

Alma Mater Studiorum Università di Bologna  
Archivio istituzionale della ricerca

Sequential protein expression and selective labeling for in-cell NMR in human cells

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

LUCHINAT, E., SECCI, E., CENCETTI, F., BRUNI, P. (2016). Sequential protein expression and selective labeling for in-cell NMR in human cells. *BIOCHIMICA ET BIOPHYSICA ACTA-GENERAL SUBJECTS*, 1860(3), 527-533 [10.1016/j.bbagen.2015.12.023].

*Availability:*

This version is available at: <https://hdl.handle.net/11585/856164> since: 2022-02-11

*Published:*

DOI: <http://doi.org/10.1016/j.bbagen.2015.12.023>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)

## Sequential protein expression and selective labeling for in-cell NMR in human cells

Enrico Luchinat<sup>a,b,\*</sup>, Erica Secci<sup>a</sup>, Francesca Cencetti<sup>b</sup>, Paola Bruni<sup>b</sup>

<sup>a</sup>Magnetic Resonance Center - CERM, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Florence, Italy.

<sup>b</sup>Department of Biomedical, Experimental and Clinical Sciences “Mario Serio”, University of Florence, Viale Morgagni 50, 50134 Florence, Italy.

\* Corresponding author at: Department of Biomedical, Experimental and Clinical Sciences “Mario Serio”, University of Florence, Viale Morgagni 50, 50134 Florence, Italy. E-mail address: eluchinat@cerm.unifi.it

This is the final peer-reviewed accepted manuscript of:

Enrico Luchinat, Erica Secci, Francesca Cencetti, Paola Bruni

*Sequential protein expression and selective labeling for in-cell NMR in human cells*

which has been published in final form in *Biochimica et Biophysica Acta (BBA) - General Subjects* Volume 1860, Issue 3, March 2016, Pages 527-533

The final published version is available online at: <https://doi.org/10.1016/j.bbagen.2015.12.023>

© 2015 Elsevier. This manuscript version is made available under the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) 4.0 International License (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

## **Abstract**

**Background:** In-cell NMR is a powerful technique to investigate proteins in living human cells at atomic resolution. Ideally, when studying functional processes involving protein-protein interactions by NMR, only one partner should be isotopically labeled. Here we show that constitutive and transient protein expression can be combined with protein silencing to obtain selective protein labeling in human cells.

**Methods:** We established a human cell line stably overexpressing the copper binding protein HAH1. A second protein (human superoxide dismutase 1, SOD1) was overexpressed by transient transfection and isotopically labeled. A silencing vector containing shRNA sequences against the HAH1 gene was used to decrease the rate of HAH1 synthesis during the expression of SOD1. The levels of HAH1 mRNA and protein were measured as a function of time following transfection by RT-PCR and Western Blot, and the final cell samples were analyzed by in-cell NMR.

**Results:** SOD1 was ectopically expressed and labeled in a time window during which HAH1 biosynthesis was strongly decreased by shRNA, thus preventing its labeling. In-cell NMR spectra confirmed that, while both proteins were present, only SOD1 was selectively labeled and could be detected by  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear NMR.

**Conclusions and General Significance:** We showed that controlling protein expression by specifically silencing a stably expressed protein is a useful strategy to obtain selective isotope labeling of only one protein. This approach relies on established techniques thus permitting the investigation of protein-protein interactions by NMR in human cells.

**Keywords:** in-cell NMR; nuclear magnetic resonance; protein-protein interactions; isotope labeling; mammalian cells.

## 1. Introduction

In-cell NMR spectroscopy has been proven to be an ideal technique to obtain atomic-level structural and functional information on biological macromolecules directly in living cells, i.e. in their native environment [1–3]. The data obtained by in-cell NMR are likely closest to the physiological conditions than any other atomic-resolution technique. In-cell NMR allows physiological processes to be investigated such as protein folding and misfolding [4–6], interactions with other proteins or nucleic acids [7–10], binding of metal ions [11–13], and ligand screening [14]. Most of the initial works focused on proteins overexpressed in *E. coli* cells, due to the high protein levels attainable and to the ease of growing and handling the cells. Although bacterial cells are considered a good model to investigate cytoplasmic proteins, they lack many cellular machineries for protein folding and post-translational modifications, which are crucial for the correct maturation of many eukaryotic proteins. Due to such limitations of the bacterial models, in-cell NMR in eukaryotic cells is a necessary step forward for studying proteins in their native environment. Several approaches have been developed to deliver a purified labeled protein into the cytoplasm of eukaryotic cells, including microinjection in *Xenopus laevis* oocytes [15], fusion with cell-penetrating peptides [16], and cell permeabilization either through pore-forming toxins [17,18] or by cell electroporation [19,20]. Alternatively, methods have been developed to overexpress the protein of interest directly into eukaryotic cells, and have been applied to yeast and insect cells [21,22] and to cultured human cells [13].

Protein overexpression in human cells has been used to follow the various steps of protein maturation right after protein biosynthesis, including folding and metal binding [5,13,23], changes in redox state [24], and targeting towards cellular compartments such as mitochondria [25]. With the same approach, ectopic co-expression of two or more proteins has permitted the observation of functional interactions with specific partners on the protein of interest [5,13,24]. The employed strategy for protein co-expression causes the two (or more) proteins to be expressed simultaneously by the cells, and does not allow any labeling selectivity between them. Therefore, as soon as the isotopically enriched medium is provided, all the expressed proteins are equally labeled, and are detected in the in-cell NMR spectra. Ideally, when studying biomolecular interactions by NMR, only one partner should be labeled in order to allow the correct interpretation of any spectral changes upon interaction. In bacterial cells, methods have been developed for this purpose, in which

orthogonal induction systems are used for two or more proteins, allowing them to be expressed at different times and with different isotope labeling [7,8]. To date, similar approaches that permit sequential protein expression and labeling in human cells are lacking.

Here we show that constitutive and transient protein expression can be combined with protein knock-down by gene silencing to allow two human proteins to be expressed at different times in human cells, permitting selective isotope labeling of one protein. This novel methodological approach was tested and optimized on two human proteins that have been extensively characterized by NMR both in-cell and *in vitro*: HAH1 [26,27] and SOD1 [5,11,13,28,29]. We are confident that this novel experimental procedure will enable the study of protein-protein interactions in mammalian cells by NMR.

## 2. Materials and Methods

### 2.1. Gene cloning

The gene encoding full-length human HAH1 (amino acids 1-68; MW: 7401.63 Da; theoretical pI: 6.70; GenBank accession number: NP\_004036.1) was amplified from cDNA by PCR and cloned in the pURD vector (kindly provided by Y. Zhao, University of Oxford) [30] between HindIII and XhoI restriction enzymes. Three sequences of small hairpin RNA (shRNA) were designed to perform silencing of HAH1, in the form of sense-loop-antisense sequences [31,32]. One sequence, shRNA-1, targets the HAH1 open reading frame (5'-

**GACGAGTTCTCTGTGGACATGACGAATCATGTCCACAGAGAACTCGTTTTTTTGGAA-3'**;

sense and antisense sequences are shown in bold) while other two target unique UTR elements present in the pURD vector: shRNA-2 targets the woodchuck hepatitis virus post-transcriptional regulatory element (WPRE: 5'-

**GCAATCAACCTCTGGATTACACGAATGTAATCCAGAGGTTGATTGCTTTTTTGGAA-3'**);

shRNA-3 targets the bovine growth hormone polyadenylation signal (bGHpA: 5'-**GGAAGACAATAGCAGGCATGCCGAAGCATGCCTGCTATTGTCTTCCTTTTTTGGAA-3'**). The

corresponding synthetic double stranded DNA oligonucleotides – flanked by compatible ends for BamHI and

HindIII restriction sites – were individually cloned in the pSilencer 2.1-U6 Hygro vector (Life Technologies) following the vector manual. pHLsec-derived vectors containing the cDNA encoding the full-length human SOD1 (pHLsec-SOD1; amino acids 1-154; MW: 15935.74 Da; theoretical pI: 5.70; GenBank accession number: NP\_000445.1) and the full-length human thioredoxin 1 (pHLsec-TRX1; amino acids 1-105; MW: 11737.50 Da; theoretical pI: 4.82; GenBank accession number: NP\_003320.2) were used for transient expression as previously described [13].

## *2.2. Generation of stable cell lines*

HEK293T-derived cell lines stably overexpressing HAH1 were obtained following a reported protocol [33]. Parental 293T cells (ATCC CRL-3216) were transfected in a 6-well dish with the pURD vector containing the HAH1 cDNA, together with the PhiC31 integrase gene (pgk-phiC31/pCB92) in a 1:3 (HAH1:phiC31) molar ratio, to facilitate integration into the host cell genome. The following day the cells were seeded into a 15 cm dish, and grown for 5 days. Once they reached ~100% confluence, stable cell line selection was performed by treating the cells with 2 µg/ml puromycin for 10 days, replacing the growth medium every 3-4 days to remove dead cells, until cell clumps (single colonies) were visible. Single colonies were manually picked from the dish and individually seeded 96-well plate wells. During the following days, 36 growing colonies (H1 – H36) were selected by manual inspection and seeded to 12-well plates. 12 lines out of these were selected based on the growth rate, and were expanded to 6-well plates and further to T25 flasks. The HAH1 expression levels were measured by Western Blot analysis, and the cell line (H22) with the highest expression levels was chosen for the NMR experiments.

## *2.3. Cell culture and transient transfection*

HEK293T and HAH1 stably-transfected HEK293T cells were maintained in Dulbecco's Modified Eagle's medium (DMEM; high glucose, Gibco) supplemented with L-glutamine (Gibco), antibiotics (penicillin and streptomycin) (Gibco) and 10% FBS (Gibco) in uncoated 75-cm<sup>2</sup> plastic flasks and incubated at 37 °C, 5% CO<sub>2</sub> in a humidified atmosphere. Stably-transfected cells were additionally supplemented with MEM NEAA (Gibco). Transient transfection was performed by DNA:polyethylenimine complex as previously described [13]. Cells expressing SOD1 with silenced HAH1 were obtained by transfection with pHLsec-SOD1 and the

three silencing vectors (3:1:1:1 ratio); unsilenced cells expressing SOD1 were obtained by transfection with pHlsec-SOD1 and empty pHlsec (1:1 ratio); control cells not expressing SOD1 were transfected as above, by replacing pHlsec-SOD1 with empty pHlsec. The total DNA was kept constant at 25 µg / 75 cm<sup>2</sup> flask in all experiments. [U-<sup>15</sup>N]-BioExpress 6000 medium (Cambridge Isotope Laboratories) was used for <sup>15</sup>N labeling. Cell samples expressing SOD1 were supplemented with 10 µM ZnSO<sub>4</sub> starting at the time of transfection.

#### 2.4. Western Blot and protein quantification

Protein content was measured by Western Blot analysis. HAH1 was stained with a rabbit monoclonal anti-HAH1 antibody (Abcam: ab154179, diluted to 53.2 ng/mL). GAPDH was used as reference (rabbit polyclonal anti-GAPDH antibody, Abcam: ab9485, diluted at 1:2,000). Goat anti-rabbit IgG (whole molecule)-peroxidase secondary antibody (Sigma:A0545) was used, diluted at 1:80,000. For detection, LiteAbloT EXTEND chemiluminescent substrate (EuroClone) was used. Purified recombinant HAH1 was used as standard for the quantification. HAH1 was expressed with a 6-His tag at C-terminal part in *E. coli* BL21(DE3) gold competent cells for 4 h at 37 °C. After cell lysis, a first purification step was performed using a HiTrap chelating (GE Healthcare) column charged with Zn(II). The protein was eluted in 20 mM Na<sub>2</sub>HPO<sub>4</sub>, 0.5 M NaCl, 5 mM imidazole, 20 mM EDTA, pH=8, buffer. After buffer exchange and digestion with Factor Xa protease (New England Biolabs) for 24 hr at 25 °C, the protein was separated from the affinity tag in the HiTrap column charged with Zn(II).

#### 2.5. Real-time PCR

Total RNA (1 µg) was extracted from cells using TRI Reagent® (Sigma-Aldrich) and reverse-transcribed by the employment of high-capacity cDNA reverse transcription kit (Applied Biosystems), according to the manufacturer's instructions. Quantification of the mRNA level of target genes was performed by Real-Time PCR using TaqMan® gene expression assays with the automated ABI Prism 7700 sequence detector system (Applied Biosystems) as previously described[34]. Samples were run in triplicate in Micro-Amp optical 96-well plates (Applied Biosystems) with a TaqMan Universal PCR Master Mix (Applied Biosystems). Simultaneous amplification of the target sequence of HAH1 (Hs01076125\_m1, FAM<sup>TM</sup>-MGB dye) together

with the housekeeping gene, human ACTB (beta-actin) (Applied Biosystems Endogenous Control VIC®/TAMRA™ Probe) was carried out with the following profile: initial denaturation for 10 minutes at 95 °C, followed by denaturation for 15 seconds at 95 °C and primer annealing and elongation at 60 °C for 1 minute for 40 cycles. The results were analyzed by ABI Prism Sequence Detection System software (version 1.7; Applied Biosystems). The  $2^{-(\Delta\Delta CT)}$  method was used as a comparative method of quantification [35]. The data are reported as mean  $\pm$  SEM of  $2^{-(\Delta\Delta CT)}$  of samples run in triplicate; statistical analysis was performed by One-way ANOVA followed by Bonferroni post-hoc test (\* $p < 0.05$ ).

## 2.6. NMR experiments

NMR experiments on cells and lysates were acquired using a 950 MHz Bruker Avance spectrometer equipped with a TCI CryoProbe. 1D  $^1\text{H}$  (zgpg30 Bruker pulse sequence, 128 scans) and 2D  $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC spectra [36] were acquired at 308 K. The total acquisition time of the  $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC for each cell sample was about 1 h (64 scans, 128 increments) or 1.5 h (96 scans, 128 increments). The supernatant of each cell sample was checked in the same experimental conditions to exclude protein leakage. The same NMR spectra were also acquired on the cell lysates. All the spectra were processed with Bruker Topspin software. To eliminate the signals arising from the  $^{15}\text{N}$  enrichment of other cellular components, the 2D  $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC spectra of the cell lysates were further processed by subtracting the spectrum of a lysate from untransfected cells, acquired in the same experimental conditions and identically processed.

## 3. Results

### 3.1. Stable cell line characterization

HEK293T cell lines stably expressing human HAH1 were obtained from single colonies of cells transfected with the HAH1 and the integrase genes, after antibiotic selection, and the relative expression levels of HAH1 in each cell line were assessed by Western Blot analysis and SDS-PAGE (Fig. 1a). The cell line in which HAH1 reached the expression levels suitable for in-cell NMR experiments was selected (HEK293T-HAH1-22, H22 hereafter, Fig. 1b). The effective concentration of HAH1 in a NMR sample of H22 cells was ~45



$\mu\text{M}$ , as estimated by 1D  $^1\text{H}$  NMR by comparing the signal intensity between 0.1 and  $-0.2$  ppm in the cell lysate spectrum with an *in vitro* sample of recombinantly expressed HAH1 at known concentration (Fig. 2a). The amide crosspeaks arising from  $[\text{U}-^{15}\text{N}]$ -labeled HAH1 in a sample of H22 cells grown in  $[\text{U}-^{15}\text{N}]$ -labeled medium were detected in the  $^1\text{H}$ - $^{15}\text{N}$  correlation NMR spectra of both cells and lysate, showing that stably transfected cells can reach constitutive proteins levels sufficient for detection by heteronuclear in-cell NMR (Fig. 2b,c and Supplementary Figure S1a,b). The chemical shifts of the detected crosspeaks corresponded to those previously assigned on  $[\text{U}-^{15}\text{N}]$ -apo-HAH1 *in vitro* (Supplementary Figure S1c) [26]. The cell line selection strategy employed here is quite versatile [30,33] and should be easily applicable to stably overexpress other proteins.

### 3.2. Transient transfection of the second protein gives poor labeling selectivity

A sequential expression strategy was tested, in which H22 cells were transiently transfected with the gene of a second protein, and the unlabeled medium was replaced with  $[\text{U}-^{15}\text{N}]$ -labeled medium at the time of transfection. SOD1 was chosen as a test case for globular proteins that are easily detected by in-cell NMR [5,11,13]; thioredoxin 1 (TRX1) was chosen to represent globular proteins that cannot be detected by in-cell solution NMR without resorting to deuterium enrichment, due to extensive weak interactions with the cellular environment [20,24,37]. In either case, transient transfection did not decrease the levels of stably expressed HAH1, and the overexpressed protein reached the same levels as previously observed in the parental HEK cells [13,24]. However, signals arising from  $^{15}\text{N}$ -labeled HAH1 were still being detected in the resulting in-cell NMR spectra (Fig. 3a,b and Supplementary Figure S2a,c) and in the corresponding lysates (Fig. 3c,d and Supplementary Figure S2b,d), due to HAH1 being still synthesized in the  $[\text{U}-^{15}\text{N}]$ -labeled medium, together with the second protein.

### 3.3. Silencing of HAH1 reduces the rate of protein synthesis

To avoid the unwanted labeling of HAH1 during the expression of the second protein, the rate of HAH1 synthesis needs to be reduced after transfection. Gene silencing by RNA interference (RNAi) is a widely

used technique to knock-down gene expression in mammalian cells. With that purpose, a silencing vector was used to transcribe sequences of short hairpin RNA (shRNA) complementary to different regions of the HAH1 mRNA. Three different sequences were individually cloned in the silencing vector (Fig. 4a). The silencing vector was co-transfected with the gene of the second protein. The relative levels of HAH1 mRNA and protein were measured at different times after transfection by RT-PCR and Western Blot, respectively. Efficient silencing was achieved by mixing together the three shRNA-encoding sequences (Fig. 4b). When comparing the levels of mRNA with those of HAH1 protein as a function of time after transfection, it was observed that the protein levels decreased much slower than those of mRNA, indicating that the protein turnover rate is low enough to permit the expression of the second protein in the time-frame following transfection, without compromising the level of HAH1 (Fig. 4c and Supplementary Figure S3).

### *3.4. Silencing combined with timed medium switch improves the labeling selectivity*

A temporal window was identified, during which the levels of HAH1 mRNA had reached a minimum, while the HAH1 protein levels did not decrease excessively (Fig. 4c). Specific labeling of the second protein could be performed during that window with minimal unwanted labeling of HAH1. To test this procedure, H22 cells were transfected with either SOD1 or the empty vector together with the vectors encoding the shRNAs against HAH1, and were kept in unlabeled medium for 14 h followed by 16 h in [U- $^{15}\text{N}$ ]-medium (Fig. 5 and Supplementary Figure S4). Corresponding control samples were produced, in which HAH1 was not silenced. The resulting in-cell NMR spectra confirmed our expectations: HAH1 signals were almost not detected in the  $^1\text{H}$ - $^{15}\text{N}$  correlation spectra acquired on both cells (Fig. 5b,d) and lysates (Fig. 5f,h), while the  $^1\text{H}$  spectra revealed the presence of unlabeled HAH1 (~80% of the levels of unsilenced HAH1, Fig. 5i-l). The relative amounts of total and labeled HAH1 in each lysate were estimated from the ratio between the aliphatic  $^1\text{H}$  signal intensities (arising from the total HAH1, Fig. 5i-l) and the amide  $^1\text{H}$ - $^{15}\text{N}$  crosspeak intensities (arising from  $^{15}\text{N}$ -labeled HAH1, Fig. 5e-h), normalized by the same ratio obtained from a sample of purified  $^{15}\text{N}$ -labeled HAH1. In the “fully”  $^{15}\text{N}$ -labeled H22 cell sample (Fig. 2c), HAH1 was ~93% enriched, close to the nominal isotopic enrichment of the growth medium (U- $^{15}\text{N}$ , 98%). The “14+16” labeling strategy alone, without HAH1 silencing, caused HAH1 to be ~32% enriched, while the silencing further increased the

labeling selectivity as HAH1 was only ~15% enriched, and was barely detected by  $^1\text{H}$ - $^{15}\text{N}$  correlation NMR. At 14 h post-transfection the number of cells had increased from  $1.2 \pm 0.1 \times 10^7$  to  $1.8 \pm 0.2 \times 10^7$  cells, while at 30 h post-transfection, the number of cells had approximately doubled ( $2.5 \pm 0.2 \times 10^7$  cells). Providing  $[\text{U}-^{15}\text{N}]$ -medium at the time of transfection would therefore result in ~50% enrichment of the cellular components with  $^{15}\text{N}$  (causing cellular background signals to be detected in the  $^1\text{H}$ - $^{15}\text{N}$  correlation in-cell NMR spectra). With the “14+16” labeling strategy, the background isotope incorporation was decreased to ~30%, thus increasing the protein/background signal ratio and the overall quality of the NMR spectra.

#### 4. Discussion

HEK293T cells stably expressing HAH1 were co-transfected with two vectors: one encoding SOD1, the other containing a shRNA which knocked-down the expression of HAH1 by degrading its mRNA via RNAi.  $^{15}\text{N}$  labeling of SOD1 was achieved by replacing the unlabeled growth medium with  $[\text{U}-^{15}\text{N}]$ -growth medium at a defined and optimized time. The optimal times for the medium switch and sample collection after SOD1 expression were chosen by measuring HAH1 protein and mRNA levels at different time points following cell transfection. Since HAH1 has a sufficiently long intracellular half-life, a time window was identified (between 14 h and 30 h following transient transfection) during which the shRNA had degraded the HAH1 mRNA, so that the rate of HAH1 synthesis was greatly decreased, while the cellular content of HAH1 had not yet decreased below the concentration required for in-cell NMR (Fig. 4c). During this time window, SOD1 was transiently overexpressed and selectively  $^{15}\text{N}$ -labeled with respect to HAH1.

The strategy described above combines stable and transient protein expression to allow two (and in principle more) proteins to be overexpressed in human cells at different times. Without any protein expression control mechanism, the stably expressed protein would still be produced by the cells after the expression of the second protein has started, resulting in poor labeling selectivity, which is critical for proper interpretation of the NMR spectra. Our strategy exploits the RNAi machinery of the cells to control the rate of protein expression as a function of time [31,32]. For the same purpose, other methods may be used, such as the

employment of synthetic small interference RNA (siRNA) molecules instead of shRNAs [38], however these approaches would require the protocol to be adapted to combine DNA transfection with siRNA treatment. Alternatively, different strategies could be pursued to control protein expression. Inducible gene expression systems, for example the tetracycline-controlled Tet-ON/Tet-OFF systems, are commonly used for that purpose [39–41]. Stable cell lines could then be created that can overexpress the proteins of interest separately, by combining alternative inducible promoters, such as the ecdysone-controlled system [42], to the tetracycline-controlled one. However, such strategies would require additional vector engineering, and expression may not reach the level required by the relatively low sensitivity of NMR spectroscopy. A limit of sequential protein expression strategies is the intracellular half-life of the first protein. For proteins having half-life shorter than the time required to express the second protein, a suitable approach could combine the stable expression of the first protein in unlabelled media with the delivery of the second protein (isotope labelled) from outside by existing approaches. This would decrease the delay between the synthesis of the first protein and the NMR data collection.

This work shows that controlled protein expression is a potentially useful approach for studying protein interactions in human cells by NMR. Although our test proteins were not supposed to interact directly, they were useful to assess the labeling selectivity and to demonstrate the feasibility of our strategy. This method represents an innovative approach for studying *bona fide* interactions between proteins at the atomic level, and in principle it can be adapted to different isotope labeling schemes or combined with the existing protein-delivery approaches. Hopefully, this work will provide a starting point for developing advanced methods that allow finer control of protein synthesis in mammalian cells for in-cell NMR applications.

## Acknowledgments

We would like to thank Prof. Lucia Banci (Univ. of Florence) for critically reading the manuscript and providing advice; Prof. A. Radu Aricescu and Dr. Yuguang Zhao (Univ. of Oxford) for providing support for the stable cell line selection. This work has been supported by “MEDINTECH: Tecnologie convergenti per aumentare la sicurezza e l’efficacia di farmaci e vaccini” (grant number: CTN01\_00177\_962865) and by

Instruct, part of the European Strategy Forum on Research Infrastructures (ESFRI) and supported by national member subscriptions. Specifically, we thank the EU ESFRI Instruct Core Centres CERM-Italy. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under BioStruct-X (grant agreement N°283570).

## **Appendix A. Supplementary data**

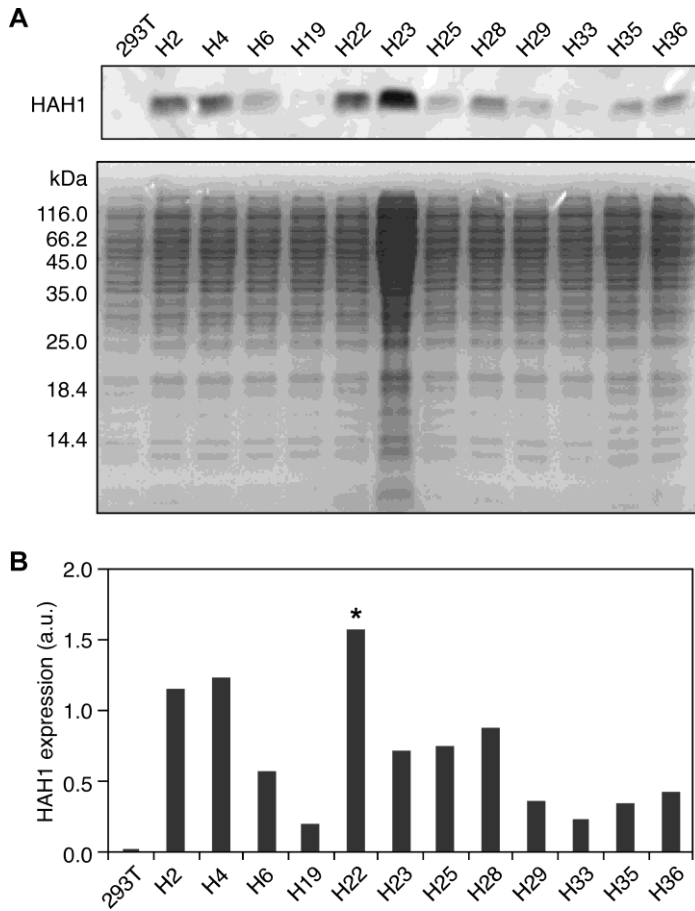
Supplementary data to this article can be found online at

## References

- [1] D.I. Freedberg, P. Selenko, Live cell NMR, *Annu. Rev. Biophys.*, 43 (2014) 171–192.
- [2] R. Hänsel, L.M. Luh, I. Corbeski, L. Trantirek, V. Dötsch, In-cell NMR and EPR spectroscopy of biomacromolecules, *Angew. Chem. Int. Ed Engl.*, 53 (2014) 10300–10314.
- [3] A.E. Smith, Z. Zhang, G.J. Pielak, C. Li, NMR studies of protein folding and binding in cells and cell-like environments, *Curr. Opin. Struct. Biol.*, 30 (2015) 7–16.
- [4] A.P. Schlesinger, Y. Wang, X. Tadeo, O. Millet, G.J. Pielak, Macromolecular crowding fails to fold a globular protein in cells, *J. Am. Chem. Soc.*, 133 (2011) 8082–8085.
- [5] E. Luchinat, L. Barbieri, J.T. Rubino, T. Kozyreva, F. Cantini, L. Banci, In-cell NMR reveals potential precursor of toxic species from SOD1 fALS mutants, *Nat. Commun.*, 5 (2014) 5502.
- [6] W.B. Monteith, R.D. Cohen, A.E. Smith, E. Guzman-Cisneros, G.J. Pielak, Quinary structure modulates protein stability in cells, *Proc. Natl. Acad. Sci. U. S. A.*, 112 (2015) 1739–1742.
- [7] D.S. Burz, K. Dutta, D. Cowburn, A. Shekhtman, Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR), *Nat. Methods*, 3 (2006) 91–93.
- [8] D.S. Burz, A. Shekhtman, In-cell biochemistry using NMR spectroscopy, *PloS One*, 3 (2008) e2571.
- [9] A.M. Augustus, P.N. Reardon, L.D. Spicer, MetJ repressor interactions with DNA probed by in-cell NMR, *Proc. Natl. Acad. Sci. U. S. A.*, 106 (2009) 5065–5069.
- [10] L.M. Luh, R. Hänsel, F. Löhr, D.K. Kirchner, K. Krauskopf, S. Pitzius, B. Schäfer, P. Tufar, I. Corbeski, P. Güntert, V. Dötsch, Molecular crowding drives active Pin1 into nonspecific complexes with endogenous proteins prior to substrate recognition, *J. Am. Chem. Soc.*, 135 (2013) 13796–13803.
- [11] L. Banci, L. Barbieri, I. Bertini, F. Cantini, E. Luchinat, In-cell NMR in E coli to monitor maturation steps of hSOD1, *PloS One*, 6 (2011) e23561.
- [12] D.S.S. Hembram, T. Haremakei, J. Hamatsu, J. Inoue, H. Kamoshida, T. Ikeya, M. Mishima, T. Mikawa, N. Hayashi, M. Shirakawa, Y. Ito, An in-cell NMR study of monitoring stress-induced increase of cytosolic Ca<sup>2+</sup> concentration in HeLa cells, *Biochem. Biophys. Res. Commun.*, 438 (2013) 653–659.
- [13] L. Banci, L. Barbieri, I. Bertini, E. Luchinat, E. Secci, Y. Zhao, A.R. Aricescu, Atomic-resolution monitoring of protein maturation in live human cells by NMR, *Nat. Chem. Biol.*, 9 (2013) 297–299.
- [14] J. Xie, R. Thapa, S. Reverdatto, D.S. Burz, A. Shekhtman, Screening of small molecule interactor library by using in-cell NMR spectroscopy (SMILI-NMR), *J. Med. Chem.*, 52 (2009) 3516–3522.
- [15] R. Thongwichian, P. Selenko, In-cell NMR in *Xenopus laevis* oocytes, *Methods Mol. Biol.* Clifton NJ, 895 (2012) 33–41.
- [16] K. Inomata, A. Ohno, H. Tochio, S. Isogai, T. Tenno, I. Nakase, T. Takeuchi, S. Futaki, Y. Ito, H. Hiroaki, M. Shirakawa, High-resolution multi-dimensional NMR spectroscopy of proteins in human cells, *Nature*, 458 (2009) 106–109.
- [17] S. Ogino, S. Kubo, R. Umemoto, S. Huang, N. Nishida, I. Shimada, Observation of NMR signals from proteins introduced into living mammalian cells by reversible membrane permeabilization using a pore-forming toxin, streptolysin O, *J. Am. Chem. Soc.*, 131 (2009) 10834–10835.
- [18] S. Kubo, N. Nishida, Y. Udagawa, O. Takarada, S. Ogino, I. Shimada, A gel-encapsulated bioreactor system for NMR studies of protein-protein interactions in living mammalian cells, *Angew. Chem. Int. Ed Engl.*, (2012).
- [19] B. Bekei, In-cell NMR spectroscopy in mammalian cells, Freie Universität Berlin, Germany, 2013.
- [20] S. Majumder, J. Xue, C.M. DeMott, S. Reverdatto, D.S. Burz, A. Shekhtman, Probing protein quinary interactions by in-cell nuclear magnetic resonance spectroscopy, *Biochemistry (Mosc.)*, 54 (2015) 2727–2738.
- [21] K. Bertrand, S. Reverdatto, D.S. Burz, R. Zitomer, A. Shekhtman, Structure of proteins in eukaryotic compartments, *J. Am. Chem. Soc.*, 134 (2012) 12798–12806.
- [22] J. Hamatsu, D. O'Donovan, T. Tanaka, T. Shirai, Y. Hourai, T. Mikawa, T. Ikeya, M. Mishima, W. Boucher, B.O. Smith, E.D. Laue, M. Shirakawa, Y. Ito, High-resolution heteronuclear multidimensional NMR of proteins in living insect cells using a baculovirus protein expression system, *J. Am. Chem. Soc.*, 135 (2013) 1688–1691.
- [23] E. Luchinat, A. Gianoncelli, T. Mello, A. Galli, L. Banci, Combining in-cell NMR and X-ray fluorescence microscopy to reveal the intracellular maturation states of human superoxide dismutase 1, *Chem. Commun. Camb. Engl.*, 51 (2015) 584–587.

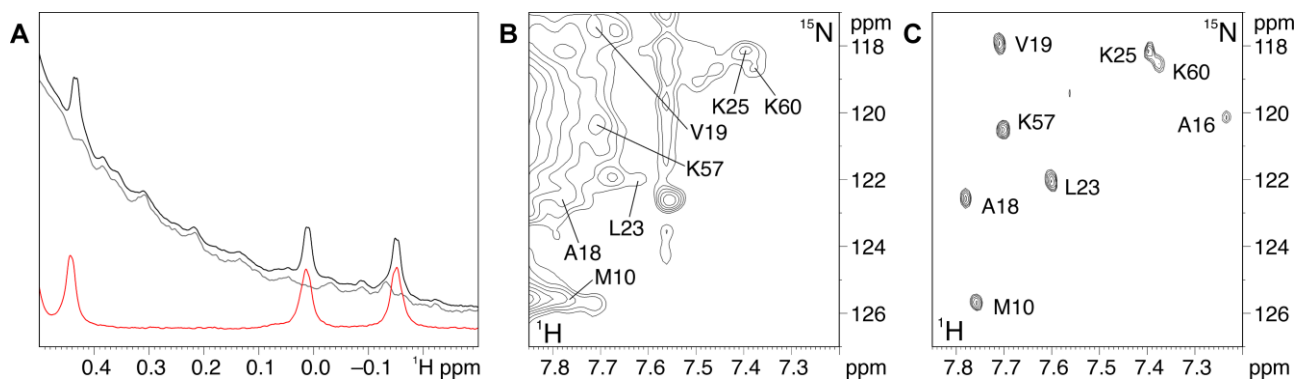
- [24] L. Banci, L. Barbieri, E. Luchinat, E. Secci, Visualization of redox-controlled protein fold in living cells, *Chem. Biol.*, 20 (2013) 747–752.
- [25] L. Barbieri, E. Luchinat, L. Banci, Structural insights of proteins in sub-cellular compartments: In-mitochondria NMR, *Biochim. Biophys. Acta*, 1843 (2014) 2492–2496.
- [26] I. Anastassopoulou, L. Banci, I. Bertini, F. Cantini, E. Katsari, A. Rosato, Solution structure of the apo and copper(I)-loaded human metallochaperone HAH1, *Biochemistry (Mosc.)*, 43 (2004) 13046–13053.
- [27] F. Arnesano, L. Banci, I. Bertini, I.C. Felli, M. Losacco, G. Natile, Probing the interaction of cisplatin with the human copper chaperone Atox1 by solution and in-cell NMR spectroscopy, *J. Am. Chem. Soc.*, 133 (2011) 18361–18369.
- [28] L. Banci, I. Bertini, F. Cantini, N. D’Amelio, E. Gaggelli, Human SOD1 before harboring the catalytic metal: solution structure of copper-depleted, disulfide-reduced form, *J. Biol. Chem.*, 281 (2006) 2333–2337.
- [29] L. Banci, I. Bertini, F. Cantini, T. Kozyreva, C. Massagni, P. Palumaa, J.T. Rubino, K. Zovo, Human superoxide dismutase 1 (hSOD1) maturation through interaction with human copper chaperone for SOD1 (hCCS), *Proc. Natl. Acad. Sci. U. S. A.*, 109 (2012) 13555–13560.
- [30] Y. Zhao, T. Malinauskas, K. Harlos, E.Y. Jones, Structural insights into the inhibition of Wnt signaling by cancer antigen 5T4/Wnt-activated inhibitory factor 1, *Struct. Lond. Engl.* 1993, 22 (2014) 612–620.
- [31] P.J. Paddison, A.A. Caudy, G.J. Hannon, Stable suppression of gene expression by RNAi in mammalian cells, *Proc. Natl. Acad. Sci. U. S. A.*, 99 (2002) 1443–1448.
- [32] G. Sui, C. Soohoo, E.B. Affar, F. Gay, Y. Shi, W.C. Forrester, Y. Shi, A DNA vector-based RNAi technology to suppress gene expression in mammalian cells, *Proc. Natl. Acad. Sci. U. S. A.*, 99 (2002) 5515–5520.
- [33] E. Seiradake, Y. Zhao, W. Lu, A.R. Aricescu, E.Y. Jones, Production of cell surface and secreted glycoproteins in mammalian cells, *Methods Mol. Biol. Clifton NJ*, 1261 (2015) 115–127.
- [34] C. Donati, F. Cencetti, P. Nincheri, C. Bernacchioni, S. Brunelli, E. Clementi, G. Cossu, P. Bruni, Sphingosine 1-phosphate mediates proliferation and survival of mesoangioblasts, *Stem Cells Dayt. Ohio*, 25 (2007) 1713–1719.
- [35] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, *Methods San Diego Calif*, 25 (2001) 402–408.
- [36] 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.
- [37] S. Reckel, J.J. Lopez, F. Löhr, C. Glaubitz, V. Dötsch, In-cell solid-state NMR as a tool to study proteins in large complexes, *ChemBiochem Eur. J. Chem. Biol.*, 13 (2012) 534–537.
- [38] M. Amarzguioui, J.J. Rossi, D. Kim, Approaches for chemically synthesized siRNA and vector-mediated RNAi, *FEBS Lett.*, 579 (2005) 5974–5981.
- [39] M. Gossen, H. Bujard, Tight control of gene expression in mammalian cells by tetracycline-responsive promoters, *Proc. Natl. Acad. Sci. U. S. A.*, 89 (1992) 5547–5551.
- [40] S. Freundlieb, C. Schirra-Müller, H. Bujard, A tetracycline controlled activation/repression system with increased potential for gene transfer into mammalian cells, *J. Gene Med.*, 1 (1999) 4–12.
- [41] H. Nishijima, T. Yasunari, T. Nakayama, N. Adachi, K. Shibahara, Improved applications of the tetracycline-regulated gene depletion system, *Biosci. Trends*, 3 (2009) 161–167.
- [42] D. No, T.P. Yao, R.M. Evans, Ecdysone-inducible gene expression in mammalian cells and transgenic mice, *Proc. Natl. Acad. Sci. U. S. A.*, 93 (1996) 3346–3351.

## Figures

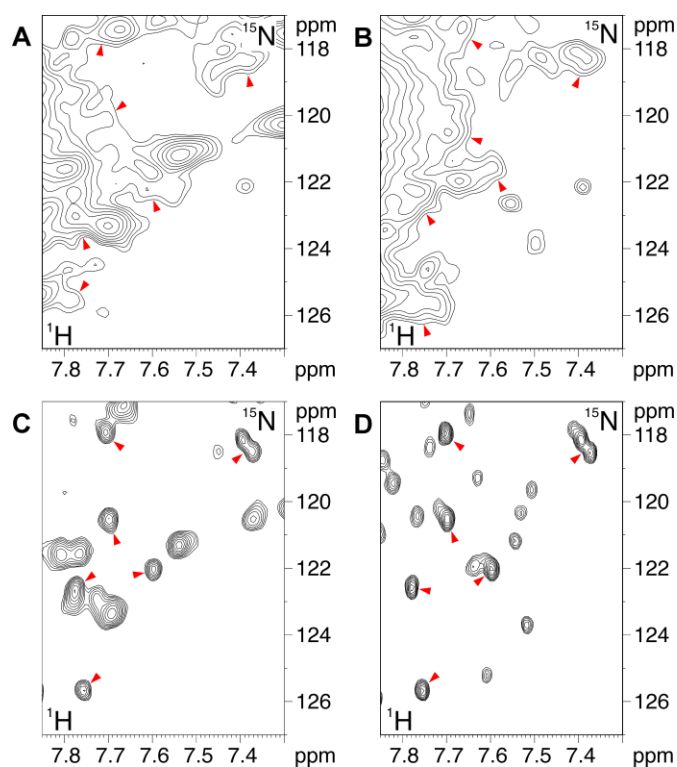


**Fig. 1.** Stable cell line characterization. (A) Western Blot and SDS-PAGE analysis of 12 HEK293T-HAH1 cell lines stably expressing HAH1 selected from 36 single colonies (H1 – H36) obtained after antibiotic selection; the parental cell line (293T) is also shown. (B) Relative HAH1 expression levels obtained by normalizing the Western Blot intensity to the SDS-PAGE lane intensity; the cell line with the highest expression was chosen (H22, marked with an asterisk).

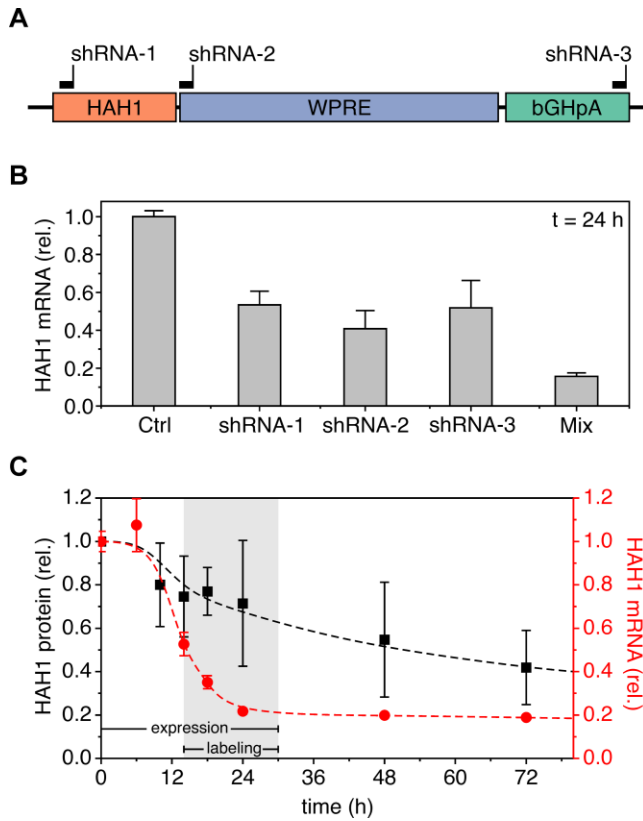




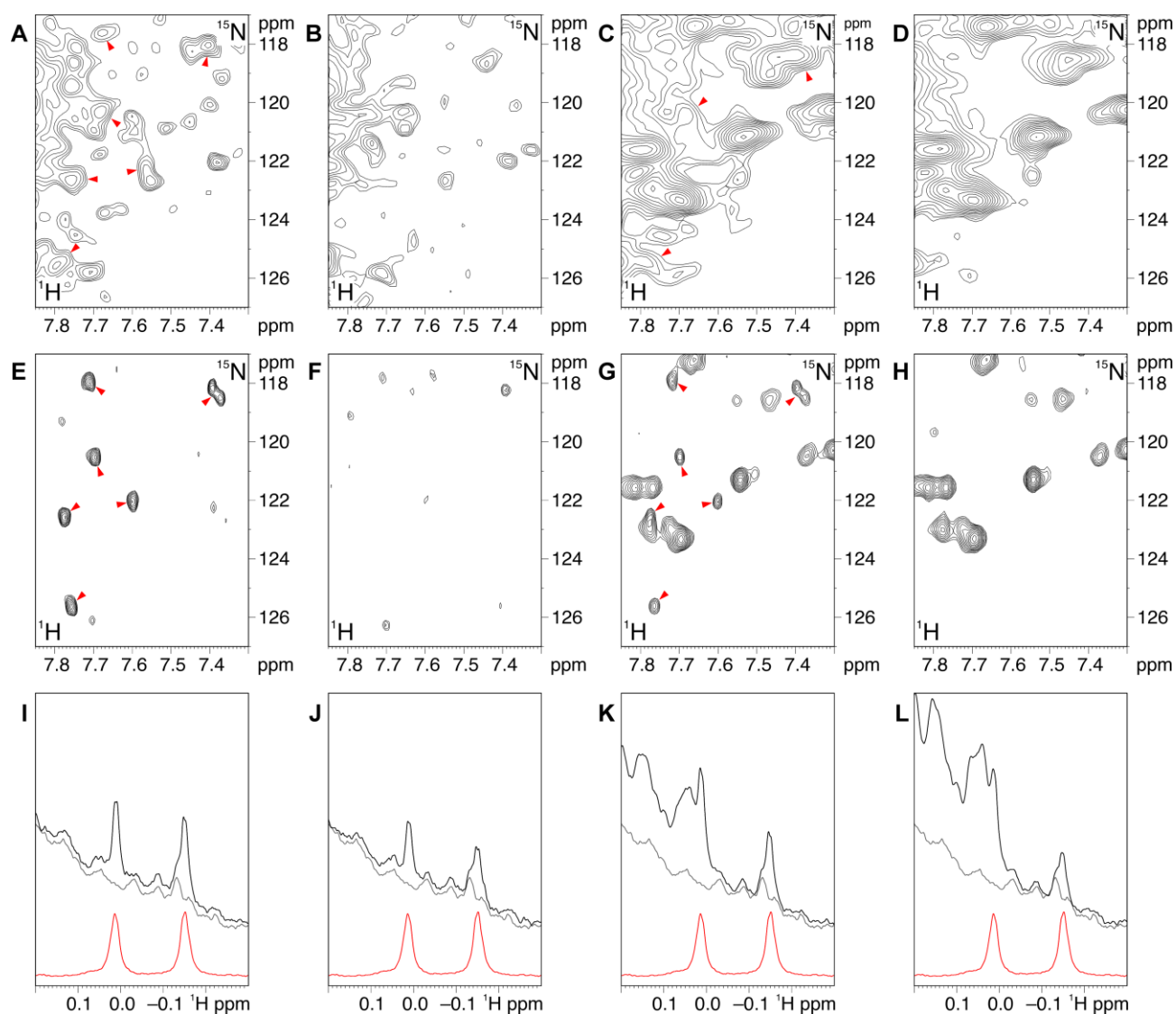
**Fig. 2.** HAH1 stably expressed in human cells can be observed by in-cell NMR. (A) Aliphatic region of the 1D  $^1\text{H}$  NMR spectrum of lysed H22 cells (black), showing signals arising from stably expressed HAH1. The spectrum of lysed HEK293T cells (gray) and the spectrum of purified HAH1 (red) are shown for reference. (B) Selected region of the  $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC spectrum recorded on a sample of  $^{15}\text{N}$ -labeled H22 cells; (C) Same region as (B) of the  $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC spectrum recorded on the corresponding cell lysate. The amide crosspeaks arising from HAH1 are labeled according to the resonance assignment. The spectrum in (C) was further processed to subtract the cellular background. See also Supplementary Figure S1.



**Fig. 3.** Transient expression and labeling of the second protein causes partial labeling of the stably expressed HAH1. Selected region of  $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC spectra of H22 cells (A,B) and the corresponding lysates (C,D) transiently transfected either with SOD1 (A,C) or with TRX1 (B,D). The cells were collected either 24 h (A,C) or 48 h (B,D) following transfection. The molar ratio of SOD1 and TRX1 relative to HAH1 (estimated from the  $^1\text{H}$  NMR spectra of the cell extracts) was  $\sim 0.6$  (A,C) and  $\sim 0.9$  (B,D), respectively. The amide crosspeaks arising from  $^{15}\text{N}$ -labeled HAH1 are indicated by red arrowheads. See also Supplementary Figure S2.



**Fig. 4.** The expression of HAH1 can be efficiently silenced with shRNA-encoding vectors. (A) Schematic representation of the HAH1 gene integrated in the H22 genome; from left to right: open reading frame from HAH1 cDNA (salmon), woodchuck hepatitis virus post-transcriptional regulatory element (WPRE, light blue), bovine growth hormone polyadenylation signal (bGHpA, light green). The recognition sites of each shRNA sequence are shown (black bars). (B) Silencing efficiency of each shRNA-encoding vector, assessed from the relative levels of HAH1 mRNA 24 h after transfection. The combination of shRNA vectors (Mix) gives the highest transfection efficiency (error bars: SEM) (C) Relative levels of HAH1 mRNA (red circles) and HAH1 protein (black squares) as a function of time after transfection, measured by Real Time RT-PCR (error bars: SEM) and Western Blot (error bars: SD, n = 3), respectively. Trends are qualitatively shown as dashed lines. The timing chosen for expression and labeling of the second protein is shown (expression in the white + gray area; labeling in the gray area). See also Supplementary Figure S3.



**Fig. 5.** HAH1 silencing combined with the appropriate labeling strategy improves the labeling selectivity. (A-D) Selected region of  $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC spectra of four samples of H22 cells. Cells were transiently transfected with empty vector (A), shRNA mix (B), SOD1 (C) or shRNA mix + SOD1 (D). The  $^{15}\text{N}$ -medium was provided after 14 h, and the cells were collected 30 h following transfection. (E-H) Selected region of  $^1\text{H}$ - $^{15}\text{N}$  SOFAST-HMQC spectra of the corresponding lysates. The molar ratio of SOD1 relative to HAH1 was  $\sim 0.65$  (c,g) and  $\sim 1.5$  (D,H). The amide crosspeaks arising from  $^{15}\text{N}$ -labeled HAH1 are indicated by red arrowheads. (I-L) Aliphatic region of the 1D  $^1\text{H}$  spectra of the same cell lysates as in (E-H) (black); 1D  $^1\text{H}$  spectra of untransfected parental HEK293T lysate (gray) and purified HAH1 (red) are shown for reference. See also Supplementary Figure S4.