

Absolut minimales Sampling in der hochdimensionalen NMR-Spektroskopie

Alexander L. Hansen und Rafael Brüschweiler*

Abstract: Standardmäßige dreidimensionale (3D-)Fourier-Transformations(FT)-NMR-Experimente von molekularen Systemen erfordern häufig lange Messzeiten, die durch das umfangreiche Sampling entlang der indirekten Zeitbereiche verursacht werden, damit eine hinreichende spektrale Auflösung erhalten werden kann. In den letzten Jahren wurde eine Fülle von alternativen Sampling-Methoden vorgeschlagen, um diesen Engpass zu vermeiden. Aufgrund ihrer algorithmischen Komplexität für vorgegebene NMR-Proben und Experimente ist es jedoch häufig schwierig, die minimale Abtastinformation und die damit verbundene minimale Messzeit zu bestimmen. Hier wird die Methode des absoluten minimalen Samplings (AMS) vorgestellt, die auf gängige 3D-NMR-Experimente angewendet werden kann. Für die Proteine Ubiquitin und Arginin-Kinase wird gezeigt, wie man für das weitverbreitete 3D-HNCO-Experiment präzise Kohlenstofffrequenzen mithilfe eines einzigen Zeitinkrements erhält, während für andere, z. B. das 3D-HN(CA)CO-Experiment, alle relevanten Informationen in nur sechs Schritten erhältlich sind, wodurch eine Beschleunigung um einen Faktor 7–50 erreicht wird.

Der große Informationsgehalt moderner mehrdimensionaler NMR-Spektren ist der Grund für ihre weite Verbreitung in der chemischen und biochemischen Forschung. Die Standard-Akquisition von NMR-Experimenten in drei (oder mehr) Dimensionen ist jedoch zeitintensiv und erfordert Messzeiten in der Größenordnung von einem Tag oder mehr. Beispielsweise werden dreidimensionale (3D-)NMR-Experimente von einheitlich ^{15}N , ^{13}C -markierten Proteinproben traditionell gemessen, indem man die Evolutionszeiten t_1 und t_2 entlang den ^{13}C - und ^{15}N -Dimensionen mit N_1 und N_2 komplexen Punkten unabhängig erfasst.^[1,2] Dies führt zu einer Gesamtmesszeit, die mit $N_1 \times N_2$, anwächst, wobei N_1 und N_2 in der Regel auf 32 bis 64 Schritte beschränkt sind, sodass die spektrale Auflösung nach der Fourier-Transformation (FT) entlang den

beiden indirekten Dimensionen ausreichend hoch ist (in der Regel in der Größenordnung von einigen zehn Hz).

In den letzten Jahren wurde eine Reihe von Methoden eingeführt, mit denen durch gezielte Reduzierung der Gesamtzahl von Inkrementen die Messzeit verkürzt werden soll.^[3] Die meisten Methoden gehören zu einer von zwei Sampling-Strategien: Bei der radialen Abtastung werden t_1 und t_2 gleichzeitig erhöht, was die Messung von einer oder mehreren geneigten Frequenzebenen ermöglicht und dadurch die Rekonstruktion der mehrdimensionalen Spektren ermöglicht,^[4,5] während beim nicht-einheitlichen Sampling (NUS) bestimmte t_1 - und t_2 -Punkte zufällig weggelassen werden, wobei vorwiegend Datenpunkte mit relativ kurzen Entwicklungszeiten gemessen werden, da diese im Durchschnitt ein höheres Signal-zu-Rausch-Verhältnis aufweisen. Die beiden Sampling-Methoden haben komplementäre Vorteile: Beim radialen Sampling wird eine modifizierte Form des Standard-Fast-Fourier-Transformations-Verfahrens angewendet, allerdings ist die Rekonstruktion des vollständigen 3D-Spektrums anfällig für Artefakte. NUS ermöglicht keine Verarbeitung mittels Standard-Fourier-Transformation und verwendet eine Regularisierungsbedingung für die spektrale Rekonstruktion, z. B. die maximale Entropie^[6,7] oder das Compressed Sensing.^[8,9] Es erfordert ein Gleichgewicht zwischen der Konsistenz des rekonstruierten Spektrums mit den experimentellen Daten und der erreichbaren Einfachheit des schlussendlichen Spektrums. Die beiden Stichprobenstrategien können für die hochwertige Rekonstruktion von mehrdimensionalen Spektren auch kombiniert werden.^[10]

Ein weiterer Ansatz ist das Angleichen in der Zeitdomäne, die ursprünglich in der linearen Vorhersage (linear prediction, LP) verwendet wurde, wobei ein verkürztes Zeitsignal verlängert wird, sodass es einer normalen Fourier-Transformation unterzogen werden kann.^[11] Obwohl die LP die spektrale Auflösung verbessert und Trunkierungsartefakte minimiert, ist sie typischerweise auf eine zweifache Verringerung der Messzeit pro indirekte Dimension begrenzt. In jüngster Zeit wurden auch andere Verfahren zur Zeitbereichsanpassung vorgeschlagen, wie die Filterdiagonalisierungsmethode^[12,13] oder CRAFT für die Erstellung von Tabellen mit Frequenzen, Linienbreiten und Amplituden von Signalen in 1D- und 2D-NMR-Datensätzen^[14,15] sowie für eine genauere Extraktion von Kreuzsignalamplituden von 2D-HSQC-Spektren.^[16] Für alle diese Verfahren ist die minimale Anzahl der benötigten Zeitschritte, um alle relevanten spektralen Informationen zu erhalten, unbekannt.

In der Praxis sind die interessantesten spektroskopischen Parameter in höherdimensionalen Spektren häufig die chemischen Verschiebungen von Kernen, die dem gleichen Spinsystem angehören, beispielsweise die Kerne von Atomen

[*] Prof. R. Brüschweiler

Department of Chemistry and Biochemistry
The Ohio State University
100 West 18th Avenue, Columbus, OH 43210 (USA)
E-Mail: bruschweiler.1@osu.edu

Dr. A. L. Hansen, Prof. R. Brüschweiler
Campus Chemical Instrument Center
The Ohio State University
460 W. 12th Avenue, Columbus, OH 43210 (USA)

Prof. R. Brüschweiler
Department of Biological Chemistry and Pharmacology
The Ohio State University
1645 Neil Avenue, Columbus, OH 43210 (USA)

Hintergrundinformationen zu diesem Beitrag sind unter:
<http://dx.doi.org/10.1002/ange.201608048> zu finden.

in Peptidbindungen des Proteinrückgrats. Für einige Experimente sind solche Informationen mit minimaler Messzeit erhältlich. Zum Beispiel kann in einem 3D-HNCO-Spektrum^[17] eines Proteins, in dem eine einzige C'-Carbonyl-Resonanz zu einem ausgewählten ¹⁵N,¹H-Frequenzpaar gehört, die Carbonyl-Resonanzfrequenz ω_c ausgehend von einem einzigen komplexen Zeitpunkt $t_1 > 0$ bestimmt werden [Gl. (1)].^[18] Dabei sind $C(t_1)$ und $S(t_1)$ [Gl. (2)] die Cosinus- bzw. Sinusmodulierten experimentellen Datenpunkte entlang der indirekten Zeitdimension t_1 des interessierenden ¹⁵N,¹H-Frequenzpaars, und $\arctan(y,x)$ ist die Doppelargument-Variante der Arcustangens-Funktion (die auch als atan2 bekannt ist), die empfindlich gegenüber den Vorzeichen der Argumente ist, wobei sich die transversalen Relaxationseffekte $A \exp(-R_2 t_1)$ mit der Zerfallsrate R_2 gegenseitig aufheben und somit ω_c als einzige Unbekannte verbleibt.

$$\omega_c = \arctan\{S(t_1), C(t_1)\}/t_1 \quad (1)$$

$$C(t_1) = A \exp\{-R_2 t_1\} \cos(\omega_c t_1) \quad (2)$$

$$S(t_1) = A \exp\{-R_2 t_1\} \sin(\omega_c t_1)$$

Die Anwendung von Gleichung (1) mit dem zweiten komplexen t_1 -Inkrement entlang der C'-Dimension eines HNCO-Experiments ist eine Erweiterung des SPEED-Verfahrens^[18] auf 3D-NMR-Daten, aber ohne den Bedarf eines ¹³C-Carbonyl-Referenzspektrums. Ergebnisse für Ubiquitin (8.5 kDa, 76 Aminosäuren) und Arginin-Kinase (AK, 42 kDa, 357 Aminosäuren)^[19] zeigen Resonanzpositionen in guter Übereinstimmung mit dem entsprechenden 3D-FT-Spektrum mit einer Wurzel aus der mittleren quadratischen Abweichung (RMSE) von $\Delta\delta = 0.18$ und 0.34 ppm (Abbildung 1A,C). Für Ubiquitin ist Ile3 ein Ausreißer infolge starker Signalüberlappung bei seinen ¹⁵N,¹H-Frequenzen (Abbildung S1 der Hintergrundinformationen). Es ist erwähnenswert, dass die FT eines einzigen Zeitbereichpunktes keine nützlichen Frequenzinformationen liefern würde.

Die Genauigkeit der experimentellen chemischen Verschiebungen kann direkt mithilfe einer nichtlinearen Kleinstste-Quadrat-Angleichung in der Zeitdomäne der Summe von Cosinus- und Sinusfunktionen, $C(t_1)$ und $S(t_1)$, an mehrere t_1 -Inkremente verbessert werden. Dies ermöglicht auch die Bestimmung von mehreren Resonanzfrequenzen, die zum selben ¹⁵N,¹H-Paar gehören [Gl. (3)]. Dabei sind im Falle der

$$C(t_1) = \sum_{k=1}^M A_k \exp(-R_2 t_1) \cos(\omega_{c,k} t_1) \quad (3)$$

$$S(t_1) = \sum_{k=1}^M A_k \exp(-R_2 t_1) \sin(\omega_{c,k} t_1)$$

gleichförmigen Abtastung $t_1 = n \Delta t_1$ ($n = 0, \dots, N-1$) und die Abtastrate die Nyquist-Frequenz $1/\Delta t_1$ (spektrale Breite). Es gibt dann nur zwei Angleichungsparameter pro Resonanz k , nämlich die Amplitude A_k und die Frequenz $\omega_{c,k}$. Für kleine N -Werte, d.h. kurze maximale Evolutionszeiten, ist die transversale Relaxation (R_2) in guter Näherung vernachlässigbar. Für größere N -Werte und längere Evolutionszeiten oder größere Proteine mit einer raschen R_2 -Relaxation kann

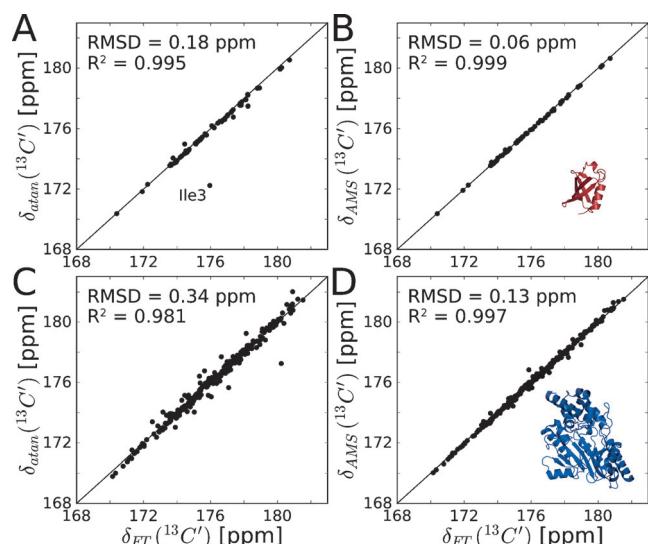


Abbildung 1. Genauigkeit des Arcustangens-Ansatzes (A,C) und der AMS-Methode (B,D) für chemische Verschiebungen. Resultate für 3D-HNCO-Datensätze von ¹³C,¹⁵N-markiertem Ubiquitin (A,B) und Arginin-Kinase (C,D). In (A) wurde der Ausreißer (Ile3) von den RMSD- und R^2 -Berechnungen ausgeschlossen. Carbonyl-Verschiebungen in (A,C) wurden aus Gleichung (1) nur mit dem zweiten komplexen t_1 -Zeitpunkt des HNCO-Experimentes ermittelt. Die AMS-abgeleiteten Verschiebungen in (B,D) entsprechen den besten Lösungen unter Verwendung der Gleichung (3) mit den ersten 4 komplexen t_1 -Punkten des HNCO-Experimentes. Chemische Referenzverschiebungen δ_{FT} (x-Achse) wurden mittels Peak-Picking aus den konventionell mit FT verarbeiteten 3D-HNCO-Experimenten bestimmt, die mit 40 und 44 komplexen t_1 -Punkten für Ubiquitin und Arginin-Kinase gemessen wurden.

eine einheitliche Wahl von $R_2 > 0$ helfen, die Angleichungsergebnisse zu verbessern. Im Prinzip ist der kleinstmögliche Wert für N gleich der Anzahl von Resonanzen (M) und deshalb der absoluten minimalen Abtastinformation (AMS). In der Praxis und in der Gegenwart von Rauschen verbessert die Verwendung eines etwas größeren $N > M$ die Robustheit und Genauigkeit der angepassten Frequenzen und Amplituden.

Abbildung 1B,D zeigt AMS-3D-HNCO-Daten unter Verwendung von Gleichung (3) für $N=4$ t_1 -Schritte, die deutlich die Genauigkeit der chemischen Verschiebung auf $\delta = 0.06$ ppm (9 Hz) für Ubiquitin und $\delta = 0.13$ ppm (26 Hz) für AK-C' verbessern. Dies entspricht einer Genauigkeit der chemischen Verschiebung, die vergleichbar mit der natürlichen C'-Linienbreite ist (siehe Hintergrundinformationen). Die Dauer des Experimentes für die AMS-Daten von AK beträgt nur 2.5 h gegenüber 28 h für den kompletten 3D-Datensatz, was einer Beschleunigung um den Faktor 11 entspricht.

Überraschenderweise ermöglicht im Falle von mehreren Resonanzen die Methode der nichtlinearen Anpassung die genaue und robuste Bestimmung der dominanten Frequenz mit wenig Einfluss durch das Vorhandensein einer zweiten, schwächeren Resonanz. Dies ist in Abbildung 2 gezeigt, in der die Güte des Angleichungsparameters χ^2 als Funktion der Frequenz für unterschiedliche Abtastpunkte N aufgetragen ist. Das globale χ^2 -Minimum befindet sich an der richtigen Position ($\delta = 172$ ppm) mit einem zweiten lokalen Minimum

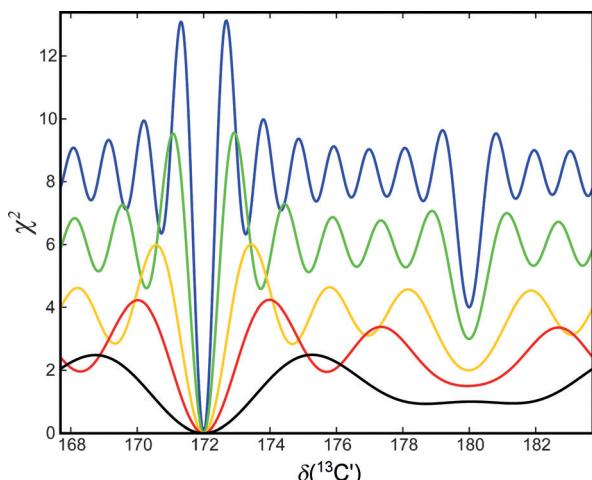


Abbildung 2. Angleichungsgüte χ^2 als Profil bei der Rekonstruktion einer einzigen chemischen Verschiebung (in ppm) von Zeitbereichsdaten, die der Summe von zwei unterschiedlichen Frequenzen entsprechen. Simulierte Daten wurden mit zwei Frequenzen bei $\delta = 172$ und 180 ppm mit relativen Populationen von 0.75 und 0.25 erzeugt und mittels Gleichung (2) angeglichen. Die Anzahl der in den Angleichungen genutzten komplexen Punkten sind 4 (schwarz), 6 (rot), 8 (gelb), 12 (grün) und 16 (blau). Wichtig ist, dass die Lage des χ^2 -Minimums des dominanten Peaks unabhängig von der Anzahl der verwendeten Punkte ist.

bei der Frequenz der Nebenresonanz ($\delta = 180$ ppm). Bei größeren N -Werten ist das Minimum schmäler und besser definiert, und in Abhängigkeit vom Signal-zu-Rausch-Verhältnis ermöglicht es eine genauere Anpassung der spektralen Parameter. Die mittleren Signal-zu-Rausch-Verhältnisse der ersten $^{15}\text{N}, ^1\text{H}$ -Ebenen der HNCO-Experimente von Ubiquitin und AK waren etwa 200:1 und 40:1.

Der AMS-Ansatz ermöglicht eine genaue Bestimmung der Resonanzpositionen und Amplituden auch für 3D-Experimente, bei denen mehr als ein ^{13}C -Kreuzpeak pro Rückgrat- $^{15}\text{N}, ^1\text{H}$ -Paar vorhanden ist, wie es beim HN(CA)CO-Experiment der Fall ist.^[20] Abbildung 3 zeigt für ausgewählte Aminosäurereste von Arginin-Kinase Querschnitte durch die vollständigen HN(CA)CO- und HNCO-3D-FT-Spektren zusammen mit den AMS-Resultaten (rote Balken). Die Ergebnisse zeigen, dass sowohl die Peak-Positionen als auch deren Amplituden sehr gut den FT-Ergebnissen entsprechen, die eine >7-fach längere Messzeiten erforderten. Dies gilt auch für HN(CA)CO-Spuren, die mehrere Resonanzen aufwiesen (Abbildung 3D,E,F), trotz eines mittleren Signal-zu-Rausch-Verhältnisses der ersten $^{15}\text{N}, ^1\text{H}$ -Ebene von nur etwa 15:1. Der Vergleich für alle zugeordneten Reste von AK (siehe Hintergrundinformationen) zeigt, dass der AMS-Ansatz Kreuzsignallisten mit hoher Genauigkeit und Zuverlässigkeit in einer optimalen Menge an Zeit zu produzieren vermag.

Der AMS-Ansatz passt direkt Resonanzfrequenzen und deren Amplituden aus Zeitbereichssignalen entlang einer indirekten Dimension an, beispielsweise die ^{13}C -Dimension, ohne eine Fourier-Transformation zu erfordern. Somit ist AMS nicht an das Unsicherheitsprinzip der FT gebunden. Beispielsweise beträgt für das HNCO-Experiment von AK

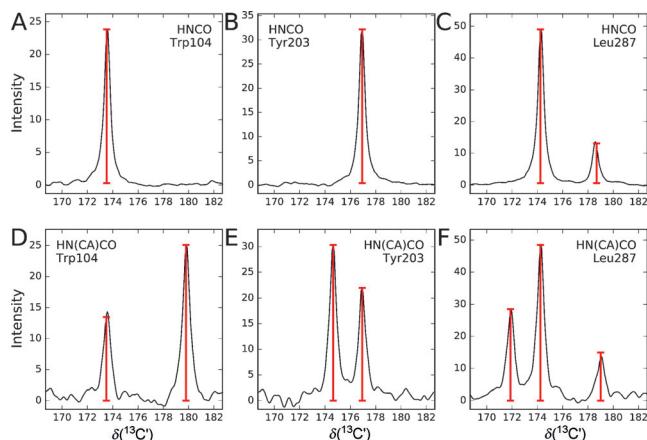


Abbildung 3. Carbonyl-FT-Querprofile und AMS-Ergebnisse von HNCO- und HN(CA)CO-Experimenten für Arginin-Kinase. In Schwarz sind die traditionellen FT-Querschnitte (mit 44 t_1 -Punkten) gezeigt, während die roten Balken den mittels AMS abgeleiteten chemischen Verschiebungen mit nur den ersten 4 oder 6 komplexen Punkten für die HNCO- und HN(CA)CO-Spektren entsprechen. Die Höhen der roten Balken geben die durch AMS ermittelten Resonanzamplituden A_k wieder [Gl. (3)].

mit $N=4$ die digitale Auflösung des Fourier-Spektrums entlang der C'-Dimension 706 Hz ($1/4$ der spektralen Breite). Es ist bemerkenswert, dass der mittlere Fehler der C'-Frequenzen des AMS nur 26 Hz oder 3.7% der digitalen Auflösung ist, was die hohe Wirksamkeit dieser Methode belegt. In den meisten 3D-NMR-Experimenten ist die Anzahl der Resonanzen entlang einer Achse klein, selbst für große Proteine, was die Menge der notwendigen Datenpunkte in der Zeitdomäne drastisch reduziert, die typischerweise weit unter 10 Inkrementen liegt. Dies macht die AMS-Methode besonders vielversprechend für mehrdimensionale Experimente, die weniger durch Empfindlichkeit begrenzt sind als durch die erforderliche Gesamtzeit für das hinreichende Abtasten der indirekten Dimensionen, was für die meisten Resonanzzuordnungsexperimente der Fall ist. Wie in CRAFT ist das Ergebnis des AMS eine Reihe von Frequenzen und Amplituden und nicht ein herkömmliches FT-Spektrum.

Ziel des AMS-Verfahrens ist die Bestimmung der 3D-Kreuzsignal-Positionen mit der geringstmöglichen Anzahl von Schritten und somit einer minimalen Menge an Zeit und Ressourcen. AMS erfordert nicht die Verwendung einer komplexen Regularisierung, Projektion oder eines komplizierten Rekonstruktionsschemas und kann auf jede beliebige Standard-NMR-Pulssequenz angewendet werden. Es gibt keine Beschränkungen bezüglich Computer-Festplatte und Speicherbeschränkung für AMS, da das komplette hochaufgelöste Spektrum nicht explizit rekonstruiert wird. Es ist möglich, eine AMS-Angleichung durchzuführen, während das NMR-Experiment noch im Gange ist. Sobald die Güte der Anpassung für die Zeitdomänensignale entlang der Dimension(en) von Interesse ausreichend hoch ist, kann das NMR-Experiment automatisch angehalten werden, was nützlich sein kann, wenn beispielsweise mehrere 3D-Pulssequenzen in der Warteschlange sind, möglicherweise in Kombination mit einem automatischen Probenwechsler, um eine

maximale Anzahl von Experimenten oder Proben in einem minimalen Zeitraum zu messen. Als Alternative kann für Proteine mit niedriger Konzentration die Betonung auf die maximale Empfindlichkeit anstelle der Auflösung verschoben werden.

AMS wurde für diejenige Art von 3D-NMR-Experimenten demonstriert, die routinemäßig in vielen Laboratorien für Proteinrückgrat-Resonanzzuordnungen verwendet werden und die kumulativ einen signifikanten Anteil der Gesamtzeit auf Hochfeld-NMR-Systemen ausmachen. Für diese Art von Experimenten bringt AMS eine vorteilhafte NMR-Zeitersparnis gegenüber der Standard-3D-FT ebenso wie gegenüber alternativen Abtastverfahren mit sich.

Danksagung

Diese Arbeit wurde von der National Science Foundation (MCB-1360966) und den NIH (5R01GM077643 und 2R01GM066041) unterstützt.

Stichwörter: Absolut minimale Probenahme · Nichtlineare Angleichung im Zeitbereich · NMR-Spektroskopie · Proteinstrukturen

Zitierweise: *Angew. Chem. Int. Ed.* **2016**, *55*, 14169–14172
Angew. Chem. **2016**, *128*, 14376–14379

- [1] A. Bax, S. Grzesiek, *Acc. Chem. Res.* **1993**, *26*, 131–138.
- [2] M. Sattler, J. Schleucher, C. Griesinger, *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 93–158.
- [3] K. Kazimierczuk, V. Orekhov, *Magn. Reson. Chem.* **2015**, *53*, 921–926.

- [4] S. Kim, T. Szyperski, *J. Am. Chem. Soc.* **2003**, *125*, 1385–1393.
- [5] E. Kupče, R. Freeman, *J. Am. Chem. Soc.* **2003**, *125*, 13958–13959.
- [6] J. C. Hoch, M. W. Maciejewski, M. Mobli, A. D. Schuyler, A. S. Stern, *Acc. Chem. Res.* **2014**, *47*, 708–717.
- [7] S. G. Hyberts, K. Takeuchi, G. Wagner, *J. Am. Chem. Soc.* **2010**, *132*, 2145–2147.
- [8] K. Kazimierczuk, V. Y. Orekhov, *Angew. Chem. Int. Ed.* **2011**, *50*, 5556–5559; *Angew. Chem.* **2011**, *123*, 5670–5673.
- [9] D. J. Holland, M. J. Bostock, L. F. Gladden, D. Nietlispach, *Angew. Chem. Int. Ed.* **2011**, *50*, 6548–6551; *Angew. Chem.* **2011**, *123*, 6678–6681; Übersichtsartikel: D. J. Holland, L. F. Gladden, *Angew. Chem. Int. Ed.* **2014**, *53*, 13330–13340; *Angew. Chem.* **2014**, *126*, 13546–13557.
- [10] H. Hassanieh, M. Mayzel, L. Shi, D. Katabi, V. Y. Orekhov, *J. Biomol. NMR* **2015**, *63*, 9–19.
- [11] J. J. Led, H. Gesmar, *Chem. Rev.* **1991**, *91*, 1413–1426.
- [12] V. A. Mandelsham, *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *38*, 159–196.
- [13] J. Chen, V. Mandelsham, A. Shaka, *J. Magn. Reson.* **2000**, *146*, 363–368.
- [14] K. Krishnamurthy, *Magn. Reson. Chem.* **2013**, *51*, 821–829.
- [15] K. Krishnamurthy, A. M. Seftel, D. J. Russell, *Magn. Reson. Chem.* **2016**, DOI: 10.1002/mrc.4449.
- [16] R. A. Chylla, K. Hu, J. J. Ellinger, J. L. Markley, *Anal. Chem.* **2011**, *83*, 4871–4880.
- [17] L. E. Kay, M. Ikura, R. Tschudin, A. Bax, *J. Magn. Reson.* **1990**, *89*, 496–514.
- [18] E. Kupče, R. Freeman, *Magn. Reson. Chem.* **2007**, *45*, 711–713.
- [19] X. Niu, L. Bruschweiler-Li, O. Davulcu, J. J. Skalicky, R. Bruschweiler, M. S. Chapman, *J. Mol. Biol.* **2011**, *405*, 479–496.
- [20] R. T. Clubb, V. Thanabal, G. Wagner, *J. Magn. Reson.* **1992**, *97*, 213–217.

Eingegangen am 17. August 2016
Online veröffentlicht am 10. Oktober 2016