

# Chemical Science

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

## ARTICLE

# Systematic Study of the Dynamics and Half-lives of Newly Synthesized Proteins in Human Cells

Cite this: DOI: 10.1039/x0xx00000x

Weixuan Chen<sup>a</sup>, Johanna M. Smeekens<sup>a</sup> and Ronghu Wu<sup>a\*</sup>Received 00th January 2012,  
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Protein dynamics are essential in regulating nearly every cellular event, and aberrant proteostasis is the source of many diseases. It is extraordinarily difficult to globally study protein dynamics and accurately measure their half-lives. Here we have developed a chemical proteomics method integrating protein labeling, click chemistry and multiplexed proteomics, which overcomes current challenges with existing methods. Labeling with both azidohomoalanine (AHA) and heavy lysine allows us to selectively enrich newly synthesized proteins, clearly distinguish them from existing proteins, and reduce the impact of heavy amino acid recycling. Moreover, multiplexed proteomics enables us to quantify proteins at multiple time points simultaneously, thus increasing the accuracy of measuring protein abundance changes and their half-lives. Systematic investigation of newly synthesized protein dynamics will provide insight into proteostasis and the molecular mechanisms of disease.

## Introduction

Protein homeostasis (proteostasis) is critical for nearly every cellular event. The proper maintenance of proteostasis allows cells to grow and proliferate, respond to environmental cues and nutrient supplies, and defend cells from being attacked by pathogens.<sup>1, 2</sup> Aberrant proteostasis is the source of many diseases, including neurodegenerative, autoinflammatory and cardiovascular diseases.<sup>3-6</sup> The maintenance of cellular proteostasis requires a balance of protein synthesis, trafficking and degradation. The degradation of newly synthesized proteins is an extremely important component of proteostasis because excess proteins are not only a burden for cells, but also a waste of cell energy.

Protein dynamics have been a long-standing interest in the biological and biomedical research fields. Early protein dynamics studies relied on the incorporation of radioactive elements into newly synthesized proteins, and the decay in radioactivity was measured over time to study the protein degradation.<sup>7, 8</sup> Although this method typically obtains information regarding overall protein degradation, the use of radioactive elements could lead to health problems. In order to analyze individual proteins, antibodies are required, which make large-scale analysis difficult. Fluorescence-based methods have also been developed to detect newly synthesized proteins and measure their half-lives.<sup>9, 10</sup> However, they typically require proteins to be tagged with a fluorescence probe and then measured individually, making comprehensive protein analysis time-consuming. Mass spectrometry (MS) combined with pulse-chase stable isotope labeling by amino acids in cell culture (SILAC) is currently a very popular method, and has been extensively applied to investigate protein turnover and degradation.<sup>11, 12</sup> However, there are several challenges with this method. First, it cannot selectively enrich targeted and labeled proteins; therefore, low abundance proteins or proteins with a high degradation rate could be missed for detection at later chase points, resulting in inaccurate abundance measurements. In addition, many existing proteins may interfere

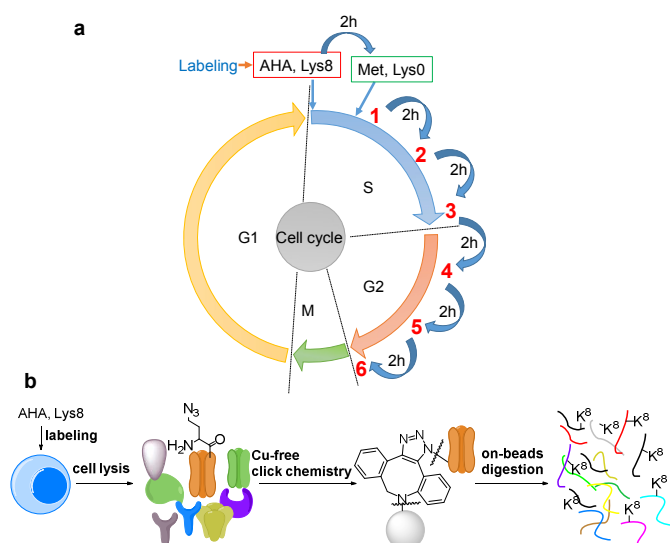
with peptide quantification during MS analysis. Second, a portion of heavy amino acids are always recycled by the cell during the chase step. Although the recycling effects are often ignored, it may have dramatic effects on the accurate quantification of protein degradation.<sup>13</sup> Ideally, the method would be high-throughput and selective for newly synthesized proteins, so that it can be used to accurately measure protein abundance changes.

Here we have developed a chemical proteomics method integrating protein labeling, click chemistry and multiplexed proteomics, which can effectively overcome the challenges with existing methods used to study protein dynamics. Newly synthesized proteins were selectively enriched and their abundance changes were quantified as a function of time. We studied protein dynamics in the cell cycle because it is highly dynamic and well-regulated, and is one of the most important events in biological systems.<sup>14</sup> The S phase, *i.e.* the synthesis phase, is a critical stage of the cell cycle during which DNA is replicated. Over 1,400 newly synthesized proteins were identified in the S phase in HepG2 cells, including three cyclins as expected. These newly synthesized proteins were selectively enriched at multiple time points, and their abundance changes were quantified. The half-lives of many newly synthesized proteins were accurately obtained. This method can be extensively applied to investigate protein dynamics in biological systems.

## Results and discussion

### The principle of the selective enrichment of newly synthesized proteins in the S phase

The introduction of bio-orthogonal groups into proteins has provided an effective method that incorporates a chemical handle for the investigation of protein dynamics and functions.<sup>15-19</sup> Several groups have been incorporated into proteins, including azido and alkylne



**Fig. 1** Experimental principle and procedure: (a) Cells were labeled with AHA and heavy lysine. After cells were labeled by AHA for two hours, normal medium containing methionine and light lysine were used to culture cells; cells were harvested every two hours; (b) Copper-free click chemistry was employed to selectively enrich newly synthesized proteins in the S phase.

groups.<sup>18, 20-27</sup> Azidohomoalanine (AHA), an analog of methionine, was employed to effectively label proteins with an azido group by replacing methionine during protein synthesis.<sup>28, 29</sup> Here, cells were first synchronized in the early S phase by the double thymidine block method,<sup>30</sup> and then double labeled with heavy lysine (+8 D) and AHA, as shown in Fig. 1a. As a result, newly synthesized proteins were labeled by AHA and heavy lysine. AHA labeling enabled the selective enrichment of newly synthesized proteins, while labeling proteins with heavy lysine allowed us to readily distinguish newly synthesized proteins from non-specifically bound existing proteins.

Newly synthesized proteins labeled with AHA were separated and enriched using magnetic beads conjugated with dibenzocyclooctyne (DBCO) (Fig. 1b).<sup>19, 31</sup> The copper-free click reaction between DBCO and the azido group incorporated into proteins is quick, efficient, and does not require any heavy metal ions that may damage the protein backbone. After proteins containing the azido group were covalently bound to beads, stringent washes were performed to remove non-specifically bound proteins. Finally, on-beads digestion was carried out and the resulting peptides were detected by LC-MS/MS.

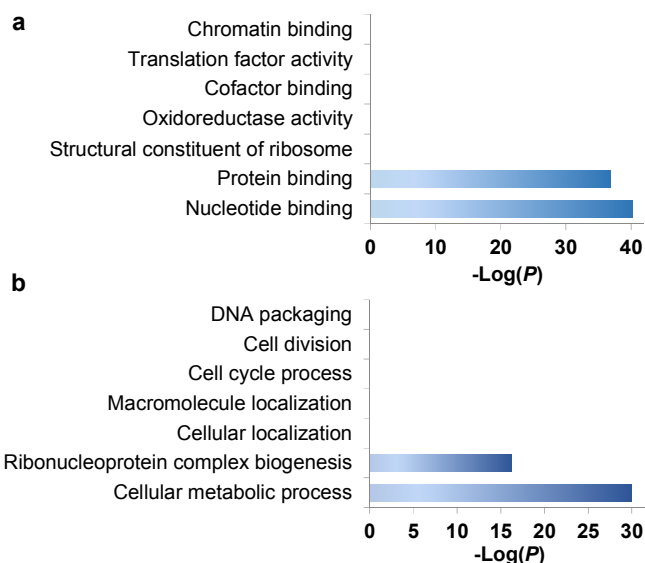
### Identification of newly synthesized proteins in the S phase

During the S phase of the cell cycle, cells prepare for their division in the M phase and their genome is doubled. Each of two functional daughter cells requires not only a complete copy of DNA, but also many essential proteins for survival. More importantly, protein abundance changes are critical for cells to regulate the cell cycle, and a well-known example is the oscillation of cyclin abundances throughout the cell cycle. Although there has been a long-standing interest to determine which proteins are synthesized in the S phase, it is extraordinarily challenging to selectively analyze them for the following reasons. First, newly synthesized proteins are often present at low abundance, which are frequently masked by many high-abundance existing proteins. Second, here we are only

interested in the dynamics of newly synthesized proteins, which need to be clearly distinguished from existing proteins in the cell. Therefore, effective separation and tagging are essential to analyze these newly synthesized proteins prior to MS analysis.

A total of 1,426 newly synthesized proteins were identified in the S phase (Table S1). The newly synthesized proteins in the S phase were clustered based on Gene Ontology (GO) analysis from the Database for Annotation, Visualization and Integrated Discovery (DAVID)<sup>32</sup> and the Protein ANalysis THrough Evolutionary Relationships (PANTHER) classification system,<sup>33, 34</sup> and the results are shown in Fig. 2. Based on molecular function, proteins with nucleotide binding functions were the most highly enriched with a very low  $P$  value of  $5.6 \times 10^{-41}$ , and those with protein binding functions were also very highly enriched ( $P = 1.4 \times 10^{-37}$ ). It is anticipated that during DNA duplication, many proteins related to nucleotide and protein binding are synthesized and active, including helicases, topoisomerases, DNA polymerases and transcription factors. Structural constituents of the ribosome were highly enriched as well because in addition to DNA, many essential proteins are required to double before the M phase, including the ribosome itself.

Based on biological process, proteins related to cellular metabolic process were the most highly enriched, and 56 newly synthesized proteins corresponded to ribonucleoprotein complex biogenesis ( $P = 5.2 \times 10^{-17}$ ). This is consistent with the enriched category of structural constituents of the ribosome discussed above. There were 90 newly synthesized proteins corresponding to the cell cycle process, and 53 proteins related to cell division ( $P = 1.1 \times 10^{-6}$ ). Protein clustering based on cellular component revealed that proteins in both the ribosome and spliceosome were highly enriched with  $P$  values of  $2.7 \times 10^{-23}$  and  $2.5 \times 10^{-10}$ , respectively. These results clearly demonstrated that in addition to DNA duplication, protein synthesis was also highly active in the early stage of the S phase.



**Fig. 2** Clustering of newly synthesized proteins identified in the S phase based on (a) molecular function and (b) biological process.

### Cyclins and transcription factors identified in the S phase

Cyclins are a family of proteins that control the progression of the cell cycle by activating cyclin-dependent kinase (CDK) enzymes, and their concentration varies in a cyclical fashion during the cell cycle. This group of proteins has been extensively studied and information regarding their abundance oscillation throughout the cell

cycle has been well-documented,<sup>35-38</sup> therefore the identification of cyclins in the S phase can be used as an excellent standard to examine the effectiveness of the method.

Cyclin D forms complexes with CDK4 or CDK6, which play key roles in the G1/S transition; the synthesis of cyclin D is initiated during G1. Cyclin E binds to CDK2 in the G1 phase, which is required for the transition from the G1 to S phase.<sup>39</sup> The concentrations of cyclins D and E reach their peaks in the transition of G1/S.<sup>40</sup> As expected, newly synthesized cyclins D and E were not identified in the S phase.

Cyclins A and B are present at the lowest concentrations in the early S phase, based on reports in the literature.<sup>39, 41</sup> Cyclin A has two distinct isoforms: A1 - the embryonic-specific form, and A2 - the somatic form. Cyclin A1 is mainly expressed during meiosis and in early stages of embryogenesis, while cyclin A2 is expressed in dividing somatic cells,<sup>42</sup> including HepG2 cells. Therefore, only cyclin A2 was identified in this work. Cyclin B interacts with CDK1, and is necessary for the progression of cells into and out of the M phase. The amount of cyclin B and the activity of the cyclin B-CDK1 complex increase until mitosis, where they decline dramatically due to degradation of cyclin B.<sup>43</sup> The synthesis of cyclin B starts in the early S phase. As expected, two cyclin B isoforms, B1 and B2, were identified. Cyclin B3 was not identified, as it is a testis specific cyclin expressed in developing germ cells in the testis, but weakly or not expressed in other tissues.<sup>44, 45</sup>

Previously siRNAs and time-lapse epifluorescence microscopy were employed to examine the roles of various candidate mitotic cyclins in chromatin condensation in HeLa cells.<sup>41</sup> Cyclin A2 helps initiate mitosis, and cyclins B1 and B2 are particularly critical for the maintenance of the mitotic state.<sup>41</sup> In this work, we identified only cyclins A2, B1 and B2, which correspond extremely well with reports in the literature,<sup>46</sup> considering the cells used and the newly synthesized proteins expected to be in the early S phase. The identification of these three cyclins firmly indicated that the enrichment is highly effective to selectively enrich newly synthesized proteins.

Transcription factors are another critical group of proteins; they regulate gene expression in cells but are normally present at very low abundance. It is extremely challenging to detect them by using discovery-based proteomics techniques because they are often buried by many high-abundance proteins. Among 1,426 newly synthesized proteins identified in the S phase, 9.6% of them were transcription factors (137 proteins, Table S2), which is slightly greater than the normal distribution of transcription factors (~8% of total genes). This further demonstrated that the current methods can effectively enrich low-abundance newly synthesized proteins.

### Quantify the abundance changes of newly synthesized proteins

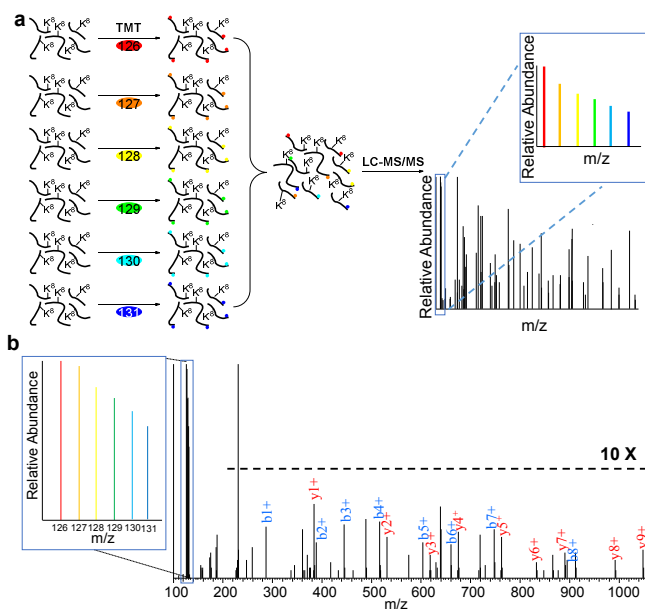
In eukaryotic cells, protein degradation is mainly mediated by two pathways: the ubiquitin-proteasome and lysosomal proteolysis. The half-lives vary greatly, and differing rates of protein degradation are related to the regulation of cellular events. Therefore, accurate measurement of protein abundance changes and their half-lives will provide insight into protein properties and functions, and cellular activities. Here we specifically quantified the abundance changes of these newly synthesized proteins in living cells, and measured their half-lives. After newly synthesized proteins were labeled by AHA and heavy lysine during the early S phase, cells were chased with regular medium containing nocodazole to prevent them from entering the M phase, and subsequently harvested at six time points (0, 2, 4, 6, 8 and 10 hours), as shown in Fig. 3. Similar to the procedure above, newly synthesized proteins bearing the azido group in the S phase were selectively separated and enriched using click chemistry. After digestion of enriched proteins, peptides from each

of the six samples were labeled with different tandem mass tag (TMT) reagents (Fig. 3a). All enriched samples were mixed based on the number of cells harvested at each time point. As a result, the half-lives were not affected by protein dilution resulting from cell growth. The fragments in the tandem mass spectra enabled us to confidently identify peptides, and reporter ion intensities were used to quantify peptide and corresponding protein abundance changes. One example is shown in Fig. 3b; based on the fragmentation, we were able to confidently identify the peptide YDVENC\*LANK# (\* refers to the alkylation of cysteine and # indicates heavy lysine). The identified peptide is from MCM3, which is a DNA replication licensing factor. The abundance changes were measured from the reporter ion intensities, as shown in Fig. 3b. The protein half-lives were calculated based on the protein abundance changes with the most commonly used exponential decay equation:<sup>10</sup>

$$P = P_0 e^{-kt} \quad (1)$$

where  $P_0$  is the intensity of the reporter ion at the first time point,  $P$  is the intensity of the reporter ion at each subsequent time point,  $k$  is the degradation rate constant and  $t$  is time. Therefore, the following equation shows the ratio of the intensities:

$$P/P_0 = e^{-kt} \quad (2)$$

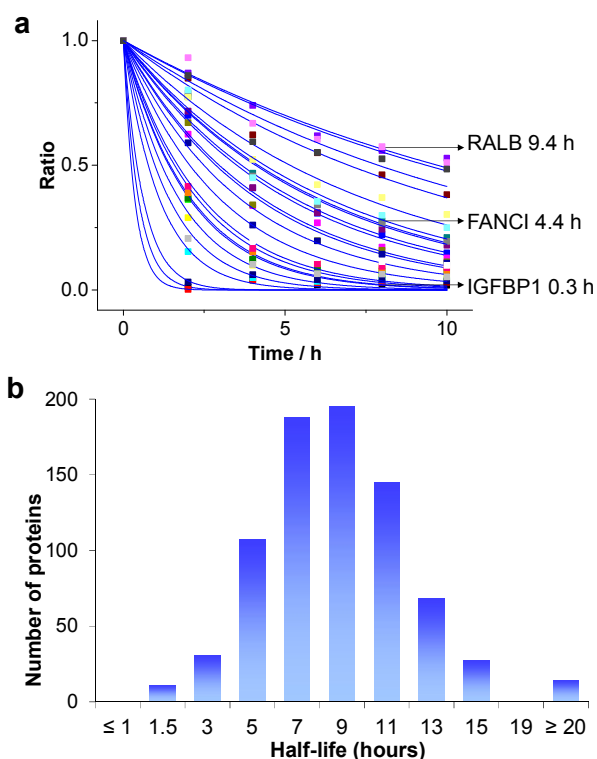


**Fig. 3** Measurement of half-lives of newly synthesized proteins in the S phase by combining copper-free click chemistry enrichment and TMT labeling: (a) Experimental procedure for quantification of proteins; (b) Example mass spectrum showing protein abundance changes.

The ratios at the six time points allowed us to calculate the half-life ( $t_{1/2}$ ) of each protein, *i.e.* the time point when only half of the newly synthesized protein remain in cells. After strict data filtering, the half-lives of 803 newly synthesized proteins in the S phase were obtained (listed in Table S3).

The quantification of peptides based on reporter ion intensities, such as with the TMT method, could result in ratio suppression due to the potential interferences from other ions.<sup>47</sup> However, here it should not be a problem because of several reasons. First, compared to whole cell lysates, the enriched samples were much simpler

because only newly synthesized proteins within the first two hours of the early S phase were selectively separated, which is consistent with the identification of only 1,400 proteins compared to ~10,000 proteins that are typically expressed in cells. Second, fractionating enriched peptides with HPLC made each of the twenty fractions much simpler. In addition, a long gradient in the LC-MS/MS method allowed peptides to be further separated. In our experiment, the combination of higher-energy collisional dissociation (HCD) and fragment detection, including reporter ions, in the high mass accuracy and high resolution Orbitrap cell allowed us to confidently identify and accurately quantify peptides. The ratios of the reporter ions were further calibrated based on the isotopic distribution. If ratio suppression existed, the calculated half-lives would be relatively long. However, the values here are shorter than those obtained with the SILAC pulse-chase method,<sup>10, 13</sup> and are in very good agreement with the values obtained using a MS-independent method.<sup>48</sup>



**Fig. 4** Half-lives of newly synthesized proteins in the S phase: (a) Abundance changes of several proteins and the simulation to determine their half-lives; (b) Distribution of protein half-lives.

#### Distribution of protein half-lives

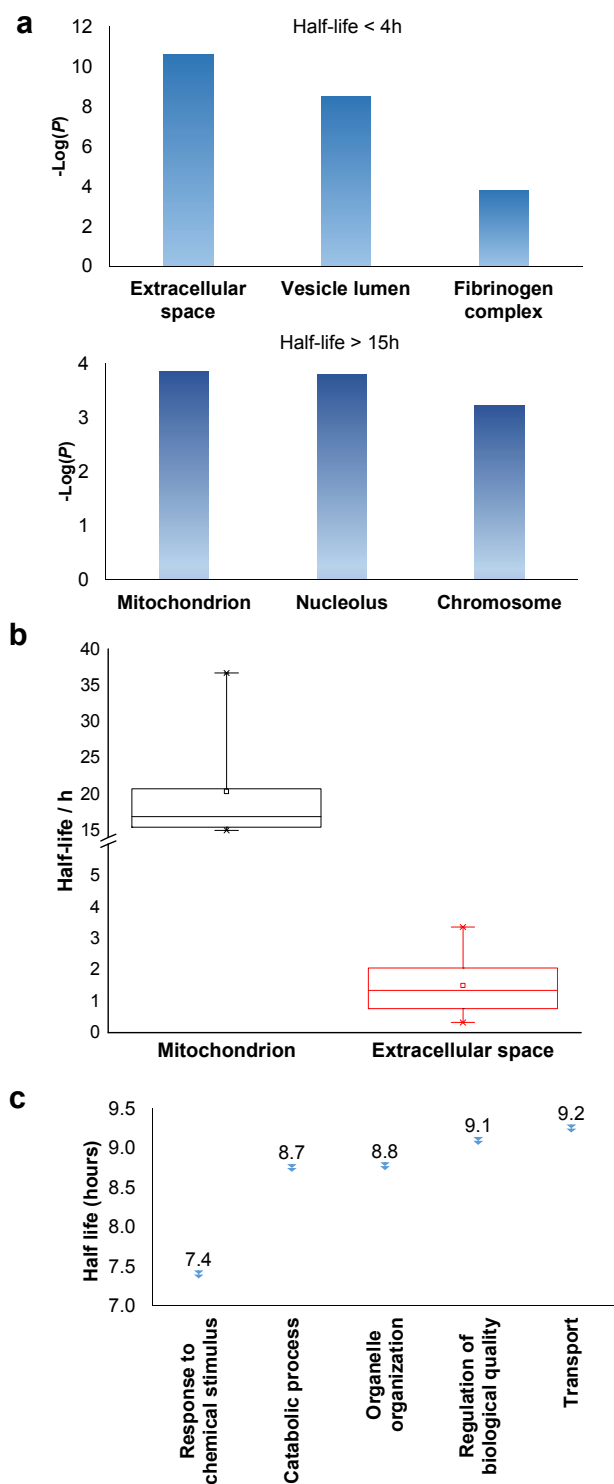
Based on the protein abundance changes measured through the multiplexed proteomics experiment, we can confidently calculate the half-lives of 803 proteins. The abundance changes of several proteins and their half-life traces based on Equation 2 are shown in Fig. 4a. For example, RALB is a GTPase, and is involved in a variety of cellular processes including gene expression, cell migration and proliferation, and membrane trafficking. The half-life was measured to be 9.4 hours in this work. IGFBP1 is insulin-like growth factor-binding protein 1, which is a well-known secreted protein. It has a very short half-life of 0.3 hours inside the cell because it is secreted after synthesis. The distribution of all protein half-lives is shown in

Fig. 4b. The bin at 9 hours (half-lives between 8–10 h) contains the greatest number of proteins. The majority of the proteins quantified have half-lives within the range of 4–14 hours. About 6% of all quantified proteins (49) have half-lives <4 hours, while 51 proteins have long half-lives (>14 hours); the median half-life is 8.7 hours.

Among the proteins with a short half-life in cells, those secreted to the extracellular region were found to be the most highly enriched. For example, F5 (coagulation factor V) has a half-life of 0.8 hours, and APOB has a half-life of 1.3 hours. These values clearly demonstrate that the method worked well because secreted proteins are known to be translocated to the extracellular space shortly after synthesis. Proteins in the vesicle lumen and the fibrinogen complex are also short-lived in cells. The fibrinogen complex is highly soluble, and found in blood plasma and involved in clot formation, which is located outside of the cell.

Among long-lived proteins (half-lives >15 hours), the top three enriched groups of proteins were located in the mitochondrion, nucleolus and chromosome (Fig. 5a). Normally proteins in the mitochondrion are relatively more stable, and 11 long-lived proteins were located in the mitochondrion. Histones are another example of typically long-lived proteins, and here, several histones were identified with half-lives over 10 hours. For example, based on 7 unique and 17 total quantified peptides, the half-life of HIST1H1B was calculated to be 15.7 hours. Previously the half-life of another histone (H2AFV) was calculated to be 13.9 hours using a MS-independent method.<sup>48</sup> As shown in Fig. 5b, the median half-life of proteins located in the mitochondrion is 16.9 hours. In striking contrast, the median half-life of proteins in the extracellular space is over ten times shorter (only 1.3 hours). Several categories of proteins with distinct functions are shown in Fig. 5c. Proteins corresponding to chemical stimuli have the shortest median half-life, while proteins with transportation functions are the longest-lived among these five categories. Eukaryotic initiation factors (eIFs) are proteins involved in the initiation phase of eukaryotic translation, and they form a complex with the ribosomal subunit to regulate protein synthesis. Half-lives of a group of eukaryotic translation initiation factors were listed in Table 1. The median half-life is 7.1 hours and the majority of proteins have half-lives less than 8 hours.

There have already been several reports regarding the measurement of protein half-lives by using SILAC-based pulse-chase combined with MS-based proteomics.<sup>10, 13</sup> The current results are quite different from those obtained from the SILAC methods. In pulse-chase SILAC experiments, heavy amino acid recycling is an inherent problem.<sup>13</sup> By using the current method, AHA and heavy lysine can still be recycled, but the possibility of recycling both AHA and heavy lysine in a single identified peptide is extremely low. Furthermore, the protein half-lives obtained from the pulse-chase SILAC experiments were based on the ratios measured independently by MS, and errors are more likely as a result.<sup>10</sup> Here we quantified proteins at six time points, and the peptide abundance changes at these time points were measured in a single MS<sup>2</sup> spectrum. The half-lives of proteins calculated from accurately measured protein abundance changes at six points are expected to be more reliable. More importantly, in our work, we can selectively enrich those newly synthesized proteins, which is beyond the reach of normal pulse-chase SILAC experiments. One drawback of AHA labeling is that AHA could affect protein folding and stability, but the comparison of our results with those reported using a MS-independent method suggests that this is less likely to be true. Protein half-lives reported here are very similar to those measured by



**Fig. 5** Clustering of proteins with different half-lives: (a) Enriched proteins among those with short and long half-lives; (b) Comparison of the median half-lives among proteins located in mitochondrion and extracellular space (the half-lives of proteins in the extracellular space refer to the intracellular half-lives); (c) The median half-lives for proteins with different functions.

a MS-independent method, *i.e.* the bleach-chase experiment with fluorescence detection.<sup>48</sup> For example, the half-lives of EIF2S2, EIF4A1 and EIF5B were calculated to be 6.3, 6.0 and 8.2 hours in this work, respectively, and they were reported to be 8.3, 7.3 and 8.9 hours with the bleach-chase method in literature.<sup>48</sup> Overall, the median half-life of 100 proteins measured in the bleach-chase experiment was 8.2 hours, which is in excellent agreement with the median half-life of 8.7 hours for 803 newly synthesized proteins in the S phase calculated here.<sup>48</sup>

## Conclusions

The cell cycle is an extremely complex and well-regulated event that is essential for life. Multiple factors contribute to the regulation of the cell cycle, including protein abundance changes, modifications and interactions. Many aspects remain unknown, including which proteins are synthesized in each stage of the cell cycle and the dynamics of these proteins. In this work, a novel chemical proteomics method integrating bio-orthogonal amino acid labeling and copper-free click chemistry was developed to selectively separate newly synthesized proteins during the early S phase of the cell cycle. More than 1,400 newly synthesized proteins were identified in the S phase, including cyclin A2, B1 and B2, which is very consistent with reports in the literature. Because this method selectively targets newly synthesized proteins, it enabled us to identify low-abundance proteins, including 137 transcription factors. By combining this method with multiplex quantitative proteomics techniques, the abundance changes of these newly synthesized proteins in the S phase were quantified after being selectively separated at six different time points and their half-lives were calculated. The median half-life of over 800 quantified proteins was 8.7 hours, which corresponds very well with the median half-life of 8.2 hours for 100 proteins in previous research. Proteins with shorter half-lives in cells are typically secreted, while those in the mitochondrion and nucleus are long-lived. The identification of newly synthesized proteins in the early S phase and the measurement of their half-lives provide very useful information to further understand the molecular mechanisms of the cell cycle, and protein properties and functions.

Although bio-orthogonal amino acid labeling, click chemistry and TMT tagging are well-documented, here we have designed and tailored a novel method integrating these approaches to globally study protein dynamics and measure their half-lives, which overcomes problems associated with existing methods. The current method has several advantages. *First*, double labeling with AHA and heavy lysine allows us to selectively enrich newly synthesized proteins and clearly distinguish them from existing proteins. *Second*, double labeling can significantly reduce the impact of heavy amino acid recycling on the measurement of protein abundance changes in the pulse-chase SILAC experiments. *Third*, interferences from protein dilution due to cell growth is eliminated because cells were arrested against cell division, and newly synthesized protein abundance changes are based on the same number of cells, not amount of proteins. *Fourth*, multiplexed proteomics enabled us to quantify proteins at several time points simultaneously, therefore increasing the accuracy of measuring protein abundance changes and the corresponding half-lives. *Fifth*, the high throughput MS-based experiment enabled us to study protein dynamics on a global level, instead of a one-to-one fashion. *Lastly*, over eight hundred protein half-lives were accurately measured, which provides valuable information to the biological community. This method can be extensively applied to investigate newly synthesized proteins and their dynamics in the biological and biomedical research fields.

**Table 1. Examples of half-lives measured for eukaryotic translation initiation factors.**

Gene Symbol	UniProt ID	Protein half-life		Annotation
		This work	Previous work <sup>48</sup>	
EIF1AX	P47813	7.7		Eukaryotic translation initiation factor 1A
EIF2S2	P20042	6.3	8.3	Eukaryotic translation initiation factor 2 subunit 2
EIF3E	P60228	8.3		Eukaryotic translation initiation factor 3 subunit E (eIF3e)
EIF3H	O15372	7.0		Eukaryotic translation initiation factor 3 subunit H (eIF3h)
EIF3L	Q9Y262	12.2		Eukaryotic translation initiation factor 3 subunit L (eIF3l)
EIF4A1	P60842	6.0	7.3	Eukaryotic initiation factor 4A-I
EIF4A2	Q14240	5.7		Eukaryotic initiation factor 4A-II
EIF4A3	P38919	6.4		Eukaryotic initiation factor 4A-III
EIF4B	P23588	7.4		Eukaryotic translation initiation factor 4B (eIF-4B)
EIF4G1	Q04637	6.8		Eukaryotic translation initiation factor 4 gamma 1
EIF4G2	P78344	9.6		Eukaryotic translation initiation factor 4 gamma 2
EIF5B	O60841	8.2	8.9	Eukaryotic translation initiation factor 5B

## Acknowledgements

This work is supported by the National Science Foundation (CAREER Award, CHE-1454501).

## Notes and references

<sup>a</sup> School of Chemistry and Biochemistry and the Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, Georgia 30332, USA

\* Correspondence: Ronghu.Wu@chemistry.gatech.edu, Phone: 404-385-1515, Fax: 404-894-7452

Electronic Supplementary Information (ESI) available: [Supplementary Text and References, six Supplementary figures, and thirteen Supplementary Tables]. See DOI: 10.1039/b000000x/

- W. E. Balch, R. I. Morimoto, A. Dillin and J. W. Kelly, *Science*, 2008, 319, 916-919.
- E. T. Powers, R. I. Morimoto, A. Dillin, J. W. Kelly and W. E. Balch, in *Annual Review of Biochemistry*, Annual Reviews, Palo Alto, 2009, vol. 78, pp. 959-991.
- J. Labbadia and R. I. Morimoto, *Annu. Rev. Biochem.*, 2015, 84, 435-464.
- S. Jaisson and P. Gillery, *Diabetologia*, 2014, 57, 1517-1527.
- A. Brehm and E. Kruger, *Semin. Immunopathol.*, 2015, 37, 323-333.
- T. Morawe, C. Hiebel, A. Kern and C. Behl, *Mol. Neurobiol.*, 2012, 46, 41-54.
- R. D. Conde and O. A. Scornik, *Biochem. J.*, 1976, 158, 385-390.
- M. Llovera, F. J. Lopez-Soriano and J. M. Argiles, *J. Natl. Cancer Inst.*, 1993, 85, 1334-1339.
- A. A. Cohen, N. Geva-Zatorsky, E. Eden, M. Frenkel-Morgenstern, I. Issaeva, A. Sigal, R. Milo, C. Cohen-Saidon, Y. Liron, Z. Kam, L. Cohen, T. Danon, N. Perzov and U. Alon, *Science*, 2008, 322, 1511-1516.
- B. Schwanhauser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen and M. Selbach, *Nature*, 2011, 473, 337-342.
- J. M. Pratt, J. Petty, I. Riba-Garcia, D. H. L. Robertson, S. J. Gaskell, S. G. Oliver and R. J. Beynon, *Molecular & Cellular Proteomics*, 2002, 1, 579-591.
- I. Fierro-Monti, J. Racle, C. Hernandez, P. Waridel, V. Hatzimanikatis and M. Quadroni, *PLoS One*, 2013, 8, 17.
- F.-M. Boisvert, Y. Ahmad, M. Gierliński, F. Charrière, D. Lamont, M. Scott, G. Barton and A. I. Lamond, *Molecular & Cellular Proteomics*, 2012, 11, 10.1074/mcp.M1111.011429.
- T. Ly, Y. Ahmad, A. Shlien, D. Soroka, A. Mills, M. J. Emanuele, M. R. Stratton and A. I. Lamond, *eLife*, 2014, 3.
- D. C. Dieterich, A. J. Link, J. Graumann, D. A. Tirrell and E. M. Schuman, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, 103, 9482-9487.
- M. M. Zhang, L. K. Tsou, G. Charron, A. S. Raghavan and H. C. Hang, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, 107, 8627-8632.
- K. H. Wang, A. Sachdeva, D. J. Cox, N. W. Wilf, K. Lang, S. Wallace, R. A. Mehl and J. W. Chin, *Nat. Chem.*, 2014, 6, 393-403.
- S. T. Dieck, L. Kochen, C. Hanus, M. Heumueller, I. Bartnik, B. Nassim-Assir, K. Merk, T. Mosler, S. Garg, S. Bunse, D. A. Tirrell and E. M. Schuman, *Nat. Methods*, 2015, 12, 411-414.
- J. M. Baskin, J. A. Prescher, S. T. Laughlin, N. J. Agard, P. V. Chang, I. A. Miller, A. Lo, J. A. Codelli and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, 104, 16793-16797.
- K. L. Kiick, E. Saxon, D. A. Tirrell and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, 99, 19-24.
- W. P. Heal, B. Jovanovic, S. Bessin, M. H. Wright, A. I. Magee and E. W. Tate, *Chem. Commun.*, 2011, 47, 4081-4083.
- E. M. Sletten and C. R. Bertozzi, *Accounts Chem. Res.*, 2011, 44, 666-676.
- L. Davis and J. W. Chin, *Nat. Rev. Mol. Cell Biol.*, 2012, 13, 168-182.
- M. Grammel and H. C. Hang, *Nat. Chem. Biol.*, 2013, 9, 475-484.
- A. Mahdavi, J. Szychowski, J. T. Ngo, M. J. Sweredoski, R. L. J. Graham, S. Hess, O. Schneewind, S. K. Mazmanian and D. A. Tirrell, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, 111, 433-438.
- T. S. Elliott, F. M. Townsley, A. Bianco, R. J. Ernst, A. Sachdeva, S. J. Elsasser, L. Davis, K. Lang, R. Pisa, S. Greiss, K. S. Lilley and J. W. Chin, *Nat. Biotechnol.*, 2014, 32, 465-U186.
- W. X. Chen, J. M. Smeeckens and R. H. Wu, *Chemical Science*, 2015, 6, 4681-4689.
- D. C. Dieterich, J. J. Lee, A. J. Link, J. Graumann, D. A. Tirrell and E. M. Schuman, *Nat. Protoc.*, 2007, 2, 532-540.
- A. J. M. Howden, V. Geoghegan, K. Katsch, G. Efstathiou, B. Bhushan, O. Boutourel, B. Thomas, D. C. Trudgian, B. M. Kessler, D. C. Dieterich, B. G. Davis and O. Acuto, *Nat. Methods*, 2013, 10, 343-+.
- J. Harper, in *Cell Cycle Control*, eds. T. Humphrey and G. Brooks, Humana Press, 2005, vol. 296, ch. 7, pp. 157-166.
- J. M. Smeeckens, W. X. Chen and R. H. Wu, *J Am Soc Mass Spectr*, 2015, 26, 604-614.
- D. W. Huang, B. T. Sherman and R. A. Lempicki, *Nat. Protoc.*, 2009, 4, 44-57.
- H. Y. Mi, B. Lazareva-Ulitsky, R. Loo, A. Kejariwal, J. Vandergriff, S. Rabkin, N. Guo, A. Muruganujan, O. Doremieux, M. J. Campbell, H. Kitano and P. D. Thomas, *Nucleic Acids Res.*, 2005, 33, D284-D288.
- P. D. Thomas, M. J. Campbell, A. Kejariwal, H. Y. Mi, B. Karlak, R. Daverman, K. Diemer, A. Muruganujan and A. Narechania, *Genome Res.*, 2003, 13, 2129-2141.
- T. Ishidate, A. Elewa, S. Kim, C. C. Mello and M. Shirayama, *Cell Cycle*, 2014, 13, 1384-1391.

## Journal Name

36. D. O. Morgan, *Annu. Rev. Cell Dev. Biol.*, 1997, 13, 261-291.
37. W. Y. Tam and M. C. Lai, *Cell Div.*, 2011, 6.
38. J. Bloom and F. R. Cross, *Nat. Rev. Mol. Cell Biol.*, 2007, 8, 149-160.
39. M. Malumbres and M. Barbacid, *Nat. Rev. Cancer*, 2009, 9, 153-166.
40. K. T. Siu, M. R. Rosner and A. C. Minella, *Cell Cycle*, 2012, 11, 57-64.
41. D. Gong and J. E. Ferrell, *Molecular Biology of the Cell*, 2010, 21, 3149-3161.
42. C. H. Yam, T. K. Fung and R. Y. C. Poon, *Cellular and Molecular Life Sciences: CMLS*, 2002, 59, 1317-1326.
43. A. Hershko, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 1999, 354, 1571-1576.
44. J.-C. Lozano, E. Perret, P. Schatt, C. Arnould, G. Peaucellier and A. Picard, *Biochemical and Biophysical Research Communications*, 2002, 291, 406-413.
45. T. B. Nguyen, K. Manova, P. Capodiecì, C. Lindon, S. Bottega, X.-Y. Wang, J. Refik-Rogers, J. Pines, D. J. Wolgemuth and A. Koff, *Journal of Biological Chemistry*, 2002, 277, 41960-41969.
46. J. D. Moore, *Nat. Rev. Cancer*, 2013, 13, 201-208.
47. S. Y. Ow, M. Salim, J. Noirel, C. Evans, I. Rehman and P. C. Wright, *J. Proteome Res.*, 2009, 8, 5347-5355.
48. E. Eden, N. Geva-Zatorsky, I. Issaeva, A. Cohen, E. Dekel, T. Danon, L. Cohen, A. Mayo and U. Alon, *Science*, 2011, 331, 764-768.