Sciencexpress Reports

Operon structure and cotranslational subunit association direct protein assembly in bacteria

Yu-Wei Shieh,[1] Pablo Minguez,[2] Peer Bork,[2,3] Josef J. Auburger,[1] D. Lys Guilbride,[1,4] Günter Kramer,[1]* Bernd Bukau[1]*

[1]Center for Molecular Biology of the University of Heidelberg (ZMBH) and German Cancer Research Center (DKFZ), DKFZ-ZMBH Alliance, Im Neuenheimer Feld 282, Heidelberg D-69120, Germany. [2]European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany. [3]Max-Delbrück-Centre for Molecular Medicine, Robert-Rössle-Strasse 10, 13125 Berlin, Germany. [4]Malaria Research Foundation Inc, Post Office Box 10420, Aspen, CO 81612, USA.

*Corresponding author. E-mail: bukau@zmbh.uni-heidelberg.de (B.B.); g.kramer@zmbh.uni-heidelberg.de (G.K.)

Assembly of protein complexes is considered a post-translational process involving random collision of subunits. We show that within the *Escherichia coli* cytosol, bacterial luciferase subunits LuxA and LuxB assemble into complexes close to the site of subunit synthesis. Assembly efficiency decreases markedly if subunits are synthesized on separate mRNAs from genes integrated at distant chromosomal sites. Subunit assembly initiates cotranslationally on nascent LuxB in vivo. The ribosome-associated chaperone Trigger Factor delays the onset of cotranslational interactions until the LuxB dimer interface is fully exposed. Protein assembly thus is directly coupled to the translation process, and involves spatially confined, actively chaperoned cotranslational subunit interactions. Bacterial gene organization into operons therefore reflects a fundamental mechanism for spatiotemporal regulation vital to effective cotranslational protein complex assembly.

Oligomeric protein complexes are thought to assemble by diffusion and random collision of subunits within the cytosol (*1*) ("trans-assembly"; see fig. S1). This mechanism, however, does not explain how unassembled subunits avoid (i) non-specific interactions (ii) aggregation (iii) quality control sequestration to proteases and chaperones and (iv) navigate crowded and occluded cellular environments. We postulated that for complex subunits translated from polycistronic mRNAs, local confinement of assembly around the translation sites ("cis-assembly", fig. S1) would promote efficient, non-stochastic assembly.

We tested our hypothesis using the bacterial *Vibrio harveyi* heterodimeric luciferase complex which contains the subunits LuxA (α) and LuxB (β) (*2*), encoded by the *luxCDABE* operon (*3*). To express *luxA and luxB* from the same or distinct operons, we integrated into the *Escherichia coli* chromosome at distinct sites (fig. S2A) plasmids harboring artificial *lux* operons (fig. S2B). We fused genes encoding monomeric variant forms of enhanced YFP or CFP to the 5 end of *luxA* or *luxB* (fig. S2B). We created four different strains (1 through 4, Fig. 1A), each with two artificial operons carrying different tag configurations of the *yfp/cfp-luxA* and *yfp/cfp-luxB* fusion genes integrated into the chromosome.

To assess luciferase assembly in the vicinity of synthesis, we assayed specific luciferase activities and FRET efficiencies in strains 1 and 2 (Fig. 1A). Both strains encode *luxA* followed by *luxB* in each operon. The fluorescent FRET pairs in strain 1 tag LuxA and LuxB within each operon, whereas in strain 2, YFP tags both LuxA and LuxB subunits in one operon and CFP tags both subunits in the other. The strains exhibit similar specific luciferase activities (Fig. 1B, left, and fig. S3A), consistent with the identical genetic order of *luxA* and *luxB* within the operons. Strain 2 however, which encodes FRET pairs in separate operons displays only 40 ± 3% of the FRET efficiency of strain 1, which encodes FRET pairs within each operon (Fig. 1B, right). This demonstrates bacterial luciferase heterodimers are predominantly assembled in cis, from subunits synthesized from the same bicistronic mRNA molecule.

To establish whether operon organization promoting cis-assembly confers any advantage to the assembly process, we compared specific luciferase activities in cells where *luxA* and *luxB* are expressed from the same (strain 3) or separate (strain 4) operons (Fig. 1A). Expression levels for the fusion proteins are comparable between the two strains (fig. S3B). Expression from separate mRNAs (strain 4) however shows only 60 ± 3% of the specific luciferase activity measured in strain 3 (Fig. 1C). This difference was not caused by subunit aggregation (fig. S3C). Synthesis of complex subunits from one bicistronic mRNA encoded by an operon therefore strongly increases assembly efficiency compared to trans-assembly.

We postulated that complex assembly might occur already cotranslationally, to effectively preempt pre-assembly diffusion. We used Selective Ribosome Profiling (SeRP) to uncover any nascent subunit interactions during luciferase complex assembly in vivo (*4*). This method (*4*, *5*) compares the distribution profile of nuclease-protected mRNA fragments (ribosome footprints) isolated from all translating ribosomes to the profile of ribosomes selected by immuno-

purification of YFP-tagged luciferase subunits (fig. S4A, IP ribosomes). We constructed two *E. coli* strains (5 and 6), each containing a single luciferase operon where either *luxA* (strain 6) or *luxB* (strain 5), but not both, is *yfp* 5′-end tagged (fig. S4B). Genetic fusion of YFP to LuxA or LuxB N-termini allows enrichment of potentially two classes of translating ribosomes: those directly translating the YFP-Lux protein (fig. S4A, IP part) versus those translating the untagged partner subunit, when and if cotranslational engagement of the YFP-Lux protein occurs (fig. S4A, co-IP part). The N-terminal YFP tagging allows immunopurification of the first class of ribosomes engaged with this subunit. Independent selection of *yfp-luxB* (strain 5) and *yfp-luxA* (strain 6) gene products generates a profile for each showing strong ribosome footprint enrichment once the YFP part (248 residues) of the nascent YFP-LuxB and YFP-LuxA chains has emerged (fig. S5, upper panel), confirming immunopurification is highly selective.

We then compare the density of ribosome footprints across *lux* genes of both datasets, the immunopurified ribosomes and the total pool of all ribosomes. We find strong evidence for the second class of translating ribosomes: immunopurification of YFP-LuxA copurifies ribosomes synthesizing LuxB. This directly demonstrates that nascent LuxB cotranslationally interacts with YFP-LuxA (Fig. 2A, right, without TF). This interaction is unlikely to occur after cell lysis since the crowded cytoplasm becomes highly diluted which greatly reduces the possibility of nascent subunit interactions (supplementary online text). As translation of *luxB* starts, footprint enrichment (reflecting interaction with YFP-LuxA) fluctuates around background levels, but rises with increasing length of nascent LuxB. We defined a stable two-fold enrichment threshold to reliably indicate initiation of cotranslational subunit interactions. According to this threshold, interaction of YFP-LuxA with nascent LuxB starts once about 120 residues of LuxB are synthesized, and becomes more robust with increasing length until a 20-fold enrichment, reflecting strongly interacting subunits, is reached near the termination of *luxB* translation. Assuming the ribosomal exit tunnel protects 30 C-terminal residues, emergence of the 90 amino-terminal residues of LuxB is evidently sufficient to initiate the assembly interactions. We also detected interactions of YFP-LuxB with nascent LuxA, at much lower enrichment (maximally 6-fold), reflecting less robust interactions (Fig. 2A, left, without TF). We conclude both nascent luciferase subunits can initiate complex assembly, but that cotranslational assembly initiates predominantly on nascent LuxB.

These experiments used a Δ*tig* mutant background lacking ribosome-associated Trigger Factor (TF) chaperone. TF interacts transiently with most nascent cytosolic proteins, after an average minimal nascent chain length of approximately 110 residues is synthesized (*4*). To test the impact of TF we expressed wild-type levels of TF in strains 3 and 4 (making strains 7 and 8 respectively) and strains 5 and 6 (making strains 9 and 10). In vivo luciferase activity measurements of strains 7 and 8 indicate no impact of TF on the enhanced efficiency of cis-assembly (fig. S6). However, SeRP experiments of strains 9 (*yfp-luxB* in bicistronic organization with unlabeled *luxA*) and 10 (*yfp-luxA* in bicistronic organization with unlabeled *luxB*) reveal that TF does affect cotranslational interactions of the nascent luciferase subunits (Fig. 2A). Comparison of the cotranslational interaction of YFP-LuxB with nascent LuxA in the absence (strain 5) and presence of TF (strain 9) shows that TF efficiently suppresses interactions of YFP-LuxB with nascent LuxA (Fig. 2A, left). In contrast, TF delays, but does not block the interaction of YFP-LuxA with nascent LuxB, shifting the minimal length of nascent LuxB that is required for cotranslational interaction from 90 to 152 residues (Fig. 2A, right), without affecting the approximately 20-fold enrichment of ribosome footprints near the end of *luxB* translation. We infer TF chaperone activity shields nascent LuxB from premature interactions with LuxA but allows the exposure of the dimerization fold for timely interactions with LuxA molecules in the ribosome vicinity.

The luciferase heterodimer interface is large (residues 17-161 in LuxA and 11-163 in LuxB). Extensive contacts occur via a parallel four-helix bundle formed by helices α2 and α3 from either subunit (residues 55-65 and 83-97 in LuxB, Fig. 2B and table S1) (*2*). Given the length of the ribosomal tunnel, α2 and α3 of nascent LuxB will be fully exposed after translation of the first 127 codons; the complete subunit interface of LuxB will be exposed after translation of codon 193. In the absence of TF, initial contacts of nascent LuxB with LuxA are established upon emergence of helix α3 (translation of the first 120 codons). With TF present we detected initial interactions of nascent LuxB with LuxA only upon exposure of around 152 N-terminal LuxB residues (translation of the first 182 codons) comprising nearly the complete subunit interface (Fig. 2A, right).

We conclude that association of fully-synthesized LuxA with nascent LuxB requires ribosome-associated exposure of critical structural elements of the dimerization interface on LuxB. TF further increases the specificity of subunit interactions by preventing association of LuxB with nascent LuxA and premature association of LuxA with short nascent chains of LuxB, but selectively allows LuxA association with long nascent chains of LuxB exposing the complete dimer interface. We propose that TF engagement with nascent LuxA prevents premature interactions of the latter with LuxB and effectively promotes timely delivery of full length LuxA to nascent LuxB. Likewise, interaction of TF with nascent LuxB may protect the nascent subunit during early stages of translation from incorrect and premature folding and interactions, until the assembly-competent part of LuxB is sufficiently exposed at the ribosomal surface for productive interaction with LuxA. Our results provide proof-of-

principle for organized and regulated protein assembly in bacteria and rationalize previous observations that suggested translational coupling affects protein complex assembly (*6*). The SeRP data provided for bacterial luciferase are consistent with a concept that a fully synthesized subunit encoded by an upstream gene interacts with the nascent subunit encoded by the downstream gene of an operon. Assembly, considered a final, post-translational step in the generation of oligomeric proteins, emerges as a process that is physically and kinetically coupled to the fundamental processes of protein biogenesis, folding and translation. Our findings add a further dimension to the concept of the operon, originally defined by Jacob and Monod as a genetic unit for coordinated regulation of transcription (*7*). The genetic organization into operons of genes for products destined for assembly into protein complexes, determines local concentration of subunits and crucial timing of assembly.

REFERENCES AND NOTES

1. Y. Phillip, G. Schreiber, Formation of protein complexes in crowded environments—from in vitro to in vivo. *FEBS Lett.* **587**, 1046–1052 (2013). Medline doi:10.1016/j.febslet.2013.01.007

2. T. O. Baldwin, J. A. Christopher, F. M. Raushel, J. F. Sinclair, M. M. Ziegler, A. J. Fisher, I. Rayment, Structure of bacterial luciferase. *Curr. Opin. Struct. Biol.* **5**, 798–809 (1995). Medline doi:10.1016/0959-440X(95)80014-X

3. C. M. Miyamoto, A. D. Graham, M. Boylan, J. F. Evans, K. W. Hasel, E. A. Meighen, A. F. Graham, Polycistronic mRNAs code for polypeptides of the Vibrio harveyi luminescence system. *J. Bacteriol.* **161**, 995–1001 (1985). Medline

4. E. Oh, A. H. Becker, A. Sandikci, D. Huber, R. Chaba, F. Gloge, R. J. Nichols, A. Typas, C. A. Gross, G. Kramer, J. S. Weissman, B. Bukau, Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell* **147**, 1295–1308 (2011). Medline

5. A. H. Becker, E. Oh, J. S. Weissman, G. Kramer, B. Bukau, Selective ribosome profiling as a tool for studying the interaction of chaperones and targeting factors with nascent polypeptide chains and ribosomes. *Nat. Protoc.* **8**, 2212–2239 (2013). Medline doi:10.1038/nprot.2013.133

6. P. Pradhan, W. Li, P. Kaur, Translational coupling controls expression and function of the DrrAB drug efflux pump. *J. Mol. Biol.* **385**, 831–842 (2009). Medline doi:10.1016/j.jmb.2008.11.027

7. F. Jacob, J. Monod, Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356 (1961). Medline doi:10.1016/S0022-2836(61)80072-7

8. A. Haldimann, B. L. Wanner, Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. *J. Bacteriol.* **183**, 6384–6393 (2001). Medline doi:10.1128/JB.183.21.6384-6393.2001

9. M. J. Casadaban, Transposition and fusion of the *lac* genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. *J. Mol. Biol.* **104**, 541–555 (1976). Medline doi:10.1016/0022-2836(76)90119-4

10. D. A. Zacharias, J. D. Violin, A. C. Newton, R. Y. Tsien, Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* **296**, 913–916 (2002). Medline doi:10.1126/science.1068539

11. D. Kentner, S. Thiem, M. Hildenbeutel, V. Sourjik, Determinants of chemoreceptor cluster formation in Escherichia coli. *Mol. Microbiol.* **61**, 407–417 (2006). Medline doi:10.1111/j.1365-2958.2006.05250.x

12. W. S. Hayes, M. Borodovsky, Deriving ribosomal binding site (RBS) statistical models from unannotated DNA sequences and the use of the RBS model for N-terminal prediction. *Pac. Symp. Biocomput.* **1998**, 279–290 (1998). Medline

13. D. A. Parsell, A. S. Kowal, M. A. Singer, S. Lindquist, Protein disaggregation mediated by heat-shock protein Hsp104. *Nature* **372**, 475–478 (1994). Medline doi:10.1038/372475a0

14. J. W. Hastings, R. P. Presswood, Bacterial luciferase: FMNH2-aldehyde oxidase. *Methods Enzymol.* **53**, 558–570 (1978). Medline doi:10.1016/S0076-6879(78)53057-7

15. D. Kentner, V. Sourjik, Dynamic map of protein interactions in the Escherichia coli chemotaxis pathway. *Mol. Syst. Biol.* **5**, 238 (2009). Medline doi:10.1038/msb.2008.77

16. V. Sourjik, A. Vaknin, T. S. Shimizu, H. C. Berg, In vivo measurement by FRET of pathway activity in bacterial chemotaxis. *Methods Enzymol.* **423**, 365–391 (2007). Medline doi:10.1016/S0076-6879(07)23017-4

17. F. C. Neidhardt, P. L. Bloch, D. F. Smith, Culture medium for enterobacteria. *J. Bacteriol.* **119**, 736–747 (1974). Medline

18. N. T. Ingolia, Genome-wide translational profiling by ribosome footprinting. *Methods Enzymol.* **470**, 119–142 (2010). Medline doi:10.1016/S0076-6879(10)70006-9

19. U. Rothbauer, K. Zolghadr, S. Muyldermans, A. Schepers, M. C. Cardoso, H. Leonhardt, A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. *Mol. Cell. Proteomics* **7**, 282–289 (2008). Medline doi:10.1074/mcp.M700342-MCP200

20. A. C. Clark, S. W. Raso, J. F. Sinclair, M. M. Ziegler, A. F. Chaffotte, T. O. Baldwin, Kinetic mechanism of luciferase subunit folding and assembly. *Biochemistry* **36**, 1891–1899 (1997). Medline doi:10.1021/bi962477m

ACKNOWLEDGMENTS

We thank A. Becker, V. van Noort, and C. Andréasson for valuable contributions; V. Sourjik and co-workers for help with in vivo FRET analysis; M. Langlotz and the FACS facility of the ZMBH for support; D. Ibberson for optimizing DNA sample preparation protocols, Y. Vainshtein for support with data analysis; B. Zachmann-Brand for technical support. Sequencing was done at Genomics and Proteomics Core (DKFZ) and Genomics Core (EMBL) facilities. This work was supported by research grants from the German Science Foundation (DFG; SFB638 and FOR1805) to G.K. and B.B.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/cgi/content/full/science.aac8171/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S8
Tables S1 to S3
References (*37*–*50*)

17 June 2015; accepted 15 September 2015
Published online 24 September 2015
10.1126/science.aac8171

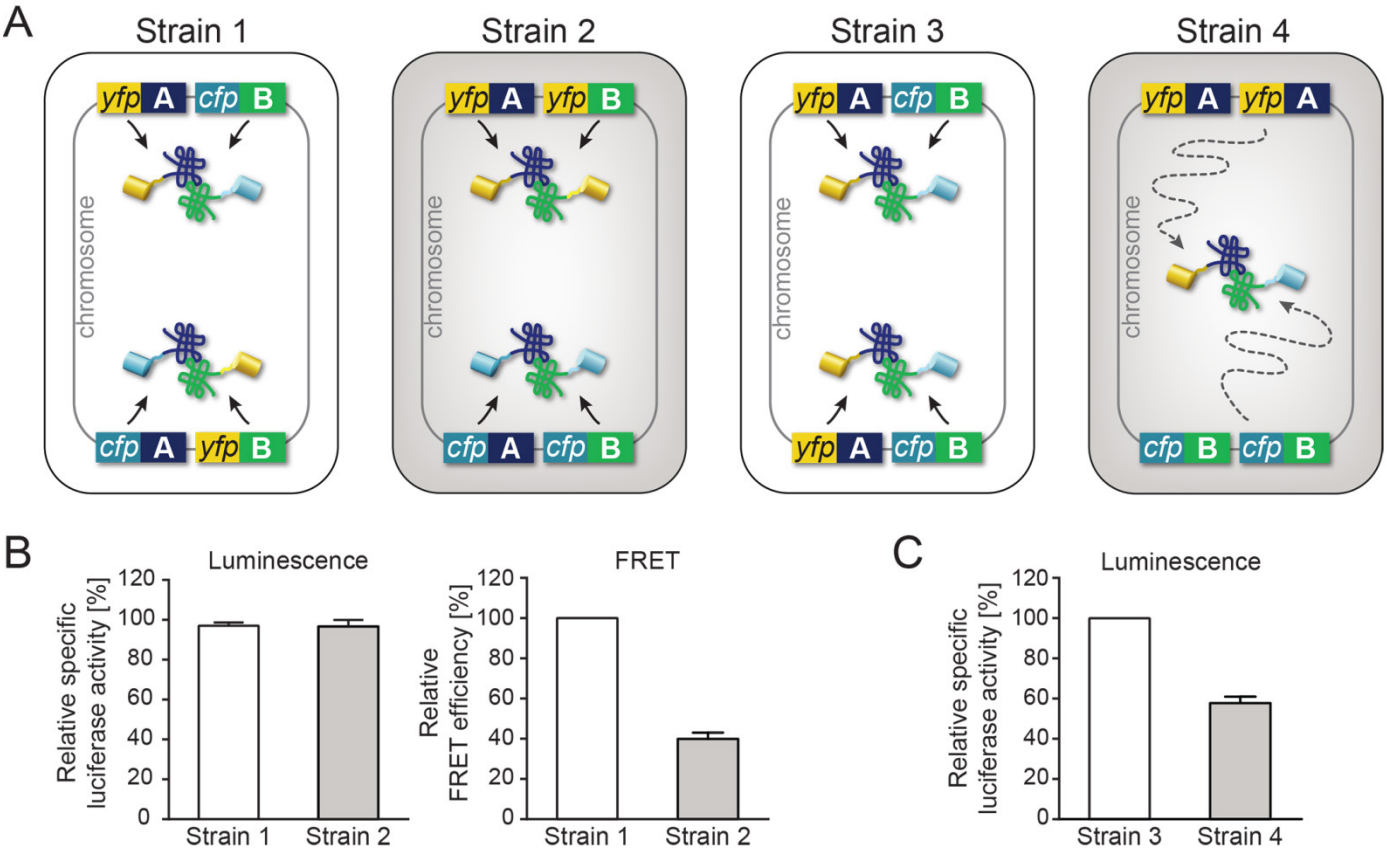

Fig. 1. Assembly of bacterial luciferase. (**A**) *E. coli* strains 1 to 4 with two bicistronic operons integrated into the chromosome. Operons encode fluorescently-tagged *luxA* or *luxB* genes under control of IPTG-inducible promoters and strong Shine-Dalgarno sequences. (**B**) Relative specific luciferase activities (mean ± SEM, $n = 3$, unpaired t test, $P = 0.9207$, fig. S3A) and relative FRET efficiencies (mean ± SEM, $n = 3$, unpaired t test, $P < 0.0001$) of strains 1 and 2 determined in vivo. (**C**) Relative specific luciferase activities of strains 3 and 4 determined in vivo (mean ± SEM, $n = 19$, unpaired t test, $P < 0.0001$, fig. S3B, right). Significance tests used software GraphPad Prism 6.

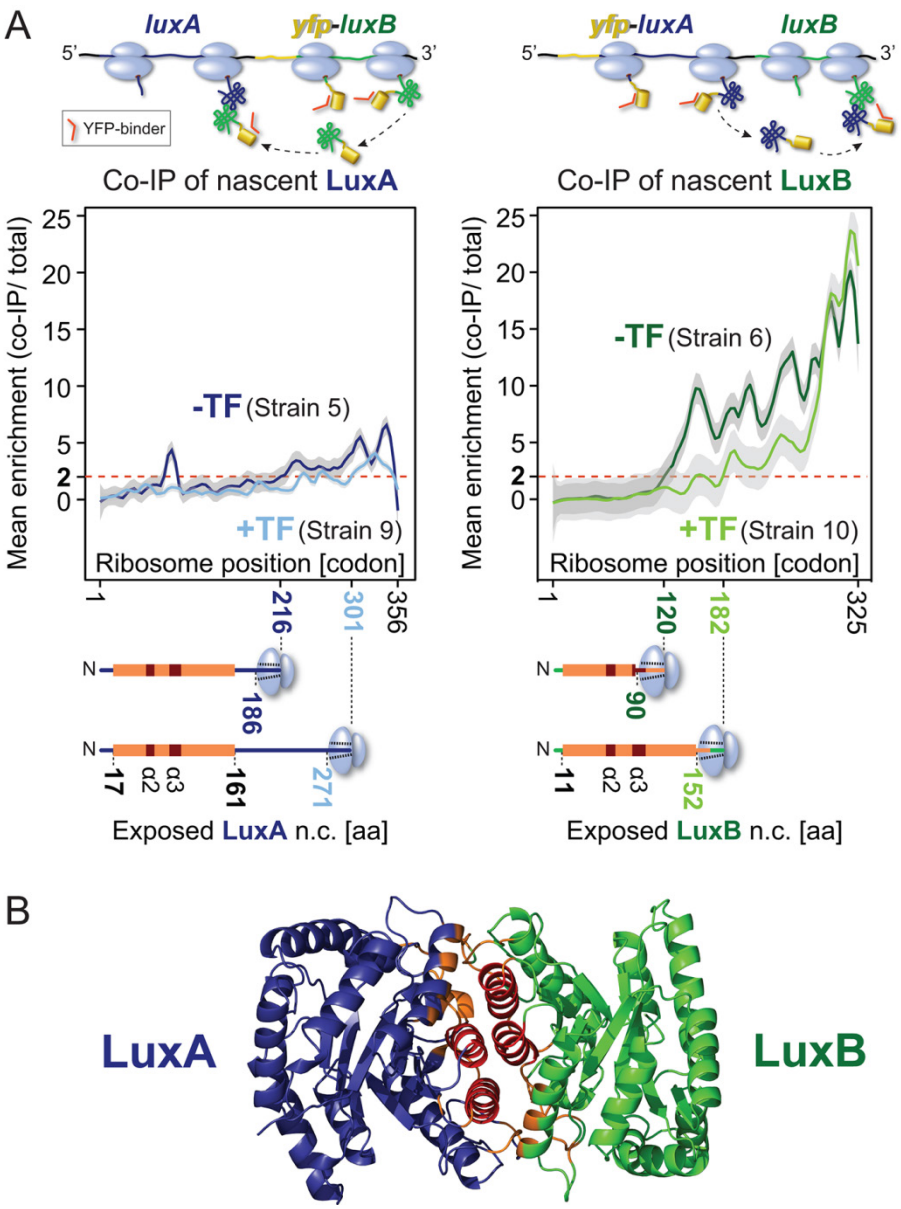

Fig. 2. Assembly of bacterial luciferase occurs cotranslationally. (**A**) Engagement of nascent LuxA by YFP-LuxB (left) and nascent LuxB by YFP-LuxA (right) without TF (strains 5 and 6, -TF) or with TF (strains 9 and 10, +TF). Upper cartoons illustrate *luxA/B* operons and the principle of SeRP analysis (fig. S4A). Mean enrichments were corrected for non-specific background binding analyzed with strains 11 and 12 (fig. S7A and table S2). Gray zone, loess curves (span parameter = 0.1) with 95% confidence interval. Bold-colored numbers, ribosome positions when the mean enrichments stably cross the two-fold threshold. Cartoons below show exposed nascent chain lengths. Orange, dimer interface, red, helices α2 and α3. (**B**) Crystal structure of the heterodimer of LuxA (blue) and LuxB (green) (PDB accession code: 1LUC). Orange, dimer interface, red, helices α2 and α3. LuxA and LuxB subunit residues involved in subunit interaction, table S1.