

# A single polyphosphate kinase-based NTP regeneration system driving cell-free protein synthesis

Received  
Accepted

DOI:

Po-Hsiang Wang,<sup>a,†</sup> Kosuke Fujishima,<sup>ab,‡,\*</sup> Samuel Berhanu,<sup>a#</sup> Yutetsu Kuruma,<sup>ac#</sup> Tony Z. Jia,<sup>ad</sup> Anna N. Khusnutdinova,<sup>e</sup> Alexander F. Yakunin<sup>ef</sup> and Shawn E. McGlynn<sup>ad\*</sup>

**Current reconstituted cell-free protein synthesis systems regenerate the ATP and GTP required for peptide bond formation using phosphocreatine and three coupled kinase reactions. Here, we replace the three-kinase system with a single polyphosphate kinase, resulting in a comparable protein yield and enabling functional expression of a heat-sensitive enzyme at 37°C.**

A major problem in contemporary biochemistry is that a sizable fraction of proteins coded in sequenced genomes do not have functional annotation.<sup>1,2</sup> A high-throughput protein synthesis platform would facilitate functional characterization and annotation, potentially to the extent of total proteome synthesis.<sup>3</sup> Cell-free protein synthesis systems derived from cell lysates or reconstituted purified components (*i.e.* the PURE system) have yielded fundamental biochemical insights<sup>4,5</sup> and also allow an opportunity for high-throughput expression of a multitude of proteins.<sup>6,7</sup> As opposed to *in vivo* cell-free protein expression in organisms like *E. coli*, systems reconstituted from purified components allow protein synthesis in an environment free of destructive nucleases and proteases,<sup>7</sup> along with the flexibility for a wide range of environmental adjustments (*e.g.*, adjustment of pH and ionic strength or addition of detergents and chaperones) that can be difficult or impossible in the *in vivo* systems. These features enable functional expression of hetero-oligomeric protein subunits, proteins with non-canonical amino acids, and toxic proteins with high levels of homogeneity.<sup>8,9</sup> Additionally, the PURE system allows protein expression using mRNA, thereby bypassing multiple steps of plasmid cloning and organism culturing. These properties substantially shorten the procedure for protein production and prevent the occurrence of undesirable gene mutations. Furthermore, the PURE

system can be applied to studying complex protein synthesis<sup>10</sup> and also protein synthesis in non-traditional, *i.e.*, synthetic, settings such as artificial cells.<sup>8,11</sup>

In mRNA translation, formation of each peptide bond requires one ATP for tRNA aminoacylation by aminoacyl-tRNA synthetases (producing an AMP) and two GTPs for elongation and translocation of polypeptides (producing two GDPs).<sup>12</sup> In contrast to cellular organisms that have intrinsic biochemical pathways to regenerate nucleoside triphosphates (NTP), regeneration of ATP and GTP in the PURE system is accomplished by three coupled reactions involving creatine kinase, myokinase and nucleoside diphosphate kinase, which transfer the phospho-moiety of phosphocreatine to ADP, AMP and GDP, respectively (**Fig. 1A**).<sup>7</sup> Harmonizing the activities of the three kinases, each with distinct kinetic properties, is challenging because myokinase and nucleoside diphosphate kinase each consume ATP produced by creatine kinase. Furthermore, creatine kinase-mediated ADP phosphorylation is reversible under assay conditions (BRENDA Database EC 2.7.3.2). Accordingly, the PURE system generally employs a high concentration of phosphocreatine for effective NTP regeneration,<sup>13</sup> which can be cost-prohibitive screening for high-throughput screening of protein function and for large-scale industrial protein production *i.e.* antibody production.<sup>12,14</sup> Developing a one enzyme-based NTP regeneration platform using a more cost-effective substrate would significantly expand the use of the PURE system.

In this study, we incorporated inorganic polyphosphates (polyP) (mixture with an average length of 25; 0.01–0.05 USD per gram) and a polyP kinase family II (PPK2) enzyme for NTP regeneration in the PURE system, establishing a simple and cost-effective mean for driving cell-free protein synthesis. PPK2 is a family of phosphotransferases capable of transferring the phospho-moiety of polyP to nucleoside mono- and diphosphates using metal cations (*e.g.* Mg<sup>2+</sup>) as the cofactor.<sup>15–17</sup> Class III PPK2s are capable of phosphorylating both nucleoside mono- or diphosphates, and have been applied to *in vitro* ATP regeneration<sup>18</sup>. PolyP has been used to supplement a PURE system to enhance ATP regeneration; nevertheless, NTP regeneration in the PURE system requires the amalgamation of multiple kinases and substrates.<sup>19,20</sup> Recently, a novel, highly active class III PPK2

<sup>a</sup> Earth-Life Science Institute, Tokyo Institute of Technology, Tokyo, Japan.

<sup>b</sup> Graduate School of Media and Governance, Keio University, Fujisawa, Japan

<sup>c</sup> JST, PRESTO, Saitama, Japan

<sup>d</sup> Blue Marble Space Institute of Science, Seattle, Washington, U.S.A

<sup>e</sup> Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, Ontario, Canada.

<sup>f</sup> Centre for Environmental Biotechnology, Bangor University, Bangor, Wales, United Kingdom.

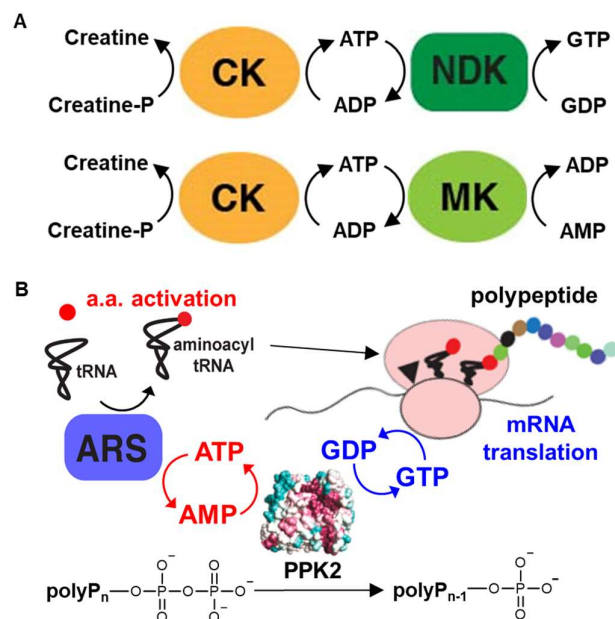
† P.-H.W. and K.F. contributed equally to this work.

\* To whom correspondence should be addressed: Earth-Life Science Institute, Tokyo Institute of Technology, Tokyo 152-8550, Japan. E-mail: [fujii@elsi.jp](mailto:fujii@elsi.jp); [mcglynn@elsi.jp](mailto:mcglynn@elsi.jp).

# Current affiliation: Institute for Extra-Cutting-Edge Science and Technology Avant-garde Research (X-star), JAMSTEC, Yokosuka, Japan

† Electronic Supplementary Information (ESI) available: Materials, methods, Fig. S1–S4, See DOI:

from *Cytophaga hutchinsonii*, CHU0107 (ENA accession ABG57400), capable of phosphorylating AMP, ADP, GMP and GDP to the corresponding nucleoside diphosphates and triphosphates was structurally and biochemically characterized.<sup>15</sup> For CHU0107, the  $K_m$  for AMP ( $0.60 \pm 0.10$  mM) and GDP ( $2.75 \pm 0.55$  mM) were within an order of magnitude, suggesting a potential for simultaneous phosphorylation of AMP to ATP via ADP and GDP to GTP.<sup>15</sup> Moreover, the truncated variant of CHU0107 without the C-terminal tail (Leu285-Asp305; denoted as CHU0107t) exhibited a two-fold higher activity in phosphorylation of both AMP (45 U/mg) and GDP (10 U/mg) compared to the wild-type protein, rendering CHU0107t a potent alternative for NTP regeneration in the PURE system (**Fig. 1B**).



**Fig. 1** NTP regeneration systems required for cell-free protein synthesis. **(A)** The conventional NTP regeneration system functions with phosphocreatine and three coupled kinases: CK, creatine kinase; MK, myokinase; NDK, nucleoside diphosphate kinase. **(B)** The single polyP kinase-based NTP regeneration system for cell-free protein synthesis developed in this study. ARS, aminoacyl-tRNA synthetase; PPK2, polyP kinase family II.

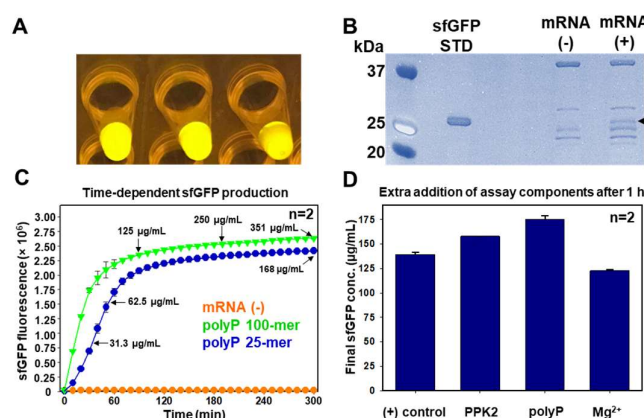
We first determined the optimal polyP concentration for NTP regeneration using customized PURE system (PUREfrex®, GeneFrontier Corp.) reagents free of phosphocreatine, ATP, GTP and the three kinases. We observed polyP-dependent AMP phosphorylation to ATP via ADP,<sup>15</sup> and ATP production was proportional to the polyP concentration in the range from 10 mM to 50 mM (based on the molar content of phosphate monomer) (**Fig. S1A**). Therefore, we used 50 mM polyP in the PURE system for optimization of CHU0107t activity.

Next, we applied CHU0107t for simultaneous polyP-dependent regeneration of ATP and GTP in the PURE system. The  $K_m$  of CHU0107t with AMP ( $22 \pm 3$   $\mu$ M) or with GDP ( $0.29 \pm 0.02$  mM) measured under PURE system conditions were at least one order-of-magnitude lower than the reported values for the wild-type CHU0107 (**Fig. S1B**).

The two-fold higher  $k_{cat}$  of CHU0107t with AMP (45/s) or GDP (7.7/s) compared to that of wild-type CHU0107 under the PURE system conditions is consistent with the previously reported data.<sup>15</sup>

For optimizing the NTP regeneration in the polyP-based PURE system, we selected the superfolder green fluorescent protein (sfGFP) (~26.5 kDa).<sup>21</sup> Production of sfGFP was estimated using fluorescence emission at 518 nm in round bottom 96-well microplates (20 microliter reactions) in a real-time PCR system. We first performed sfGFP synthesis using 3 mM GTP and 0.2 mM ATP since the  $K_m$  of CHU0107t with AMP is ~15 times lower than that with GDP (the concentration for both ATP and GTP in the commercial PURE system is 3 mM). The polyP-based PURE system showed a time-dependent sfGFP production against the mRNA-free negative control (**Fig. S2A**). Moreover, most of sfGFP was produced within the first five hours of reaction. The sfGFP production is consistent with the visible green colour of reaction mixtures after incubation (**Fig. 2A**) and was also confirmed using SDS-PAGE (**Fig. 2B**). In order to improve the protein yield, we tested multiple concentrations of  $Mg^{2+}$ , polyP and CHU0107t in the polyP-based PURE system, and determined the optimal concentrations to be 18 mM  $Mg^{2+}$ , 1.0  $\mu$ M CHU0107t, and 50 mM polyP (**Figs. S2A–C**). The optimal concentrations of ATP and GTP were found to be 0.1 mM and 1.0 mM, respectively (**Figs. S2DE**). After optimization, the polyP-based PURE system produced sfGFP to a final concentration corresponding to  $168 \pm 4$   $\mu$ g/mL (determined using purified recombinant sfGFP standard), which is comparable to the values reported for commercial PURE systems (0.1–0.4 mg/mL).<sup>13,22,23</sup>

Based on the observation that sfGFP production in the polyP-based PURE system becomes much slower after one hour of incubation (note that the fluorescence emission intensity of sfGFP is not linear at concentration >10  $\mu$ g/mL) (**Fig. 2C**), we hypothesized that this might be due to either the loss-of-function of CHU0107t or depletion of polyP over time. We also tested the potential influence of  $Mg^{2+}$  chelation by polyP oligomers during incubation. Therefore, after one hour of incubation we added either CHU0107t (1  $\mu$ M),  $Mg^{2+}$  (3 mM) or polyP (10 mM) to the reaction mixtures (values are the final concentrations of the added component) (**Fig. S2F**). After additional incubation for 70 min, polyP addition enhanced the sfGFP production by ~25% compared to the (+) control (**Fig. 2D**), while the CHU0107t addition showed minor effect. In contrast, increasing  $Mg^{2+}$  concentration hampered the production of sfGFP. Thus, our data suggest that the reduction in sfGFP production in the polyP-based PURE system after two hours might be due to both polyP depletion and  $Mg^{2+}$  chelation by phosphate. We therefore analysed time-dependent polyP consumption in the reaction mixtures using a toluidine blue-stained TBE-urea gel (6%).<sup>24</sup> We found that after two hours of incubation, the longer-chain fraction of the added polyP 25-mer mixture was mostly consumed, while the shorter-chain fraction remained unchanged (**Fig. S3**), suggesting that CHU0107t has a preference to long-chain polyP as substrate in the polyP-based PURE system.

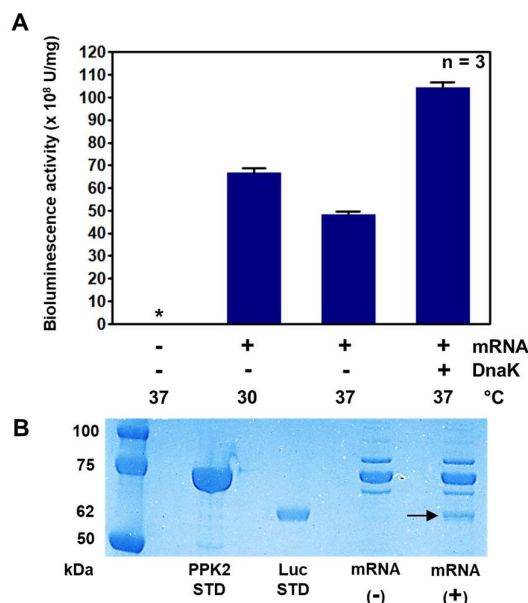


**Fig. 2** Small-scale sfGFP synthesis (20  $\mu$ L) with the polyP-based PURE system in a 96-well microplate. **(A)** sfGFP production in triplicate trials after five hours of incubation (fluorescence emission of sfGFP in a blue light gel illuminator). **(B)** SDS-PAGE analysis of the polyP-based PURE system reaction mixtures after incubation (untrimmed gel image available in **ESI**) **(C)** sfGFP production in the reaction mixtures without sfGFP mRNA (mRNA (-)), as well as with sfGFP mRNA and polyP in different lengths (polyP 25-mer or polyP 100-mer) as the energy source. The arrows point to fluorescence intensities of purified recombinant sfGFP standards (31.3–250  $\mu$ g/mL) and the final concentration of the reaction mixtures. Note that the sfGFP fluorescence intensity is not linear above 10  $\mu$ g/mL and the final concentrations were measured by dilution to the linear range (**ESI**). **(D)** Final sfGFP concentrations in the reaction mixtures that received an extra addition of polyP, CHU0107t (PPK2), Mg<sup>2+</sup> or only buffer ((+) control) during incubation. All the component stocks were dissolved in the same PURE system buffer and were added to reaction mixtures that were incubated for 60 min, and the reaction mixtures were further incubated for 70 min. Assays were performed in duplicate trials.

Therefore, we replaced the original polyP 25-mer mixture with a polyP 100-mer mixture (Kerafast®) at the same concentration (50 mM based on the molar content of phosphate monomer) as the energy source in the polyP-based PURE system. With longer polyP substrates, the system produced over 350  $\mu$ g of sfGFP per ml ( $351 \pm 3$   $\mu$ g/mL), a two-fold increase compared to the short chain polyP (**Fig. 2C**). Moreover, during 1.5–5 hours of incubation, sfGFP production in the PURE system with the polyP 100-mer increased from 125 to 351  $\mu$ g/mL, whereas that in the PURE system with the 25-mer mixture only achieved a final concentration of 168  $\mu$ g/mL (**Fig. 2C**). Enhanced production with the polyP 100-mer is consistent with the mRNA-stabilizing effect of long-chain polyP described previously.<sup>19</sup>

The versatility of the polyP-based regeneration system in protein synthesis applications was also tested using a heat-sensitive enzyme, *Photinus pyralis* firefly luciferase (~62 kDa), as a model protein. The *P. pyralis* firefly luciferase is unstable at temperatures greater than 30°C,<sup>25</sup> whereas the polyP-based PURE system is generally operated at 37°C. We were able to detect firefly luciferase bioluminescence from reaction mixtures incubated at 37°C for three hours (**Fig. 3A**), and the polyP-based system produced active firefly luciferase with a final concentration corresponding to ~50  $\mu$ g/mL of a

commercial firefly luciferase standard ( $1 \times 10^{11}$  U/mg), which is ~30% lower than the final concentration of active firefly luciferase in reaction mixtures incubated at 30°C (**Fig. S4**). The firefly luciferase production was confirmed using SDS-PAGE (**Fig. 3B**). Moreover, supplementing native *E. coli* chaperones DnaKJ and GrpE to the polyP-based PURE system doubled the active firefly luciferase production at 37°C (**Fig. 3A**).



**Fig. 3** Functional expression of *P. pyralis* firefly luciferase in the polyP-based PURE system. **(A)** Average bioluminescence activity (at 562 nm) in the supernatants of polyP-based PURE system reaction mixtures incubated for three hours in the presence (mRNA(+)) and absence (mRNA(-)) of firefly luciferase mRNA and/or native *E. coli* chaperone DnaK. Reaction mixtures were incubated at either 30°C or 37°C. The \* represents undetectable bioluminescence activity. **(B)** SDS-PAGE analysis of the supernatants of the polyP-based PURE system reaction mixtures after incubation. The arrow indicates the position of Luc. Abbreviations: PPK2, family II polyP kinase CHU0107t; Luc, firefly luciferase (~62 kDa) (untrimmed gel image available in **ESI**).

Energy delivery is a critical aspect of cell-free protein synthesis,<sup>4</sup> and the one-enzyme energy-regenerating system described here simplifies (from a three-kinase system to an one kinase system) and cheapens the reaction (by using polyP) while yielding protein amounts comparable to that reported for current commercially-available systems. The use of long-chain polyP as substrate for reconstituted cell-free protein synthesis is also cost-effective since long-chain polyP can be readily obtained from insoluble phosphate glass (~0.01 USD per gram) using an established simple purification method.<sup>19</sup> Additionally, we have shown that the polyP-based regeneration system is compatible with the expression of heat-sensitive proteins (e.g. firefly luciferase) at 37°C, and that the yield of active firefly luciferase can be further enhanced by the addition of native *E. coli* chaperones DnaKJ and GrpE. Further optimization of the reaction components

and conditions may result in a simpler and more cost-effective commercial kit, leading to wider applications including large-scale production of polypeptides and high-throughput protein screening in functional genomics studies. Given that polyP is a ubiquitous, naturally occurring energy source as well as an active metabolic regulator for all cellular organisms,<sup>24,25</sup> incorporation of the polyP-based cell-free protein synthesis system in synthetic biomembrane vesicles could lead to artificial cell systems more akin to their natural counterparts.

P.-H.W. and K.F. conceptualised this research; A.K., A.F.Y. and P.-H.W. purified the recombinant CHU0107t and recombinant sfGFP; P.-H.W. characterized kinetic properties of CHU0107t; P.-H.W. and K.F. developed polyP-based NTP regeneration system for cell-free protein synthesis; P.-H.W. and T.Z.J. performed the electrophoretic analysis of polyP; P.-H.W. and S.B. performed the firefly luciferase activity assays; P.-H.W., K.F., Y.K., S.B. and S.E.M. participated in experimental design; all authors participated in the data analysis; P.-H.W. and S.E.M. wrote the paper with assistance from all authors.

This work is supported by JSPS KAKENHI Grant Numbers 10751084 to S.E.M. and JP18K14354 to T.Z.J. K.F. is supported by FirstLogic-Astrobiology donation program. A.K. and A.F.Y. are supported by the NSERC Industrial Biocatalysis Strategic Network grant IBN and Genome Canada (GAPP). Y.K. is supported by JST, PRESTO (Grant Number JPMJPR18K5). We also gratefully acknowledge support of customized PUREfrex reagents and codon-optimized linear DNA of *P. pyralis* firefly luciferase from Dr. Takashi Kanamori at GeneFrontier Corp. (Japan).

## Conflicts of interest

There are no conflicts to declare.

## References

- 1 A. D. Hanson, A. Pribat, J. C. Waller and Valérie de Crécy-Lagard, *Biochem. J.*, 2009, **425**, 1–11.
- 2 K. W. Ellens, N. Christian, C. Singh, V. P. Satagopam, P. May and C. L. Linster, *Nucleic Acids Res.*, 2017, **45**, 11495–11514.
- 3 T. Sawasaki, T. Ogasawara, R. Morishita and Y. Endo, *Proc. Natl. Acad. Sci.*, 2002, **99**, 14652–14657.
- 4 F. Katzen, G. Chang and W. Kudlicki, *Trends Biotechnol.*, 2005, **23**, 150–156.
- 5 W. M. and Nirenberg and J. H. Matthaei, *Proc. Natl. Acad. Sci.*, 1961, **47**, 1588–1602.
- 6 T. Niwa, B.-W. Ying, K. Saito, W. Jin, S. Takada, T. Ueda and H. Taguchi, *Proc. Natl. Acad. Sci.*, 2009, **106**, 4201–4206.
- 7 Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Ishikawa and T. Ueda, *Nat. Biotechnol.*, 2001, **19**, 751–755.
- 8 H. Matsubayashi, Y. Kuruma and T. Ueda, *Angew. Chemie - Int. Ed.*, 2014, **53**, 7535–7538.
- 9 Y. Kuruma, T. Suzuki and T. Ueda, in *Cell-free protein production*, eds. Y. Endo, K. Takai and T. Ueda, Humana Press, 2010, pp. 161–173.
- 10 S. Murakami, R. Matsumoto and T. Kanamori, *Sci. Rep.*, 2019, **9**, 1–13.
- 11 Y. Kuruma, T. Suzuki, S. Ono, M. Yoshida and T. Ueda, *Biochem. J.*, 2012, **442**, 631–638.
- 12 A. V. Zavialov and M. Ehrenberg, *Cell*, 2003, **114**, 113–122.
- 13 Kazuta, Y., Matsuura, T., Ichihashi, N. and Yomo, T, *J. Biosci. Bioeng.*, 2014, **118**, 554–557.
- 14 B. W. Ying, H. Taguchi, H. Ueda and T. Ueda, *Biochem. Biophys. Res. Commun.*, 2004, **320**, 1359–1364.
- 15 B. P. Nocek, A. N. Khusnutdinova, M. Ruszkowski, R. Flick, M. Burda, K. Batyrova, G. Brown, A. Mucha, A. Joachimiak, Ł. Berlicki and A. F. Yakunin, *ACS Catal.*, 2018, **8**, 10746–10760.
- 16 A. E. Parnell, S. Mordhorst, F. Kemper, M. Giurrandino, J. P. Prince and N. J. Schwarzer, *Proc. Natl. Acad. Sci.*, 2018, **115**, 3350–3355.
- 17 Ogawa, M., Uyeda, A., Harada, K., Sato, Y., Kato, Y., Watanabe, H., Honda, K. and T. Matsuura, *ChemBioChem*, 2019, doi:10.1002/cbic.201900303.
- 18 A. Kameda, T. Shiba, Y. Kawazoe, Y. Satoh and Y. Ihara, *J. Biosci. Bioeng.*, 2001, **91**, 557–563.
- 19 H. Itoh, Y. Kawazoe and T. Shiba, *J. Microbiol. Methods*, 2006, **64**, 241–249.
- 20 F. Caschera and V. Noireaux, *Metab. Eng.*, 2015, **27**, 29–37.
- 21 J. D. Pédelacq, S. Cabantous, T. Tran, T. C. Terwilliger and G. S. Waldo, *Nat. Biotechnol.*, 2006, **24**, 79–88.
- 22 K. Jackson, T. Kanamori, T. Ueda and Z. H. Fan, *Integr. Biol.*, 2014, **6**, 781–788.
- 23 B. Lavickova and S. J. Maerkl, *ACS Syn. Biol.* 2019, **8**, 455–462.
- 24 L. Achbergerová and J. Nahálka, *Microb. Cell Fact.*, 2011, **10**, 1–14.
- 25 M. J. Gray, W. Wholey, N. O. Wagner, C. M. Cremers, A. Mueller-schickert, N. T. Hock, A. G. Krieger, E. M. Smith, R. A. Bender, J. C. A. Bardwell and U. Jakob, *Mol. Cell*, 2014, **53**, 689–699.