COMMUNICATION

ARCHE-NOAH: NMR Supersequence with Five Different CEST Experiments for Studying Protein Conformational Dynamics

Rodrigo Cabrera Allpas,[a] Alexandar L. Hansen [b] and Rafael Brüschweiler [a,b,c*]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

An NMR NOAH-supersequence is presented consisting of five CEST experiments for studying protein backbone and side-chain dynamics by ^{15}N-CEST, carbonyl-^{13}CO-CEST, aromatic-$^{13}C_{ar}$-CEST, $^{13}C_{\alpha}$-CEST, and methyl-$^{13}C_{met}$-CEST. The new sequence acquires these experiments in a fraction of the time required for the individual experiments saving over four days of NMR time per sample.

NMR is the method of choice for studying conformational exchange of biomolecules in solution providing atomic-resolution information of dynamics over timescales ranging from picoseconds to seconds.[1,2] Various NMR experiments exist for the study of chemical exchange, including line shape analysis, $R_{1\rho}$, Carr-Purcell-Meiboom-Gill (CPMG), Chemical Exchange Saturation Transfer (CEST).[3] CEST is a powerful experiment providing both kinetics and thermodynamics parameters of exchange processes (k_{ex}, p_a, p_b) between the ground state *a* and the "invisible" excited state *b*, structural information in terms of their chemical shift differences ($\Delta\Omega$) and laboratory frame relaxation parameters R_1, R_2 of the ground state.[4] For example, CEST derived relaxation parameters can then be used to calculate S^2 order parameters via a lean model-free approach (MFA) with good accuracy.[4] Furthermore, the chemical shifts of the excited states can be used to develop realistic models of the excited state structure.[5]

The most common NMR spin probes for CEST in proteins are ^{15}N and methyl-^{13}C; however, other spin probes such as $^{13}C_{\alpha}$, $^{13}C_{\beta}$, and ^{13}CO have been reported as well.[5–10] These less commonly used NMR probes can give complementary information about the dynamic processes. For example, $^{13}C_{\alpha}$ and ^{13}CO chemical shifts are very sensitive to protein backbone secondary structure.[7,8] The most used probe, ^{15}N does not report the dynamics of proline residues and is more susceptible to H^N exchange with the water solvent, thereby losing dynamic information due to solvent exchange. For the study of aromatic residues, aromatic carbon probes have also been reported, but only for CPMG relaxation dispersion type experiments,[11] despite their facility to form clusters and their important role in protein-protein interaction interfaces.[12] Clearly, there is much conformational exchange information that can be gained for each residue of a protein using all available nuclei as probes. However, CEST experiments are rather time consuming to perform, since they are ran in a pseudo-3D manner for the added resolution, taking a day to a week of experiment time for each type of spin. Hence, measuring the less commonly used spin probes can be prohibitively expensive in terms of spectrometer time, despite the complementary information one can obtain from each of them.

In recent years, NMR by Ordered Acquisition using ^{1}H-detection (NOAH) supersequences have gained popularity for studying small molecules.[13–16] The basic idea behind NOAH is to combine multiple separate pulse sequences (modules) into a supersequence, which employs only a single recovery delay (d_1) throughout the pulse program. Because d_1 is by far the longest time delay in an NMR experiment, a supersequence consisting of two separate pulse sequences with only a single recovery delay instead of two would effectively render the second experiment for "free" provided that other delays of the second experiment are minimal compared to d_1 as is almost always the case. These supersequences are designed to use "leftover" magnetization from previous modules in subsequent modules without the lengthy d_1 delays between modules to recover magnetization via spin-lattice relaxation. It is therefore critical to ensure that suitable pulse-sequence elements are employed that preserve unused magnetization for subsequent modules of the supersequence. Recently, small molecule applications of NOAH included as many as ten experiments, thereby drastically reducing the total experimental time,[17] and NOAH supersequences have been developed for complex mixture analysis in metabolomics.[18–20]

Despite much progress of NOAH for small molecule NMR, only few applications exist for biomacromolecules. Early work of sequential NMR acquisition for proteins has been proposed

[a] *Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, U.S.A.*
[b] *Campus Chemical Instrument Center, The Ohio State University, Columbus, Ohio 43210, U.S.A.*
[c] *Department of Biological Chemistry and Pharmacology, The Ohio State University, Columbus, Ohio 43210, U.S.A.*

† Electronic supplementary information (ESI) available: Details of the NMR methodology used, details of the pulse sequence parameters (standalone and ARCHE-NOAH), side-by-side comparison of standalone vs ARCHE-NOAH modules spectra, and selective pulses parameters.

for protein resonance assignment based on a non-standard assignment strategy.[21,22] The NOAH approach was also applied to drug-protein interactions by combining 2D ^{1}H-^{15}N HSQC for the ^{15}N-labeled protein with a ^{1}H-^{13}C-BIRD-HSQC for the ligand present at natural abundance.[23] Recently, our group proposed a combination of two CEST pulse sequences using the NOAH approach for uniformly ^{15}N, ^{13}C-labeled proteins. The NOAH-(^{15}N/^{13}C)-CEST produced the same dynamics results as two independently acquired ^{15}N-CEST and methyl-^{13}C-CEST experiments, but required only a fraction of the time needed in the traditional approach.[24] A key element of the approach is to use a ^{13}C-start methyl CEST experiment that follows a regular $^{1}H^{N}$ start ^{15}N-CEST. In this way, $^{1}H^{N}$ magnetization is exhausted during the first module while retaining fresh ^{13}C magnetization for the second module. While one would expect some major sensitivity loss when starting with thermal ^{13}C instead of ^{1}H magnetization, ^{1}H composite-pulse decoupling during the extended saturation period T_{ex} typically between 100 ms – 500 ms in the ^{15}N-CEST leads to a heteronuclear Nuclear Overhauser Effect (hetNOE) resulting in a significant enhancement of the ^{13}C spin magnetization at the start of the 2nd module.

Here, we report a new NOAH supersequence, named ARCHE-NOAH for **A**ll-**R**esidue-**CH**emical-**E**xchange-**NOAH**, that measures 5 protein CEST modules in a single experiment with a single recovery delay d_1 (Figure 1). It acquires CEST experiments of nearly all usable probes of amino-acid residues in proteins, namely backbone ^{15}N-CEST, $^{13}C_{\alpha}$-CEST, and ^{13}CO-CEST along with side-chain ^{13}C-aromatic-CEST and ^{13}C-methyl-CEST in a substantially accelerated manner. ARCHE-NOAH uses band-selective pulses for all carbons, thereby carefully preserving the starting ^{13}C magnetization for the subsequent CEST module (see below). NMR experiments were performed on an 850 MHz Bruker Ascend magnet equipped with an Avance III HD console and a triple resonance inverse cryoprobe. In order to assess the performance of the ARCHE-NOAH supersequence, it was compared with the individual modules as standalone ^{1}H-start sequences recorded with the same parameters applied to 1 mM colicin E7 immunity protein Im7 in 97%/3% H_2O/D_2O at 298 K. The resulting spectra were processed using NMRPipe, followed by peak-picking, peak assignment, and cross-peak quantification. CEST profiles were analyzed using the ChemEx software (http://www.github.com/gbouvignies/chemex). More details of the methods and NMR acquisition parameters can be found in ESI†.

The comparison of the sensitivity of the supersequence and NMR time with its standalone ^{1}H-start counterparts is summarized in Table 1. As should be expected from the aforementioned hetNOE enhancement, the ^{13}C-CEST module experiences most enhancement from ^{1}H-decoupling blocks in the preceding sequences. We find differential hetNOEs for the different types of ^{13}C atoms, namely $^{13}C_{met} > {}^{13}C_{\alpha} \cong {}^{13}C_{ar} > CO$, which can be explained by the number of protons attached to the carbon consistent with the literature.[25]

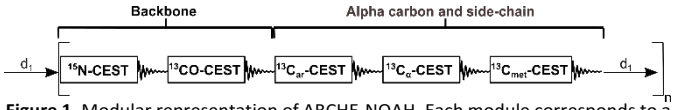

Figure 1. Modular representation of ARCHE-NOAH. Each module corresponds to a CEST experiment for a total of five modules, three for backbone and two for side-chain dynamics information (Car=aromatic carbons, Cα=alpha carbons, Cmet=methyl carbons). One scan is accomplished by using only one single recovery delay (d1) and the supersequence is repeated for n number of scans.

The total experiment time saving of the ARCHE-NOAH sequence is 65% corresponding to almost 100 hours (> 4 days) of instrument time. Shorter experiment times are especially critical for samples with limited lifetime for which ARCHE-NOAH provides an efficient way to acquire five different CEST types. In fact, the acquisition time of the new supersequence is faster than acquiring the two most used CEST sequences, ^{15}N-CEST and $^{13}C_{met}$-CEST, as standalone experiments, while providing three extra CEST experiments that carry unique information, albeit at a lower sensitivity.

One way to compare measurement times and sensitivity is to scale down the signal-to-noise of the 5 standalone experiments to the signal-to-noise of the ARCHE-NOAH experiment and calculate how much experimental time would be needed to match their signal-to-noise. In this case, the time saving is 29%. While in theory this is an accurate way of calculating the time saved by our method, in practice one cannot reduce the experimental time of the standalone experiments to this extent, since each experiment requires a minimal number of total scans. For example, it is not possible to lower the number of scans of the ^{13}CO-CEST by a factor of 10 to obtain the desired experimental time to match the signal-to-noise even with non-uniform-sampling methods. Therefore, in practice the time saving is much closer to 65%.

Table 1. Signal-to-noise ratios of standalone and ARCHE-NOAH experiments and their individual and combined experiment time.

NMR experiment	S/N[a]	Experiment time (h min)	Scaled experiment time (h min)[b]
^{15}N-CEST	1086	32 h 25 min	27 h 52 min
^{13}CO-CEST	299	29 h 17 min	2 h 41 min[c]
$^{13}C_{ar}$-CEST	331	28 h 35 min	5 h 42 min[c]
$^{13}C_{\alpha}$-CEST	454	29 h 59 min	5 h 34 min[c]
$^{13}C_{met}$-CEST	980	29 h 57 min	32 h 46 min
Total experiment time		150 h 13 min	74 h 35 min
NOAH-^{15}N-CEST	1007		
NOAH-^{13}CO-CEST	90		
NOAH-$^{13}C_{ar}$-CEST	148	52 h 59 min	-
NOAH-$^{13}C_{\alpha}$-CEST	196		
NOAH-$^{13}C_{met}$-CEST	1025		

[a] Average signal-to-noise of all non-overlapped signals of the respective sequence. Noise was calculated by taking the median absolute deviation (MAD) of the processed NMRPipe files in an empty region with no signals.

[b] Scaled experiment time (h min) for the standalone NMR experiments obtained by the formula $\left(\frac{S/N_{NOAH}*\sqrt{t_{standalone}}}{S/N_{standalone}}\right)^2$.

[c] Due to minimal sampling requirements for these pseudo-3D CEST experiments, this projected measurement time is not feasible in practice.

Additionally, it is worth nothing that if we were to leave out the least sensitive experiment ^{13}CO-CEST, our time savings after applying the correction done above would be 33% for a ^{15}N-$^{13}C_{ar}$-$^{13}C_{\alpha}$-$^{13}C_{met}$-CEST sequence. If a 2nd CEST block is left out, the corrected time savings would be 33% for ^{15}N-$^{13}C_{\alpha}$-$^{13}C_{met}$-CEST and 36% for a ^{15}N-$^{13}C_{ar}$-$^{13}C_{met}$-CEST. If both ^{13}CO- and $^{13}C_{ar}$-CEST are left out, the resulting time saving for ^{15}N-$^{13}C_{met}$-CEST is 37% (Table 2). Thus, leaving out one or several CEST blocks that have lower sensitivity only results in a modest time gain, while losing the protein dynamics information contained in those experiments. Nevertheless, our sequence is sufficiently flexible that if one is not interested in a particular CEST module, the module can be left out and one still benefits from an overall time savings with the remaining modules. In practice, this means that if one has a uniformly ^{15}N,^{13}C-labeled protein and is planning to acquire a ^{15}N-CEST and ^{13}C-methyl-CEST, there is no actual time gain by *not* acquiring all five CEST experiments with ARCHE-NOAH with the added benefit of obtaining complementary protein dynamic information.

Next, we demonstrate that the CEST profiles we obtained using the NOAH supersequence are essentially identical to the profiles obtained by the corresponding standalone ^{1}H-start sequences. For backbone CEST profiles (Figure 2), we have the ^{15}N-CEST and ^{13}CO-CEST experiments, which form the first and second block of the supersequence. Unlike ^{13}CO-CEST, ^{15}N-CEST does not suffer from signal-to-noise losses as is best visible in the baseline of the profile. Despite the increased noise level of the ^{13}CO-CEST profiles, they accurately match the reference profiles, and the asymmetry of the main dip reflecting the presence of exchange can still be clearly discerned. A comparison between standalone and NOAH of these two experiments can be found in ESI† (Figures S1-S2).

Table 2. Signal-to-noise ratios of standalone and NOAH experiments and their individual and combined experiment time in hours and minutes.

ARCHE-NOAH supersequence[a]	Time savings (%)	Scaled time savings (%)
NOAH-(N-CO-C_{ar}-C_{α}-C_{met})	65%	29%
NOAH-(N-C_{ar}-C_{α}-C_{met})	60%	33%
NOAH-(N-C_{α}-C_{met})	53%	34%
NOAH-(N-C_{ar}-C_{met})	53%	36%
NOAH-(N-C_{met})	39%	37%

[a] Each CEST module on the NOAH supersequence is defined in the following way: N = ^{15}N-CEST, CO = ^{13}CO-CEST, C_{ar} = ^{13}Car-CEST, C_{α} = C_{α}-CEST, and C_{met} = C_{met}-CEST.

The α-carbons, aromatic carbons, and methyl-CEST profiles are shown in Figure 3. Their slightly lower signal-to-noise is less of an issue. However, the $^{13}C_{\alpha}$-CEST, both when part of ARCHE-NOAH and as standalone experiment, has a significant residual water signal that perturbs the baseline of nearby peaks; therefore, we limited the comparison to residues further away from the water resonance (4.7 ppm > ^{1}H chemical shifts and ^{1}H chemical shifts > 4.9 ppm). The largest difference in our profiles is seen in a slight baseline offset of the methyl carbons. This is because of two 180^{o} pulses that perturb them in the preceding $^{13}C_{\alpha}$-CEST module during the INEPT transfers, tipping the aliphatic carbon magnetization to -z. Despite using selective 180^{o} pulses to get the methyl carbon magnetization back to +z shortly after the INEPT transfers, they result in minor ^{13}C-R_1 relaxation rate differences. This difference is about ~0.1 s^{-1} and only occurs if the C_{α}-module is used. The shoulder peaks in all these profiles, even if minor, are however accurately reproduced. Lastly, although not visible in the profiles themselves, a small systematic chemical shift variation of about 0.005 ppm (4.2 Hz) can be seen in the NOAH versus the standalone spectra. This effect, which is due to extra heating during the decoupling in each CEST block in the supersequence, is inconsequential for the analysis. A comparison between standalone and NOAH is given in ESI† (Figures S3-S5).

The majority of the NMR acquisition parameters used to run the ^{1}H-start standalone sequences are the same as those used in the ARCHE-NOAH supersequence. One unavoidable difference between standalone and the NOAH-CEST is the mandatory use of band-selective pulses in the ^{13}CO-, $^{13}C_{ar}$-, and $^{13}C_{\alpha}$-CEST modules. The replacement of regular broadband pulses by selective pulses is necessary to preserve carbon magnetization for the next modules of the supersequence. We used common selective pulses available in Topspin and their implementation can be automated using the WaveMaker app in Topspin 3.6 and above. Additionally, module-specific delays have unique identifiers to improve readability and ease of use and, as suggested in Table 2, the user can specify which of the inner-most CEST modules, i.e. ^{13}CO-, $^{13}C_{ar}$-, or $^{13}C_{\alpha}$-CEST, should be included (or excluded). More information about the selective pulses employed in the ARCHE-NOAH supersequence can be found in Table S1 of the ESI†.

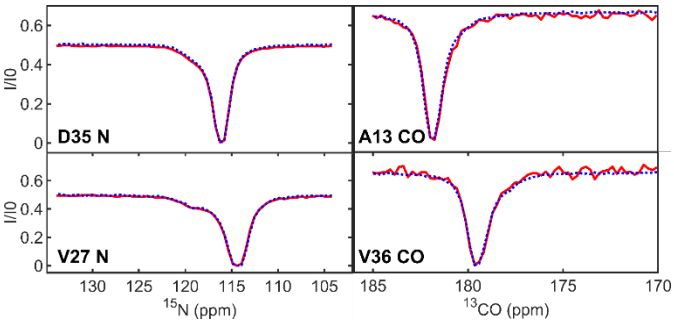

Figure 2. Representative ^{15}N-CEST profiles (left side) for Im7 residues D35 and V27, and ^{13}CO-CEST profiles (right side) for residues A13 and V36 from the supersequence (red, solid line) and the standalone ^{1}H-start sequence (blue, dotted line). Each panel shows residues that undergo two-site exchange with a second minimum, "shoulder" feature, or asymmetry in the profile.

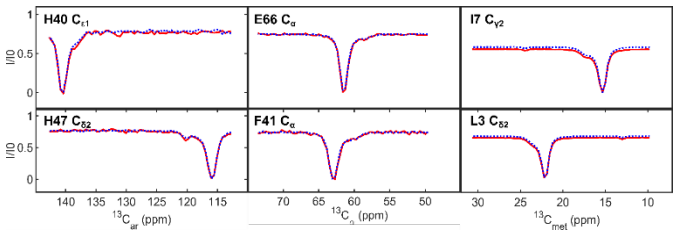

Figure 3. Representative $^{13}C_{ar}$-CEST profiles (left side) for Im7 residues H40 and H47, $^{13}C_\alpha$-CEST profiles (middle) for Im7 residues E66 and F41, and $^{13}C_{met}$-CEST profiles (right side) for residues I7 and L3 from the supersequence (red, solid line) and the standalone ^{1}H-start sequence (blue, dotted line). Each panel shows residues that undergo two-site exchange with a second minimum, "shoulder" feature, or asymmetry in the profile.

The ARCHE-NOAH can be applied to obtain chemical shifts of an "invisible" excited state of five NMR probes in one single supersequence for a uniformly ^{15}N,^{13}C-labeled sample. Three of these probes (^{13}CO, $^{13}C_{ar}$, or $^{13}C_\alpha$) have been seldomly used in the past due to the additional spectrometer time requirements, whereas the new NOAH sequence collects them for "free". Moreover, the total time of the ARCHE-NOAH with all five modules (53 h) is less than acquiring the ^{15}N- and $^{13}C_{met}$-CEST experiments as standalone ^{1}H-start experiments (62 h). The chemical shifts from the probes obtained from our sequence provide a unique window into the structure of excited states of proteins.[26] Although our method focused on covering most NMR probes encountered in amino-acid residues, a variation of our sequence with the four CEST modules (^{15}N, ^{13}CO, $^{13}C_{ar}$, $^{13}C_\alpha$) can be directly applied to uniformly ^{15}N,^{13}C-labeled RNA/DNA samples. In this case, the ^{15}N, ^{13}CO, $^{13}C_{ar}$ modules report on the nitrogenous base and the $^{13}C_\alpha$ module covers the ribose moiety offering a comprehensive view of their dynamics.[27,28]

In summary, the ARCHE-NOAH approach introduced here merges five different CEST modules (^{15}N-CEST, ^{13}CO-CEST, ^{13}C-aromatic-CEST, $^{13}C_\alpha$-CEST, ^{13}C-methyl-CEST) into a single supersequence with only a single recovery delay. The reported time savings are about 65%, which amounts to a saving of about 100 hours of high-field NMR instrument time. Hence, it is most advantageous to use ARCHE-NOAH when experiment time is of the essence, for example due to protein stability, for the screening of protein dynamics in active enzyme-ligand systems, or when studying a protein under different experimental conditions, such as variable temperature or pressure. The only requirement is a uniformly ^{15}N,^{13}C-labeled protein sample, which can be the same sample used for backbone and side-chain resonance assignments and can be produced using a well-established and most economic labelling strategy. Moreover, the band-selective pulses used in the supersequence are readily available on standard Bruker spectrometers. The resulting spectra and CEST profiles compare well with the standalone experiments and are void of artifacts. The ARCHE-NOAH sequence facilitates the exploration also of less commonly studied atomic sites in amino-acid residues, providing novel complementary information on protein dynamics never studied before. It offers the comprehensive analysis of biomolecular dynamics and exchange processes that one would otherwise miss or simply choose not to pursue due to instrument time constraints.

We thank Ms. Xinyao Xiang for providing the Im7 sample used in this work. This work was supported by the U.S. National Science Foundation (grant MCB-2103637 to R.B.). All NMR experiments were performed at the Campus Chemical Instrument Center NMR facility at Ohio State University. The ARCHE-NOAH pulse sequence and AU programs used for processing can be found at https://github.com/RCabreraAllpas.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1 I. R. Kleckner and M. P. Foster, Biochim. Biophys. Acta BBA - Proteins Proteomics, 2011, 1814, 942–968.
2 T. R. Alderson and L. E. Kay, Cell, 2021, 184, 577–595.
3 P. Vallurupalli, A. Sekhar, T. Yuwen and L. E. Kay, J. Biomol. NMR, 2017, 67, 243–271.
4 Y. Gu, A. L. Hansen, Y. Peng and R. Brüschweiler, Angew. Chem. Int. Ed., 2016, 55, 3117–3119.
5 P. Vallurupalli, G. Bouvignies and L. E. Kay, J. Am. Chem. Soc., 2012, 134, 8148–8161.
6 G. Bouvignies and L. E. Kay, J. Biomol. NMR, 2012, 53, 303–310.
7 A. L. Hansen, G. Bouvignies and L. E. Kay, J. Biomol. NMR, 2013, 55, 279–289.
8 P. Vallurupalli and L. E. Kay, Angew. Chem., 2013, 125, 4250–4253.
9 Y. Zhou and D. Yang, J. Biomol. NMR, 2015, 61, 89–94.
10 D. Long, A. Sekhar and L. E. Kay, J. Biomol. NMR, 2014, 60, 203–208.
11 U. Weininger, M. Respondek and M. Akke, J. Biomol. NMR, 2012, 54, 9–14.
12 E. Lanzarotti, L. A. Defelipe, M. A. Marti and A. G. Turjanski, J. Cheminformatics, 2020, 12, 30.
13 Ē. Kupče and T. D. W. Claridge, Angew. Chem. Int. Ed., 2017, 56, 11779–11783.
14 Ē. Kupče and T. D. W. Claridge, J. Magn. Reson., 2019, 307, 106568.
15 T. D. W. Claridge, M. Mayzel and Ē. Kupče, Magn. Reson. Chem., 2019, 57, 946–952.
16 J. R. J. Yong, A. L. Hansen, Ē. Kupče and T. D. W. Claridge, J. Magn. Reson., 2021, 329, 107027.
17 Ē. Kupče, J. R. J. Yong, G. Widmalm and T. D. W. Claridge, JACS Au, 2021, 1, 1892–1897.
18 A. L. Hansen, E. Kupče, D.-W. Li, L. Bruschweiler-Li, C. Wang and R. Brüschweiler, Anal. Chem., 2021, 93, 6112–6119.
19 F. Tang and E. Hatzakis, Anal. Chem., 2020, 92, 11177–11185.
20 F. Tang, H. S. Green, S. C. Wang and E. Hatzakis, Molecules, 2021, 26, 310.
21 P. Bellstedt, Y. Ihle, C. Wiedemann, A. Kirschstein, C. Herbst, M. Görlach and R. Ramachandran, Sci. Rep., 2014, 4, 4490.
22 C. Wiedemann, P. Bellstedt, A. Kirschstein, S. Häfner, C. Herbst, M. Görlach and R. Ramachandran, J. Magn. Reson., 2014, 239, 23–28.
23 V. M. R. Kakita, K. Rachineni, M. Bopardikar and R. V. Hosur, J. Magn. Reson., 2018, 297, 108–112.
24 R. C. Allpas, A. L. Hansen and R. Brüschweiler, Chem. Commun., 2022, 58, 9258–9261.
25 P. M. Macdonald and R. Soong, J. Magn. Reson., 2007, 188, 1–9.
26 A. J. Baldwin and L. E. Kay, Nat. Chem. Biol., 2009, 5, 808–814.
27 B. Zhao, S. L. Guffy, B. Williams and Q. Zhang, Nat. Chem. Biol., 2017, 13, 968–974.
28 B. Zhao, J. T. Baisden and Q. Zhang, J. Magn. Reson., 2020, 310, 106642.