Journal of the Korean Magnetic Resonance Society **2020**, *24*, 77-85
DOI 10.6564/JKMRS.2020.24.3.077

Stable Isotope Labeling of Proteins in Mammalian Cells

KyungRyun Lee and Jung Ho Lee*

Department of Chemistry, Seoul National University, Seoul 08826, Republic of Korea

Received Sep 10, 2020; Revised Sep 18, 2020; Accepted Sep 18, 2020

Abstract Stable isotope enrichment in proteins is necessary for high-resolution nuclear magnetic resonance (NMR) experiments. Although methods for ^{13}C, ^{15}N and ^{2}H-enrichment in prokaryotic cells are well established, full processing and correct folding of complex protein systems require higher organisms as the expression host. In the present study, we review recent efforts to enrich stable isotopes in mammalian cells for protein NMR studies.

Keywords Mammalian expression, stable isotope labeling, In cell NMR

Introduction

Detailed NMR analysis of biomolecules is facilitated by ^{13}C and ^{15}N frequency labeling. As the ^{13}C, ^{15}N and ^{2}H natural abundance is very low, it is crucial to enrich these stable isotopes during recombinant protein expression. One of the most widely used microorganism for protein expression is *Escherichia coli*, which is superior to other expression systems because of its fast growth rate, high cell density, inexpensive growth media and close to 100% isotope labeling efficiency.[1,2] However, complex and large human proteins often require eukaryotic chaperones and post-translational modifications (PTMs), which include phosphorylation, glycosylation and disulfide bridges, to correctly fold into their fully active forms.[3] Therefore, yeast, insect, plant and mammalian cells are also utilized as protein expression systems.

Although difficult, it is essential to select the right expression system for protein isotope enrichment. Regarding expression levels, there is an optimal host system for each protein.[4] For example, the protein yield of human glutamic acid decarboxylase (GAD65) was 4.5 mg (*E. coli*), 3.5 mg (yeast), 50 mg (insect) and 1.7 mg (mammalian) per liter of cell culture.[5-8] Thus, eukaryotic expression helps in GAD65 production, although it is not strictly required. The protein yield of human interleukin-6 (IL-6) was 7.5 g (*E. coli*), 280 mg (yeast) and 1 mg (insect) per liter of cell culture.[9-11] Clearly, *E. coli* provides the highest expression level for human IL-6. Generally, the cost of cell culture increases, while the efficiency of isotope incorporation decreases as we go to higher organisms.[12]

A question arises in which cases to use mammalian expression system. Interestingly, certain proteins are only expressed in mammalian cells. For example, human T-synthase was not expressed in insect cells, but was expressed in mammalian cells because only the latter possess Cosmc chaperones.[13] Moreover, the function of human proteins can be fine-tuned by PTMs provided by mammalian cells.[14] For instance, a recombinant FVIII-Fc fusion protein was successfully expressed in HEK293 cells, while retaining its full activity with native PTMs.[15] Therefore, majority of recombinant biopharmaceutical proteins are based on mammalian expression systems. Protein data bank (PDB) statistics from the last 10 years reveal a clear

* Address correspondence to: **Jung Ho Lee,** Department of Chemistry, Seoul National University, Gwanak-ro 1, Gwanak-gu, Seoul, 08826, Republic of Korea, Tel: 82-2-880-4363; E-mail: jungho.lee@snu.ac.kr

increasing trend in the use of mammalian expression systems (Figure 1). This trend reflects the increasing interest toward complex proteins that require advanced PTMs and cellular machineries for correct folding.[16]

To study these complex proteins using NMR, it is necessary to optimize the protein yield and the efficiency of stable isotope incorporation in mammalian cells. Several attempts have been made to efficiently incorporate ^{15}N and ^{13}C isotopes in the proteins of interest in mammalian cells. This review summarizes these trials and provides perspectives on stable isotope labeling and *in cell* NMR.

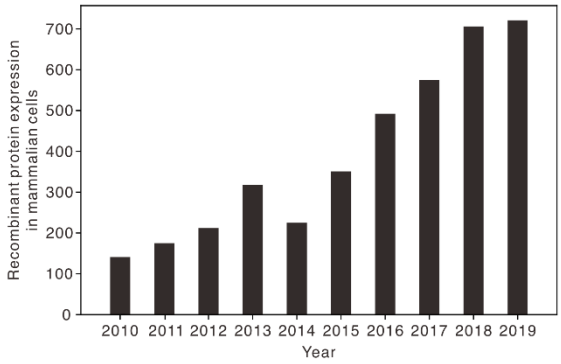

Figure 1. Annual incidences of recombinant protein expression in mammalian cells for structural studies. Data used in the histogram is from the protein data bank (PDB).[17]

Isotope Incorporation Efficiency

The first uniformly isotope-labeled protein was urokinase. It was expressed in Sp2/0 cells using hydrolysate from bacteria or algae, which was grown in $^{15}NH_4Cl$ and ^{13}C-labeled glucose (or $^{13}CO_2$).[18] To avoid isotope dilution, fetal bovine serum (FBS) was dialyzed against phosphate buffered saline. Glutamine and cysteine, which were degraded during hydrolysis, were separately added to support cell growth. Doubly-labeled ^{15}N, ^{13}C-glutamine and singly-labeled ^{15}N-cysteine was used. Based on the ^{15}N-coupled ^{1}H spectrum, the signal intensities around 11.6 ppm revealed 95% ^{15}N labeling efficiency for urokinase, although this may be an upper limit due to the low signal-to-noise ratio of the spectrum. In their later work, the presence of unlabeled glucose in the medium led to amino-acid-type-dependent ^{13}C dilution.[19] For example, the labeling efficiency of alanine, serine and glycine residues were in the 60%–75% range, whereas minimal dilution was observed for the amino acids synthesized by the citric acid cycle. A similar approach was taken by Lustbader et al., where glucose and pyruvate were both labeled with ^{13}C.[20] As a result, mass spectrometric studies confirmed greater than 90% uniform ^{13}C, ^{15}N labeling in CHO cells. In a relatively recent study, Egorova-Zachernyuk et al. used yeastolate and algal lysate to achieve uniform ^{15}N labeling at the efficiency of 90% based on the FT-IR amide band shift.[21]

Despite the cost reduction by using various lysates as a source of labeled amino acids, the aforementioned methods are rather labor intensive. One straightforward strategy for uniform isotope labeling in mammalian cells is to use Bioexpress 6000 commercial medium from Cambridge Isotope Laboratories (CIL). This method generally achieves greater than 80% labeling efficiency.[22,23] Furthermore, 90% uniformly ^{15}N isotope labeled TGF-beta 1 was created by adding individual ^{15}N-labeled amino acids and ^{15}N-labeled choline to the medium.[24] Considering the high cost of commercial medium, it may be good to check *a priori* whether the protein of interest can be well expressed in mammalian cells and whether the size of the ^{13}C, ^{15}N-labeled protein is suitable for traditional NMR assignment experiments.

Amino acid type selective (AATS) labeling is an alternative isotope labeling method. AATS isotope labeling is well illustrated in rhodopsin studies, where specific types of amino acids were replaced by the corresponding labeled ones.[25-29] However, isotope scrambling, which causes unwanted labeling of other amino acids can be problematic. This effect is more pronounced for ^{15}N-labeled amino acids than ^{13}C-labeled amino acids.[24] To minimize scrambling, we can perform AATS on lysine or threonine, which do not serve as substrates for transaminases, or add a transaminase inhibitor to the growth medium.[30] Incorporation of ^{15}N-labeled alanine, cysteine, glycine, histidine, isoleucine, leucine, methionine, phenylalanine, proline, serine, tryptophan, tyrosine

and valine residues was reported to have labeling efficiencies higher than 80%, except for serine residues with 50% labeling efficiency.[30]

The overall isotope labeling efficiencies are summarized in Table 1. High efficiency of ^{15}N and ^{13}C isotope labeling was achieved by introducing a pure ^{15}N and ^{13}C isotope source to the medium, adding extensively dialyzed FBS (or no FBS) and increasing the cell passage numbers using the isotope-enriched medium.[21] Notably, the cell morphology can change because the composition of isotope-enriched medium differs from that of the optimal base medium + 10% FBS.[21] Whether the change in cell morphology can affect the native PTMs, protein yield and *in cell* NMR studies should be carefully assessed.

Deuterium (^2H) labeling suppresses ^1H–^1H dipolar interactions and is especially useful for studying large proteins.[31] Although *E. coli* can be grown in D_2O minimal medium to achieve uniform ^2H enrichment, D_2O is generally toxic to mammalian cells. However, an interesting study showed that GFP protein was successfully expressed in A549 human cell line in 70% D_2O medium.[23] ^2H-labeled amino acids can potentially be used in AATS, but the cost would be prohibitively expensive.

Protein Yield

In general, a high (>100 μM) protein concentration is required for NMR experiments even with high field magnets and cryogenic probes. Thus, it is important to increase the isotope incorporation efficiency and enhance the protein yield. Interestingly, a high percentage of undialyzed FBS in the medium will generally result in a high protein yield at the price of a low efficiency of isotope incorporation. Several attempts have been made to increase the expression level of proteins in mammalian cells grown in the labeled media.

Table 1 summarizes the protein yields in different culture conditions. Because the yield substantially depends on the type of protein, different rows of the table should be compared with caution. In addition to the careful choice of the basal medium, isotope source and the percentage of FBS as illustrated in table 1, below efforts dramatically improved the yield of protein production.

First, rigorous cell selection methods have improved the protein yield. Since stable cell lines typically produce more proteins than transiently transfected cells, stable HEK293S cell lines comprising the opsin gene were created, which produced about 2 mg of isotope labeled rhodopsin protein per liter of medium.[29] Another method involved cloning cells with limited dilution technique, seeding one cell per well from the transfected cells, and recovering a clone with a high protein expression. This approach increased the production of human chorionic gonadotropin (hCG) by approximately ten-fold compared to the control without clone selection.[20]

Exploiting a dihydrofolate reductase (DHFR) deficient CHO cell line is another approach to obtain high levels of protein expression.[24] After transfecting the cell line with the plasmid containing both the gene of interest and the DHFR gene, cells are grown in a medium lacking hypoxanthine and thymidine to increase the number of plasmids in cells. Treating the cells with methotrexate, a DHFR inhibitor, can further amplify the copy number of the plasmids, resulting in elevated protein expression levels. Using this strategy, up to 1.2 mg of mature TGF-β1 per liter of medium was produced, despite the inherently low expression level of TGF-β1.[24] While the aforementioned methods can improve the overall protein yield, the procedure can be time consuming.

Second, viral transfection methods can be extremely efficient in mammalian cells. A recent study carried out by Sastry et al. highlights the improved protein expression via efficient delivery of a transgene to A549 cells by utilizing adenovirus infection with yields near 45 mg of HIV-1 gp120 outer domain per liter of cell growth.[23] In this study, recombinant adenovirus type 5 (Ad5) was constructed based on the recombination of shuttle vector carrying HIV-1 gp120 outer domain gene with a cosmid containing Ad5 genomic DNA with deletion of the E1 region responsible for virus replication and with the nonfunctional E3 region. Interestingly, this system enabled a high level of GFP expression around 50 mg

per liter of various media: ^{15}N or $^{15}N/^{13}C$ Bioexpress 6000, unlabeled DMEM and unlabeled DMEM with 70% D_2O.[23] Nevertheless, a long time is required for constructing recombinant adenovirus and the mammalian cells may display inflammatory response to capsid proteins.[23,32]

Third, degraded amino acids in hydrolysates can be externally added. For isotope labeling of proteins in mammalian cells, unlabeled amino acids in the chemically defined commercial basal media are omitted and the labeled amino acids are added instead. The amino acid composition is altered while using bacteria, algae or yeast lysates as a source of labeled amino acids. Previous study by Hansen et al. reported that hydrolysates without amino acid supplements could not support Sp2/0 cell growth, but the addition of glutamine and cysteine enabled obtaining 30 mg of urokinase from 1 L culture.[18] According to Lustbader et al., the addition of labeled cysteine, glutamine and arginine, where the first two were degraded during hydrolysis, was necessary for improving hCG expression when algal hydrolysate was used.[20] By using the optimized media, an hCG production level of 10 mg per liter of medium was achieved, which indicates an approximately six-fold increase compared to the expression level in unoptimized media.[20]

In cell NMR

The biological processes in cellular context can be assessed via *in cell* NMR spectroscopy.[33,34] In contrast to the traditional *in cell* NMR, where purified ^{15}N, ^{13}C-enriched proteins were introduced from outside the cell by methods such as microinjection, cell penetrating peptides, pore-forming toxins or electroporation,[35] endogenous protein expression may more faithfully reproduce the cellular condition. This is because endogenous expression can avoid protein damages during the purification steps and protein maturation can occur in its native environment. However, low NMR sensitivity and overflow of NMR signals from cellular background often create major huddles and restricts the analysis of cellular proteins. To alleviate these issues, a recent *in cell* NMR study used endogenous protein expression employing HEK293T cells transfected with a plasmid containing a strong CAG promoter and labeled in commercial Bioexpress 6000 media.[36] High levels of protein expression and isotope incorporation for the overexpressed protein clearly indicated that *in cell* NMR can be performed using the endogenous protein expression method. Therefore, development of isotope labeling strategies in mammalian cells can benefit both *in vitro* and *in cell* observation of proteins.

Perspectives

Lower organisms such as *E. coli* and yeast generally provide high protein yield and good isotope labeling efficiency. However, as we proceed to more complex and large protein systems, species-specific PTMs and cellular machineries involved in protein folding may be necessary to obtain fully functional proteins, or in certain cases, obtain any protein at all. Unquestionably, a protein will be in its most native form, when it is expressed in the species from which it originates.

The disadvantages for using mammalian expression system includes high cost and low protein yield compared to other prokaryotic or low eukaryotic expression systems. Furthermore, the inability of mammalian cells to grow in minimal media lowers the efficiency of stable isotope labeling, a prerequisite for NMR analysis.

We have summarized the efforts in this field to overcome these challenges. For uniform labeling, early developments focused on the use of bacterial and algal hydrolysates as isotope-labeled amino acid sources. Later, yeast and algal autolysates were used; these do not require the addition of isotope-labeled glutamine and cysteine, which are degraded in hydrolysates. Although the commercial Bioexpress 6000 media provides higher protein yield than the lysates, the latter provides slightly higher isotope incorporation efficiency (Table 1), presumably because lysates are often used without FBS. Furthermore, transfecting the mammalian cells by adenovirus or using a strong promoter helps to increase the yield and thereby reduces the cost of

protein production. In addition to uniform labeling, AATS labeling was introduced and extensively applied to NMR studies of rhodopsin. The advantages of AATS labeling are low cost and simplicity of NMR spectra, because one can judiciously select the amino acids to be isotope labeled. However, residues are assigned usually by individual mutations for AATS.

Effective isotope labeling strategies of proteins in mammalian cells are being developed for studies of complex proteins by NMR and to capture the genuine state of proteins in cellular context. As we delve into more complex and important biological systems, the significance of efficient methods for stable isotope labeling in mammalian cells will become increasingly highlighted in NMR studies.

Table 1. Summary of isotope labeling of proteins in mammalian cells.

Cell line	Protein	Basal medium	Source of ^{13}C, ^{15}N isotopes	FBS (%)	Yield (mg/L)	Incorporation (%)	Method to evaluate incorporation	Ref.
Sp2/0	Urokinase	Hybridoma SFM[a,b]	[U-^{15}N] bacterial hydrolysate[c]	5[d]	30	95	NMR	[18]
Sp2/0	Urokinase	Hybridoma SFM[a,b]	[U-^{15}N, ^{13}C] algal hydrolysate[c]	5[d]	-	60-75[e]	NMR	[19]
CHO	hCG	CHO-S-SFM[a,f]	[U-^{15}N, ^{13}C] algal hydrolysate[g]	-	10	> 90	Mass spec.	[20]
CHO	IgG1	NYSF404[a,f]	[U-^{15}N, ^{13}C] algal amino acid mixture[h]	2[d]	-	-	-	[37]
1B10.7	IgG2a	NYSF404[a]	[U-^{15}N] amino acid mixture[i]	-	-	-	-	[38]
CHO	Opsin	DMEM/F12[a]	[U-^{15}N] yeastolate and algal autolysate	-	-	> 90	FT-IR	[21]
HEK 293T	SOD1, Mia40	DMEM[a]	[U-^{15}N] algal autolysate[j]	1.5	-	-	-	[39]
CHO	TGF- β1	MEM-α[a]	[U-^{15}N] amino acids[k]	1[d]	1	> 90	NMR	[24]

HEK 293T	Proteome	CGM-6000-N	CGM-6000-N[l]	1.5	-	84	FT-IR	[22]
HEK 293T	Various proteins[m]	CGM-6000-N	CGM-6000-N[l]	1.5	-	-	-	[36]
A549	HIV-1 gp120	CGM-6000-N	CGM-6000-N[l]	-	45[n]	85	Mass spec.	[23]
A549	HIV-1 gp120	CGM-6000-CN	CGM-6000-CN[l]	-	44[n]	85	Mass spec.	[23]
CHO	TGF- β1	MEM-α[a]	[^{15}N, ^{13}C], [^{15}N], and [^{13}C] amino acids[o]	1[d]	0.5-1.0[s]	Varies depending on residues	NMR	[24]
CHO	TGF-β1	MEM-α[a]	[U-^{15}N, ^{13}C] and [2-^{15}N] amino acids[o]	1[d]	0.7	Varies depending on residues	NMR	[40]
CHO	hu-sCD2$_{105}$	MEM-α[a]	[^{15}N] Lys	15[p]	-	-	-	[41]
AN02	AN02 antibody	RPMI 1640[a]	[ring-^{2}H] Tyr, Phe, and Trp	1	20	> 95	NMR	[42]
HEK 293S	Rhodopsin	CGM-6750	CGM- 6750[q]	-	2	-	-	[29]
HEK 293S	Rhodopsin	CGM-6750	CGM- 6750[t]	-	2.1	-	-	[29]
HEK 293S	Rhodopsin	DMEM[a]	[α-^{15}N] Lys	10[d]	-	-	-	[25]
HEK 293S	Rhodopsin	DMEM[a]	[α,ε-^{15}N] Trp	10[d]	-	-	-	[26]
HEK 293S	Rhodopsin	DMEM[a]	[U-^{15}N, ^{13}C] Trp or [U-^{15}N] His or [U-^{13}C] His	10[d]	-	-	-	[27]
HEK 293S	Rhodopsin	DMEM[a]	[α,ε-^{15}N] Trp[r]	10[d]	-	-	-	[28]

The symbol '-' indicates that the information was not explicitly mentioned in the article

[a] Unlabeled amino acids substituted by the labeled amino acids were omitted

[b] Purchased from Gibco

[c] Supplemented with [^{15}N] Cys and [^{15}N] Gln (or [^{15}N/^{13}C] Gln)
[d] Dialyzed to remove unlabeled amino acids
[e] ^{13}C incorporation data of only Ala, Ser and Gly are available
[f] Carbohydrates and organic acids were omitted
[g] Supplemented with [^{15}N/^{13}C] Cys, Gln, Arg, ^{13}C-sodium pyruvate and ^{13}C-glucose
[h] Purchased from Taiyo Nippon Sanso Co. and supplemented with additional [U-^{15}N,^{13}C] amino acids, ^{13}C-glucose, ^{13}C-sodium pyruvate and ^{13}C-succinic acid from Shoko Co. as indicated in Ref. [37]
[i] Purchased from Chlorella Industry Co. and supplemented with additional [U-^{15}N] amino acids from CIL as indicated in Ref. [38]
[j] [U-^{15}N] *Streptomyces platensis* autolysate purchased from CIL and supplemented with unlabeled glucose
[k] [U-^{15}N] amino acids (except for Trp) and Trp (^{15}N labeled only on the backbone) were purchased from Merck and CIL, respectively. ^{15}N-choline was supplemented
[l] Commercial [U-^{15}N] or [U-^{15}N,^{13}C] labeling media (Bioexpress 6000) from CIL
[m] SOD1, CCS, HAH1, Mia40, Cox17 and profilin1
[n] Adenovirus vector-based expression
[o] Individual amino acids were purchased from CIL and Merck according to the labeling scheme in Ref. [24] and [40]
[p] FBS concentration was raised to 15% in the production phase
[q] [U-^{15}N] Gly, Gln, Leu, Lys, Ser, Thr, Trp and Val
[r] Additional ^{13}C labeled amino acids were added as indicated in Ref. [28]
[s] Yields from the three labeling schemes indicated in Ref. [24]
[t] [U-^{15}N,^{13}C] Gly, Gln, Leu, Lys, Ser, Thr and Val

Acknowledgments

This research was supported by the National Research Foundation of Korea grant, funded by the Ministry of Education, Science and Technology (2019R1C1C1009685).

References

1. G. L. Rosano and E. A. Ceccarelli, *Front Microbiol*. **5**, 172 (2014)
2. M. Ikura, L. E. Kay and A. Bax, *Biochemistry* **29**, 4659 (1990)
3. G. Duan and D. Walther, *PLOS Comput. Biol.* **11**, e1004049 (2015)
4. M. Merlin, E. Gecchele, S. Capaldi, M. Pezzotti and L. Avesani, *BioMed Res. Int.* **2014**, 136419 (2014)
5. K. M. Davis, T. Foos, C. S. Bates, E. Tucker, C.-C. Hsu, W. Chen, H. Jin, J. B. Tyburski, J. V. Schloss, A. J. Tobin and J.-Y. Wu, *Biochem. Biophys. Res. Commun*. **267**, 777 (2000)
6. M. Powell, L. Prentice, T. Asawa, R. Kato, J. Sawicka, H. Tanaka, V. Petersen, A. Munkley, S. Morgan, B. R. Smith and J. Furmaniak, *Clin. Chim. Acta* **256**, 175 (1996)
7. A. J. Moody, K. R. Hejnæs, M. O. Marshall, F. S. Larsen, E. Boel, I. Svendsen, E. Mortensen and T. Dyrberg, *Diabetologia* **38**, 14 (1995)
8. T. Matsuba, M. Yano, N. Abiru, H. Takino, S. Akazawa, S. Nagataki and K. Yasukawa, *J. Biochem*. **121**, 20 (1997)
9. T. W. Kim, B. H. Chung and Y. K. Chang, *Biotechnol. Prog*. **21**, 524 (2005)
10. H. Li, Y. Wang, A. Xu, S. Li, S. Jin and D. Wu, *FEMS Yeast Res*. **11**, 160 (2011)
11. C. W. Zhao, J. X. Wang, D. H. Xiao and X. K. Ma, *Sci. China* **37**, 1073 (1994)

12. P. T. Reddy, R. G. Brinson, J. T. Hoopes, C. McClung, N. Ke, L. Kashi, M. Berkmen and Z. Kelman, *MAbs* **10**, 992 (2018)
13. T. Ju, Q. Zheng and R. D. Cummings, *Glycobiology* **16**, 947 (2006)
14. K. W. Moremen, M. Tiemeyer and A. V. Nairn, *Nat. Rev. Mol. Cell Biol*. **13**, 448 (2012)
15. R. T. Peters, G. Toby, Q. Lu, T. Liu, J. D. Kulman, S. C. Low, A. J. Bitonti and G. F. Pierce, *J. Thromb. Haemost*. **11**, 132 (2013)
16. R. M. Addabbo, M. D. Dalphin, M. F. Mecha, Y. Liu, A. Staikos, V. Guzman-Luna and S. Cavagnero, *J. Phys. Chem. B* **124**, 6488 (2020)
17. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Research* **28**, 235 (2000)
18. A. P. Hansen, A. M. Petros, A. P. Mazar, T. M. Pederson, A. Rueter and S. W. Fesik, *Biochemistry* **31**, 12713 (1992)
19. A. P. Hansen, A. M. Petros, R. P. Meadows, D. G. Nettesheim, A. P. Mazar, E. T. Olejniczak, R. X. Xu, T. M. Pederson, J. Henkin and S. W. Fesik, *Biochemistry* **33**, 4847 (1994)
20. J. W. Lustbader, S. Birken, S. Pollak, A. Pound, B. T. Chait, U. A. Mirza, S. Ramnarain, R. E. Canfield and J. M. Brown, *J. Biomol. NMR* **7**, 295 (1996)
21. T. A. Egorova-Zachernyuk, G. J. Bosman and W. J. Degrip, *Appl. Microbiol. Biotechnol*. **89**, 397 (2011)
22. E. Mitri, L. Barbieri, L. Vaccari and E. Luchinat, *Analyst* **143**, 1171 (2018)
23. M. Sastry, L. Xu, I. S. Georgiev, C. A. Bewley, G. J. Nabel and P. D. Kwong, *J. Biomol. NMR* **50**, 197 (2011)
24. S. J. Archer, A. Bax, A. B. Roberts, M. B. Sporn, Y. Ogawa, K. A. Piez, J. A. Weatherbee, M. L. S. Tsang and R. Lucas, *Biochemistry* **32**, 1152 (1993)
25. J. Klein-Seetharaman, P. Reeves, M. Loewen, E. Getmanova, J. Chung, H. Schwalbe, P. Wright and H. Khorana, *Proc. Natl. Acad. Sci. U.S.A*. **99**, 3452 (2002)
26. J. Klein-Seetharaman, N. V. K. Yanamala, F. Javeed, P. J. Reeves, E. V. Getmanova, M. C. Loewen, H. Schwalbe and H. G. Khorana, *Proc. Natl. Acad. Sci. U.S.A*. **101**, 3409 (2004)
27. A. B. Patel, E. Crocker, P. J. Reeves, E. V. Getmanova, M. Eilers, H. G. Khorana and S. O. Smith, *J. Mol. Biol*. **347**, 803 (2005)
28. K. Werner, I. Lehner, H. K. Dhiman, C. Richter, C. Glaubitz, H. Schwalbe, J. Klein-Seetharaman and H. G. Khorana, *J. Biomol. NMR* **37**, 303 (2007)
29. K. Werner, C. Richter, J. Klein-Seetharaman and H. Schwalbe, *J. Biomol. NMR* **40**, 49 (2008)
30. Y. Arata, K. Kato, H. Takahashi and I. Shimada, *Methods Enzymol*. **239**, 440 (1994)
31. L. E. Kay and K. H. Gardner, *Curr. Opin. Struct. Biol*. **7**, 722 (1997)
32. C. S. Lee, E. S. Bishop, R. Zhang, X. Yu, E. M. Farina, S. Yan, C. Zhao, Z. Zheng, Y. Shu, X. Wu, J. Lei, Y. Li, W. Zhang, C. Yang, K. Wu, Y. Wu, S. Ho, A. Athiviraham, M. J. Lee, J. M. Wolf, R. R. Reid and T.-C. He, *Genes Dis*. **4**, 43 (2017)
33. F.-X. Theillet, A. Binolfi, B. Bekei, A. Martorana, H. M. Rose, M. Stuiver, S. Verzini, D. Lorenz, M. van Rossum, D. Goldfarb and P. Selenko, *Nature* **530**, 45 (2016)
34. H. Wen, Y. J. An, W. J. Xu, K. W. Kang and S. Park, *Angew. Chem. Int. Ed*. **54**, 5374 (2015)
35. E. Luchinat and L. Banci, *IUCrJ*. **4**, 108 (2017)
36. L. Barbieri, E. Luchinat and L. Banci, *Nat. Protoc*. **11**, 1101 (2016)
37. Y. Yamaguchi, M. Nishimura, M. Nagano, H. Yagi, H. Sasakawa, K. Uchida, K. Shitara and K. Kato, *Biochim. Biophys. Acta* **1760**, 693 (2006)
38. K. Shindo, K. Masuda, H. Takahashi, Y. Arata and I. Shimada, *J. Biomol. NMR* **17**, 357 (2000)
39. C. Fuccio, E. Luchinat, L. Barbieri, S. Neri and M. Fragai, *J. Biomol. NMR* **64**, 275 (2016)

40. A. P. Hinck, S. J. Archer, S. W. Qian, A. B. Roberts, M. B. Sporn, J. A. Weatherbee, M. L. S. Tsang, R. Lucas, B.-L. Zhang, J. Wenker and D. A. Torchia, *Biochemistry* **35**, 8517 (1996)
41. D. F. Wyss, J. M. Withka, M. H. Knoppers, K. A. Sterne, M. A. Recny and G. Wagner, *Biochemistry* **32**, 10995 (1993)
42. J. Anglister, T. Frey and H. M. McConnell, *Biochemistry* **23**, 1138 (1984)