.5
7.0
6.5
1H (ppm)


(a)
CB
Y20
Y23
Y53
CG/CD/CE
CE
CG
CD
CZ
HD
HE
(b)
CB
F34
F3
F38
CG/CD/CE/CZ
CG
HZ
HD
HD
(c)
CD2
CG
CE1
HE1
HD2
^{1}H (ppm)
^{13}C (ppm)


CG
CZ2
CE3
CZ3
CH2
CD1
CD2
CE2
HZ2
HD1
HH2
HE3
HZ3
110
115
120
125
130
135
140
13c (ppm)
7.8
7.6
7.4
7.2
7.8
6.8
6.6
6.4
1H (ppm)


(a) Tyrosine
CB
CE
CG
CD
CZ
HD
HE
39.3 ± 3.2
117.7 ± 4.2
128.3 ± 10.3
132.4 ± 5.4
155.4 ± 14.3
(b) Phenylalanine
CB
CZ
CE
CD
CG
HD
HE
HZ
40.0 ± 3.8
129.0 ± 4.3
130.3 ± 6.0
131.1 ± 6.2
137.5 ± 11.0
(c) Histidine
CD2
CG
CE1
HD2
HE1
119.9 ± 5.9
131.3 ± 7.2
137.2 ± 5.9
(d) Tryptophan
CG
CZ2
CE3
CZ3
CH2
CD1
CD2
CE2
HZ2
HD1
HE3
HZ3
HH2
110.0 ± 9.5
114.0 ± 4.7
120.1 ± 5.7
121.1 ± 4.8
123.5 ± 5.2
126.3 ± 4.5
126.8 ± 13.7
137.6 ± 9.9


(a)
CB
Y59
F4
F45
H68
CD2
CE
CZ
CZ
CE
CE
CD
CD
CD
CG
CG
CG
CE1
CZ
40
42
44
46
120
130
140
150
7.8
7.4
7.2
7.0
6.8
1H (ppm)
(b)
CB
F34
Y42
F18
F12
38
42
46
HE3
HD1
CG
H32
CD2
CE
F
F
HH2
CD
CG
CE2
CE1
CG
110
115
120
125
130
135
140
CZ
152
156
160
7.4
7.2
7.0
6.8
1H (ppm)
13C (ppm)

Lysozyme at ^{13}C natural abondance: 1.1 %