scientific reports

Check for updates

OPEN

A galactose-based auto-expression system improves T7-inducible protein production in *Escherichia coli*

James Bosco[1]✉, Emily Gagliano[1], Kassandra L. Boshae[1], John P. Statz[1], Timothy B. Wheeler[1], DeAnna Cuello[1], Ashlyn Sliter[1], Christian Newby[1], Bernice Lin[1], Aysha Demeler[3], C. Logan Pierpont[1], Cindee Yates-Hansen[1], Matthew J. Sydor[1], Maria E. Ferrini[1], Kellie C. Kuch[1], Brandon S. Cooper[1], Beverly J. Piggott[1], Sarah J. Certel[1], Kasper B. Hansen[1], Stephen R. Sprang[1], Bruce Bowler[1], Levi McClelland[1], Mehmet Berkmen[2] & Ekaterina Voronina[1]✉

Protein production using *Escherichia coli* is a cornerstone of modern biotechnology. In this study, we developed a novel auto-expression medium to maximize protein production. Each *E. coli* strain tested was capable of auto-expression in response to galactose, including strains in which the endogenous *lacZ* had been disrupted. This provides key evidence that galactose can regulate the *lac* operon independent of known *lac* operon-regulated metabolism. The enhanced capabilities of the novel auto-expression medium were documented across protein production systems including (1) increased yields for routinely expressed proteins (e.g. eGFP), (2) improved expression of human cytochrome *c* within a dual expression system, (3) robust auto-expression in *lacZ*-deficient strains producing proteins with challenging disulfide bonds, and (4) reproducible 8-fold increase in SpCas9 yields, at ≥ 95% purity. This novel medium can streamline production and improve yields for routine as well as challenging proteins, accelerating recombinant protein production and creating new opportunities in biotechnology and structural biology.

The global protein production market, currently valued at approximately 400 billion USD annually, attributes nearly a third of this revenue to protein production in *Escherichia coli*[1–4]. The T7 expression system, developed by William Studier in 1986, represents a cornerstone of *E. coli*-based protein production processes due to its simplicity and cost-effectiveness[2–4]. Despite extensive optimization of the T7 expression system over the years, challenges in scalability and reproducible expression under standard conditions persist, particularly for certain proteins of interest (POI)[4,5].

Central to the T7 system is the coupling of the cellular machinery that governs the *lac* operon in *E. coli* to the expression of the T7 RNA polymerase. This system is regulated by two main mechanisms: carbon catabolite repression of the transcriptional master regulator cyclic adenosine monophosphate (cAMP) receptor protein (CRP)[6,7] and transcriptional repression by the *lac* repressor, LacI[8–10]. CRP bound to cAMP enhances transcription, while glucose reduces the CRP-cAMP levels resulting in transcriptional downregulation[6]. Concurrently, LacI binds to the operator site of the *lac* operon, blocking transcription by RNA polymerase[10]. Thus, high levels of *lac* operon expression only occur when two conditions are met: glucose levels are low, allowing transcriptional activation by CRP-cAMP; and lactose/allolactose levels are high, binding to LacI and inducing conformational changes that disrupt LacI binding to the *lac* operator and permit transcription[6,7]. In BL21(DE3) *E. coli*, the integrated *lac* operator site upstream of the T7 polymerase genetic locus mediates activation of T7 RNA polymerase gene in response to the appropriate inducers[8–10]. T7 polymerase, now expressed, recognizes the T7 promoter on the plasmid carrying the POI open reading frame[8–10]. Notably, the T7 promoter is not recognized

[1]Division of Biological Sciences, Department of Chemistry and Biochemistry, Center for Biomolecular Structure and Dynamics, Center for Structural and Functional Neuroscience, Center for Environmental Health Sciences, University of Montana, Missoula, MT 59812, USA. [2]New England Biolabs, 240 County Road, Ipswich, MA 01938, USA. [3]Department of Chemistry and Biochemistry, Alberta RNA Research and Training Institute, University of Lethbridge, Lethbridge, AB T1K 3M4, Canada. ✉email: james.bosco@umontana.edu; ekaterina.voronina@umontana.edu

by the endogenous bacterial RNA polymerase, ensuring exclusive expression of the target POI in response to the inducers.

The two most widely employed inducers for T7 expression system, IPTG (Isopropyl ß-D-1-thiogalactopyranoside) and lactose utilized in auto-expression protocols, activate T7 RNA polymerase expression through *lac* operon regulatory elements. Some alternative T7 induction systems include those induced by arabinose, rhamnose, and tetracycline[5], or leaky expression induced by the contaminant lactose found in tryptone[4]. Induction by IPTG, a galactose derivative that mimics allolactose and binds to LacI, directly lowers the ability of the *lac* repressor to bind to the *lac* operon, promoting robust protein expression. This method, while effective, can introduce cellular stress due to overwhelming levels of POI expression leading to cytotoxicity[11,12]. Auto-expression alleviates cytotoxicity by using native regulatory mechanisms exploiting the sequential utilization of glucose and alternative carbon sources like lactose or galactose. The production of POIs occurs as the cell naturally transitions from glucose to secondary carbon sources like lactose, which induces the *lac* operon[1,2,4]. Auto-expression not only aligns with the cell's metabolic state but also reduces stress, as it avoids the abrupt changes induced by high levels of protein expression when IPTG is added acutely[2].

Given that auto-expression proves generally more effective and less stressful for the cells, particularly when expressing cytotoxic proteins[2,4], we focused on maximizing protein production using auto-expression. To address bottlenecks within the protein production workflow brought on by low yields of challenging POIs, we developed an auto-expression medium denoted as Bosco Broth (BB). BB facilitates higher expression of POIs compared with conventional autoinduction protocols[2,4]. We show that BB improves production for a broad range of expression systems and *E. coli* strains, expanding upon the current T7 system by utilizing galactose rather than conventional lactose as an inducer for auto-expression. We show that BB-powered auto-expression is suitable for diverse protein production processes relevant for academic research and the biotechnology industry.

Results

BB is the superior media for eGFP production

We hypothesized that auto-expression using BB can be measured and monitored using eGFP fluorescence. Using fluorescence as a proxy for the expression of a GFP fusion protein, especially when utilizing auto-expression, can be a robust readout within *E. coli* cultures. Capturing GFP fluorescence throughout the expression process provides key insights regarding POI status in the following ways: (1) demonstration that the POI is being expressed, given that GFP is fluorescent and is capable of being visualized, (2) at minimum, the POI is present in the soluble fraction of the *E. coli* since GFP is not fluorescent if localized to inclusion bodies[13,14] and (3) fluorescence readings as a function of protein expression can be measured in minutes, whereas expression monitoring by SDS-PAGE gel analysis following small scale sample lysis is time consuming and laborious[15,16]. However, it should be noted that a GFP protein could be proteolytically cleaved from a misfolded fusion protein, thus an analysis checking that the molecular weight corresponds to the full-length fusion protein is essential. If premature cleavage is realized, the released GFP will fluoresce resulting in false positive expression data. Thus, final yields of POI must be visualized using conventional methods.

To evaluate and compare the capacity of BB media to produce proteins, BB was tested in comparison to ThermoFisher's MagicMedia™ (MM), Terrific Broth (TB), Lysogeny Broth (LB), and Studier broth (ZYM). Expression of the target protein in BB media is induced by 0.055 M galactose. The other media contain alternative inducers such as 0.2% lactose (ZYM), 1 mM IPTG (LB and TB), or proprietary contents (Magic Media). In addition, glucose is added to repress POI expression during early bacterial growth in BB and ZYM. BB's full recipe can be found in Materials and Methods. *E. coli* BL21(DE3) cells expressing 6xHis-eGFP from a pET21a plasmid were monitored by OD_{600} for biomass and by fluorescence emission (516 nm) for eGFP production in an orbital shaker culture[16]. BB supported a similar bacterial optical density (OD_{600}) compared to TB, MM, and ZYM (Fig. 1a). We monitored eGFP fluorescence intensity within BL21(DE3) *E. coli* without lysing the cells to measure eGFP accumulation over time. Intriguingly, despite similar OD_{600} values, BB achieved significantly higher eGFP fluorescence (Fig. 1b,d). To validate eGFP fluorescence signal as a robust metric of eGFP protein production, one replicate from each media recipe was lysed after 36 h, soluble eGFP was purified from each culture in parallel using affinity chromatography, and the total yield of eGFP in each sample was determined with correction for protein purity. Total yields of eGFP without normalization to OD of cultures display the same approximate 3x difference observed during the in-cell fluorescence measurements (Fig. 1b,c). We concluded that in-cell fluorescence measurements accurately report key protein production values without significant fluorescent artifacts. Further normalization of eGFP fluorescence intensity relative to OD_{600} and total culture volume revealed that the enhanced eGFP expression capability in BB was a function of higher cell-specific yield per OD unit (Fig. 1d). Notably, LB was able to produce significantly more eGFP per OD_{600} compared to TB, MM, or ZYM. However, total protein yield from BB outcompetes LB as BB supports a higher density of *E. coli* (Fig. 1b,c).

We hypothesized that the increased yield of eGFP in BB or LB compared to ZYM, TB, and MM resulted from improved solubility of the recombinantly expressed protein[13]. To test this hypothesis, we compared eGFP distribution between the soluble and insoluble fractions in each culture condition. We observed distinct distribution of eGFP between the soluble and insoluble fractions in each media both by a total protein gel staining, as well as by a western blot against the 6xHis-eGFP (Fig. 1e). These qualitive observations supported our hypothesis predicting that BB would alleviate protein solubility shortcomings of other media. We conclude that BB can produce more eGFP compared to other popular media, in-cell fluorescence can be used to monitor and maximize protein production in auto-expression processes, and that higher folding efficiency in BB contributes to higher target protein yields.

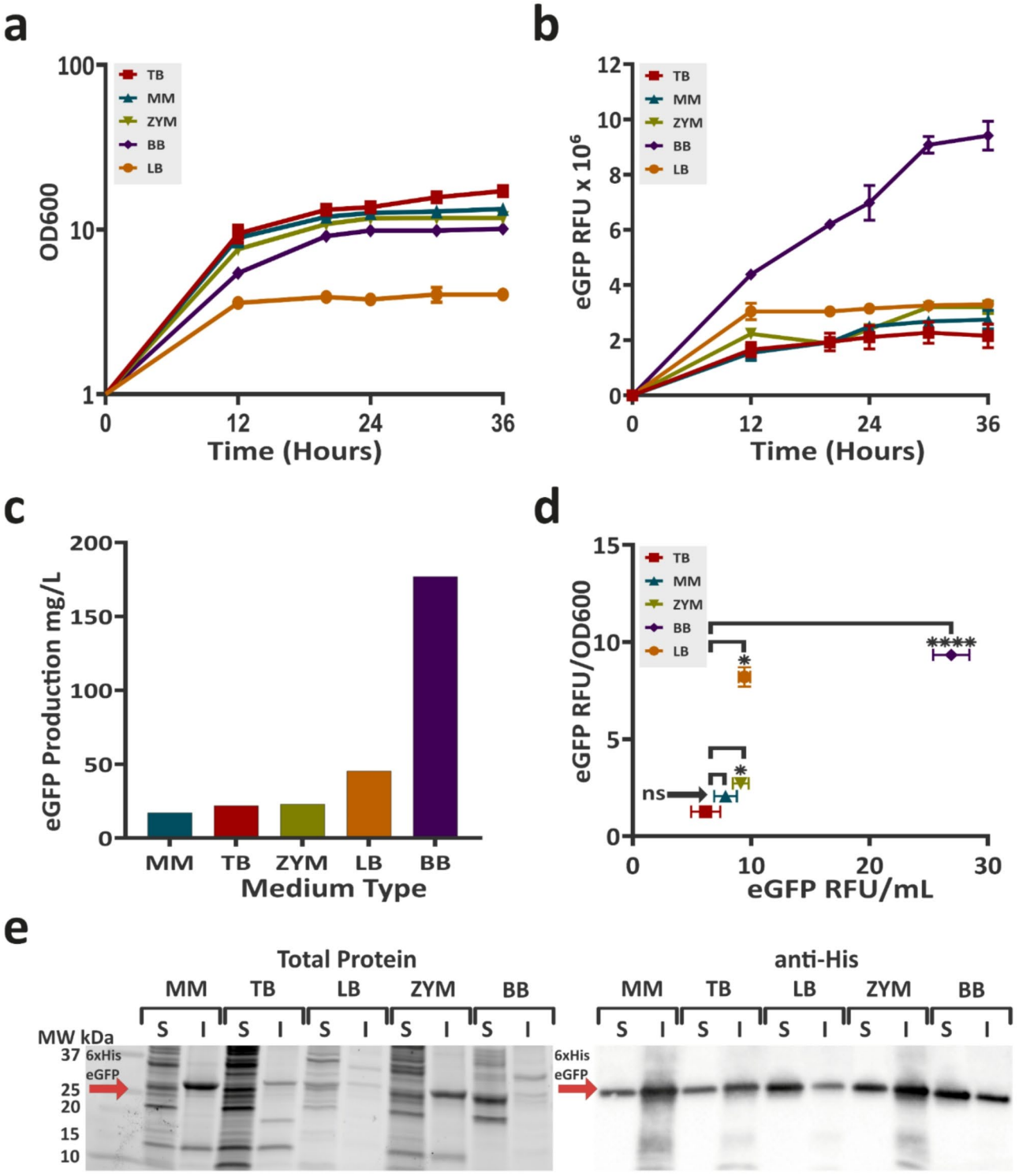

Fig. 1. Improved eGFP production in BB. (**a**) OD600 (log scale) of BL21(DE3) *E. coli* cultures expressing eGFP in different media through time. (**b**) eGFP fluorescence (ex. 485 nm, em. 516 nm) of bacterial cultures with respect to time for varying media. Samples are the same as in panel (**a**). (**c**) Yields of purified eGFP from varying media at 36 h. (**d**) eGFP fluorescence intensity normalized per cell density (OD600) or culture volume. The values in panels (**a**, **b**, and **d**) represent averages of 3 replicate cultures, error bars are SD. eGFP fluorescence per culture volume values in each media were compared by one-way Anova with Brown-Forsythe and Welch post-test to the values obtained in TB. Asterisks denote statistical significance, **** $P < 0.00001$, * $P < 0.01$. These experiments were performed once. (**e**) SDS-PAGE of soluble (S) and insoluble (I) fractions of eGFP-expressing cultures in different media, as indicated above each gel. Left: total protein stain-free detection. Right: detection of eGFP by a Western blot with anti-His-tag antibodies. Uncropped gel images are shown in Fig. S1.

BB increases efficiency of a dual expression system

We next investigated BB's performance in a challenging co-expression scenario, where the target protein needs to be produced along with an essential accessory factor. We chose human cytochrome *c* (Hu-Cyt*c*), which requires simultaneous production of yeast heme lyase encoded by the same plasmid as the Hu-Cyt*c*[17], as a co-expression model. Heme lyase is required to covalently attach a heme group onto apo-Cyt*c*, ultimately producing a functional holo Cyt*c*[17–19] (Fig. 2a). Heme lyase and Hu-Cytc are expressed from a bicistronic open reading frame[17]. Historically, either TB or 2xYT have been used to express Hu-Cyt*c*[17,19] where 2xYT is a nutrient dense medium with similar ingredients as TB only different by the ratio of tryptone and yeast extract, the presence of NaCl, and the absence of glycerol[20]. However, in our hands, 30–40% of the cultures grown in either TB or 2xYT fail to express Hu-Cyt*c* despite being grown simultaneously in the same orbital shaker with other successful cultures. By contrast, when expressed in BB, 12 of 12 cultures displayed a vibrant pink color indicative of proper heme group assembly required for Hu-Cyt*c* function, demonstrating consistent expression in every BB culture (Fig. 2b). There was not an appreciable difference between the pellets with respect to color among different replicates cultured in parallel (data not shown). Hu-Cyt*c* protein purification from three of the representative replicate 1 L cultures confirmed production of the same molecular weight and purity protein as the standard

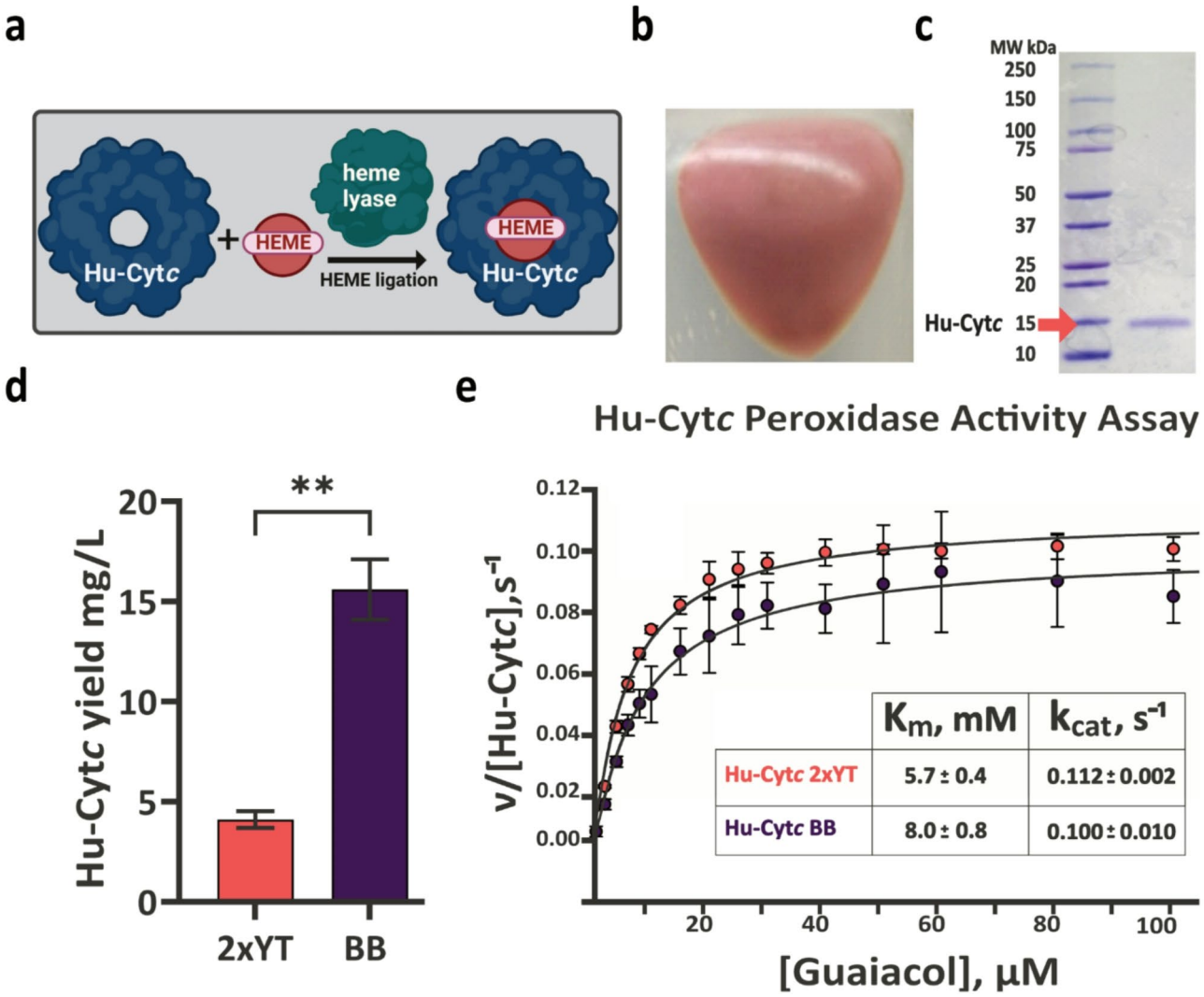

Fig. 2. Co-expression of Hu-Cyt*c* and yeast heme lyase in BB vs. 2xYT. (**a**) Schematic representation of the dual expression system that produces a functional Hu-Cyt*c* enzyme. (**b**) Pelleted BL21(DE3) *E. coli* expressing Hu-Cyt*c*. (**c**) Total protein gel of a representative Hu-Cyt*c* purification. Whole gel image is shown in Fig. S2a. (**d**) Total protein yield comparison between successful 2xYT replicates and BB. Plotted values represent averages of 3 replicates per media, error bars are SD. Protein yields were compared by t-test, asterisks denote statistical significance ** $P < 0.001$. (**e**) Michaelis-Menten plots for Hu-Cyt*c* produced in 2xYT[21] and BB, in red and purple, respectively. Plotted are average values from three technical replicates, error bars are SD. The solid curves fit to the Michaelis-Menten equation, and Michaelis-Menten parameters for the peroxidase activity of Hu-Cyt*c* holoenzyme produced in 2xYT[21] and BB.

induction (Fig. 2c, S2). When compared to successfully grown 2xYT cultures following the standard expression protocol, BB produced approximately three times as much Hu-Cyt*c* (15 mg/L *versus* 5 mg/L) (Fig. 2d).

To confirm that proteins expressed in BB were active, functional enzymes, we measured peroxidase activity of the purified Hu-Cyt*c* protein by monitoring the production of tetraguaiacol from guaiacol in the presence of H_2O_2[19,21–27] and compared the kinetic data to previously reported values for protein expressed in 2xYT[19]. We generated Michaelis-Menten plots with respect to guaiacol concentration to determine the k_{cat} and K_m parameters (Fig. 2e). Overall, the plots were similar, and the k_{cat} values obtained from the fits were the same within the margin of error (two-tailed t-test $P > 0.05$) (Fig. 2e). The K_m values were not strongly affected by the change in production media (Fig. 2e). Following confirmation of enzymatic activity, we conclude that in this dual expression system, BB reliably produced active, functional enzyme yielding enzymatic parameters consistent with previously reported data[19].

BB maintains a higher number of cells producing target proteins in a dual expression system

BB might outperform other media in dual-expression systems by several mechanisms, including maintaining a higher number of protein-producing cells in the culture, achieving higher protein output per cell, or allowing efficient simultaneous expression of two proteins rather than skewing production to a single dominant protein. To understand which potential scenario best describes BB's enhanced Hu-Cyt*c* expression, we used flow cytometry and a fluorescent dual expression system consisting of mStayGold green fluorescent protein and mTagBFP2 blue fluorescent protein. These two fluorophores were expressed from a bi-cistronic open reading frame construct (Fig. 3a)[28,29], which can allow for a similar level of biosynthesis, with a subtle production bias favoring the ORF that is transcribed first within the system. We allowed for possible lower cell positivity counts as well as POI expression levels by using the high quantum yields intrinsic to mStayGold and mTagBFP2. This provided our system with a broad dynamic range for fluorescence detection[28,29]. Flow cytometry provided the capability to determine the fraction of bacteria expressing fluorescent proteins, the level of protein produced per cell per medium, and the relative amounts of mStayGold and mTagBFP2 found in dual fluorescent cells. We chose LB and TB as the two media comparisons because TB is similar in its base composition to BB, and LB historically supports high fluorescent protein expression. As a result, LB has been used to produce an array of fluorescent proteins (Fig. 1d)[28].

We followed the published growth conditions for mStayGold[28] and analyzed cells 5 days post inoculation (BB) or IPTG induction (LB and TB) by flow cytometry. All samples were gated for intact cells (Fig. S4), and the percentage of cells producing blue-green, dual-fluorescent signal was established using uniform gating for all cultures based on a non-fluorescent control culture (Fig. S4). BB produced 10-fold more dual fluorescent cells compared with LB and TB (Fig. 3b,c). Additionally, the mean fluorescence intensity for each fluorophore in the dual-positive cells was similar among the three media, with a minor increase in BB (Fig. 3d). In every media tested, TagBFP2 fluorescence intensity trended higher than that of mStayGold, although the difference was statistically significant only in TB, which was confirmed by microscopy (Fig. 3e, S3). Although bicistronic constructs are expected to produce similar levels of coexpressed proteins, differences in relative fluorescence of two FPs might stem from distinct translational efficiency, folding, or protein stability. We conclude that BB supports a larger population of cells that simultaneously produce two target proteins in dual-expression systems compared to LB and TB. The BB-cultured cells displaying both blue and green fluorescence also showed higher mean fluorescence intensities, thus achieving similar or better protein output per cell (Fig. 3e). Additionally, dual expression in TB and perhaps other media may become biased towards one of the two proteins of interest, ultimately preventing production of the desired product, especially if the desired product requires the presence of a chaperone protein for complete maturation. Overall, the increased effectiveness of BB for dual-expression applications results from a combination of a greater number of cells engaged in protein production and an unbiased expression of both POIs.

BB produces natively folded disulfide-rich proteins using galactose as the inducer

BB relies on galactose to induce the T7 system, although galactose is not the canonical inducer of the *lac* operon. Previous literature has described galactose as a capable inducer in some *E. coli* genetic backgrounds such as BL21(DE3)[2,4]. To test BB's auto-expression capability in *E. coli* strains beyond the canonically used BL21(DE3) strain and its derivatives, we used BB to support expression by SHuffle T7 Express *E. coli*, which have been engineered to facilitate production of correctly disulfide-bonded POIs in the cytoplasm[30]. During SHuffle Express engineering, T7 RNA polymerase was integrated into *lacZ*, disrupting bacterial ability to convert lactose to allolactose[30]. Before attempting production of challenging disulfide protein targets, we first tested the ability of various carbon sources including galactose to induce expression in *E. coli B*-derived SHuffle T7 Express using a minimal medium (M9) as a base in agar plates. Each M9 variant contained one additional carbon source (Fig. 4a). Base M9, glycerol supplemented, and glucose supplemented M9 plates served as negative controls as glucose and glycerol do not induce the T7 expression system[2,4].

We used eGFP fluorescence as a readout for POI production in all experiments, allowing on-plate expression measurements. Robust eGFP expression was observed on plates containing lactose, galactose, and IPTG for BL21(DE3) (Fig. 4b). For plates containing glycerol, excess glucose, or the M9/base glucose negative control, there was little fluorescence intensity detected per colony; furthermore, excess glucose downregulated eGFP expression compared to M9/base glucose presumably through catabolite repression (Fig. 4b). As expected, we observed significant eGFP expression in SHuffle T7 Express in response to IPTG (Fig. 4b). However, eGFP production in the lactose plates was considerably reduced compared to IPTG, likely due to disruption of *lacZ* (Fig. 4b). Intriguingly, galactose was as efficient as IPTG in inducing eGFP expression in the SHuffle T7 Express strain (Fig. 4b). We conclude that galactose can induce the T7 expression system through an alternative, non-characterized mechanism. We denote this induction mechanism hereafter as auto-expression.

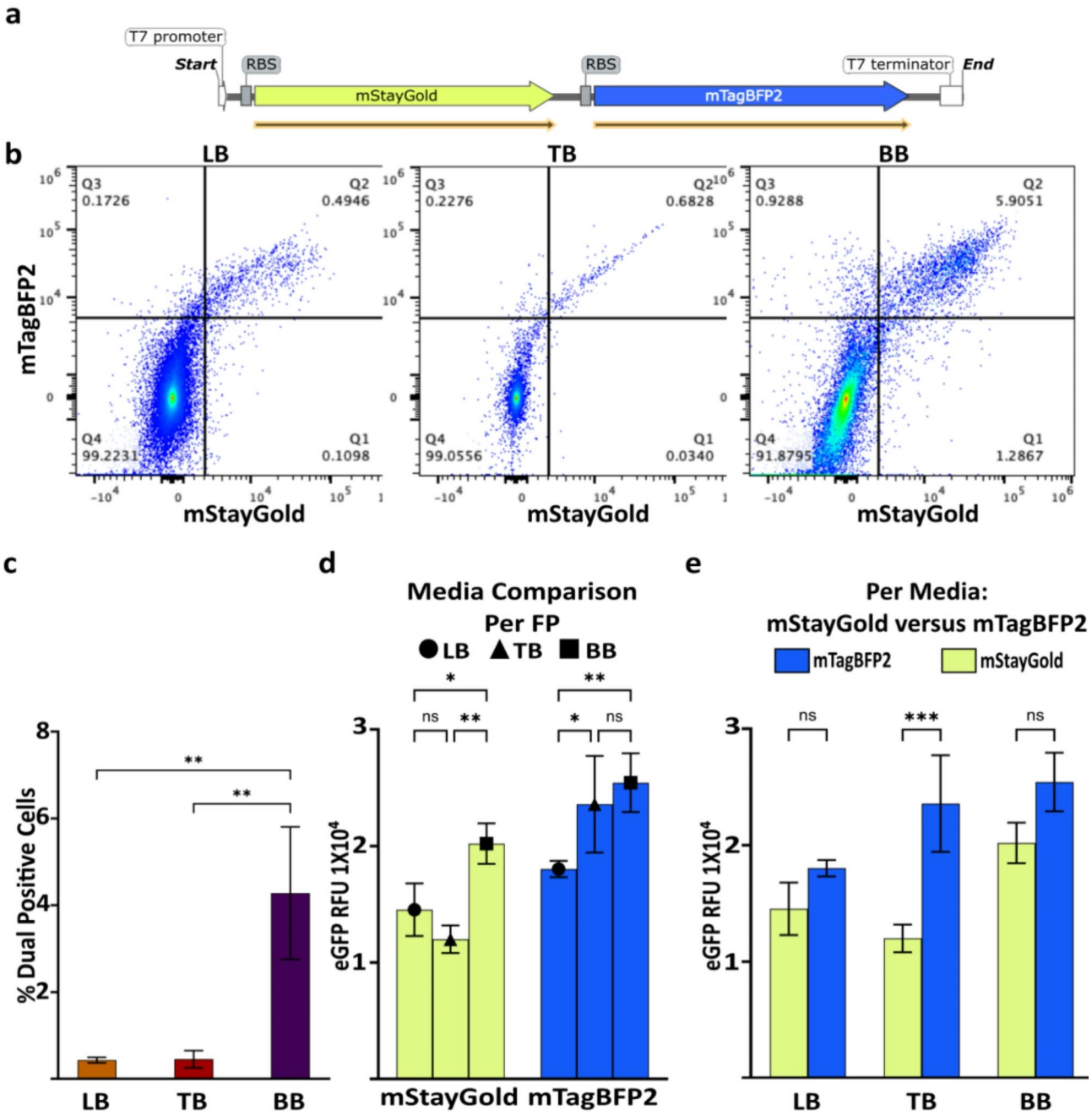

Fig. 3. BB allows a higher percentage of dual-expressing cells. (**a**) Schematic for the dual expression construct expressing mStayGold and mTagBFP2 fluorescent proteins. (**b**) Flow cytometry profiles of dual-expressing *E. coli* cells in indicated media. Cells were divided into four subpopulations: mStayGold + (Q1), mTagBFP2 + (Q3), dual-positive (Q2) and non-fluorescent (Q4). Frequency of each subpopulation is indicated in the corresponding gate. Representative analysis of one replicate for each media. (**c**) Average fraction of dual positive cells for three independent biological replicates for each media. Error bars are SD. Cell fractions were compared by one-way Anova, followed by Tukey post-test, asterisks denote statistical significance ** $P < 0.001$. (**d**) Mean in-cell fluorescence intensity comparison for each fluorescent protein between media. (**e**) Mean in-cell fluorescence intensity comparison between mStayGold and mTagBFP2 in each media. Plotted values in d and e are averages of three independent replicates and error bars are SD; (**d**) and (**e**) analyze the same data and are plotted separately for clarity. Analysis was done with two-way Anova, followed by Sidak's post-test, asterisks denote statistical significance, *** $P < 0.0001$, ** $P < 0.001$, *$P < 0.01$.

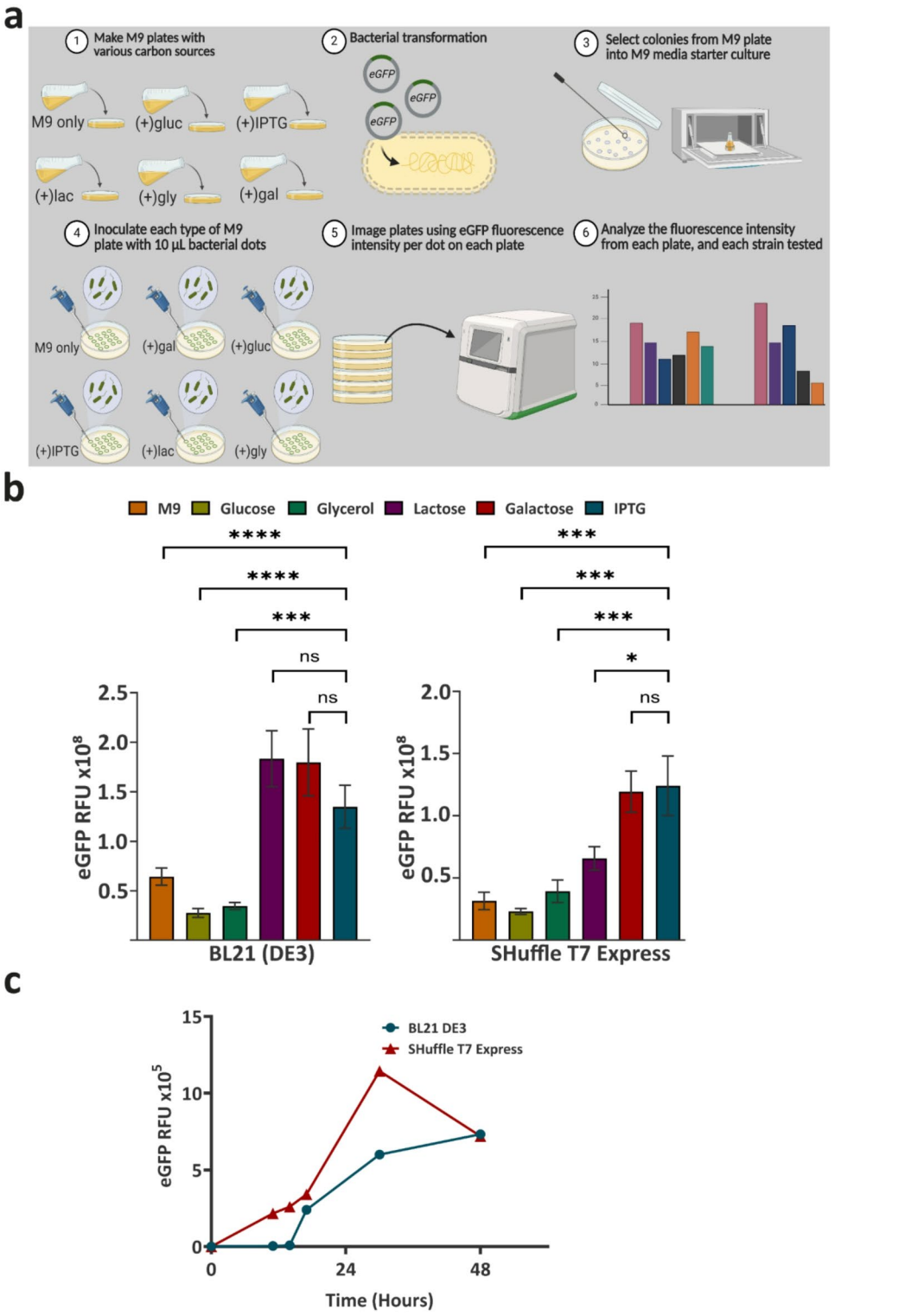

Fig. 4. eGFP production from both auto-induction and auto-expression media in BL21(DE3) and SHuffle T7 Express *E. Coli*. (**a**) workflow for an eGFP in-colony fluorescence assay experiment. This panel was generated with BioRender. (**b**) eGFP in-colony fluorescence levels with respect to carbon source and strain using defined media. Plotted are average values from 12–16 replicates per culture condition, error bars are SD. Relative fluorescence intensities for each strain were compared by one-way Anova followed by Kruskal–Wallis post-test, asterisks denote statistical significance **** $P<0.00001$, *** $P<0.0001$, * $P<0.01$. (**c**) In-cell fluorescence of eGFP in BL21(DE3) and SHuffle T7 Express *E. coli* in BB liquid cultures. This experiment was performed once.

Having established that the SHuffle T7 Express *E. coli* strain is capable of auto-expression in response to galactose on minimal media, we tested whether SHuffle T7 Express would support auto-expression when cultured in BB. We used eGFP for tracking expression, and observed robust eGFP expression, on par with a positive control BL21(DE3) (Fig. 4c). We conclude that SHuffle T7 Express is capable of auto-expression in response to galactose in both minimal media as well as BB.

Following confirmation that BB can be used for auto-expression in the SHuffle T7 express strain, we used this system to produce isolated agonist binding domains (ABD) from GluN1 and GluN2A subunits within N-methyl-D-aspartate (NMDA) receptors[31,32]. NMDA receptors have received considerable attention as targets for therapeutic agents in neurological and psychiatric disorders[33], and we chose GluN1 and GluN2A ABDs as their production represents a significant bottleneck in studies of the structural and pharmacological properties of NMDA receptors. The GluN1 and GluN2A ABDs contain three disulfide bonds each, necessitating the usage of *E. coli* strains such as SHuffle T7 Express for proper folding (Fig. 5a)[31,32]. In a functional NMDA receptor, two GluN1 subunits assemble with two GluN2A subunits to form a heterotetrametric receptor, while recombinantly expressed isolated ABDs form GluN1-GluN2A heterodimers (Fig. 5b)[32,34].

To test whether the proteins produced in BB achieve correct folding and disulfide bonding, we purified the two ABDs separately and crystallized the GluN1/GluN2A ABD-heterodimer complex. The crystal structure of the GluN2A-GluN1 ABD heterodimer in complex with the competitive GluN1 antagonist 7-chlorokynurenic acid (7-CKA) and the GluN2A agonist glutamate was solved at a 2.05 Å resolution (Fig. 5c,d). This structure displays proper disulfide pairing for each NMDA receptor ABD and corroborates previously reported GluN2A-GluN1 ABD heterodimer structures (Fig. 5d). This was further supported by aligning the backbone atoms of this structure with the 5,7-dichlorokynurenic acid (DCKA)/glutamate-bound GluN2A-GluN1 ABD-heterodimer structure (PDB ID 4NF4)[34] (Fig. 5c). This alignment gave an RMSD of 0.185 angstrom, and aligns each ligand, 7-CKA and glutamate, with DCKA and glutamate (Fig. 5e). We conclude that SHuffle T7 Express *E. coli* can be used in conjunction with BB to promote future structural and biochemical research with disulfide bond-containing proteins.

BB increases production of recombinant SpCas9

Having established that BB improves protein production for a broad range of systems, we focused on validating the combination of using BB media and tracking protein expression through fluorescence for optimizing recombinant expression of one of the critical production targets, *Streptococcus pyogenes* CRISPR-associated DNA endonuclease Cas9 (SpCas9)[35–37]. The goal of these experiments was to validate BB's utility for producing an active enzyme rather than compare the protein yields between BB and other media. To monitor production of SpCas9, we used a SpCas9-siriusGFP fusion protein (Fig. 6a)[35]. We expressed SpCas9-siriusGFP in BB and observed robust accumulation of siriusGFP through time (Fig. S5). Following apo-SpCas9 purification as per the published protocol[35], we achieved an 8-fold increase in >95% pure apo-SpCas9 (Fig. 6a) in direct comparison to previously published yields per liter (Table 1, line 3 vs. 5). Importantly, this production improvement was relative to apoSpCas9 protein yields of comparable purity>95% due to the necessity for high SpCas9 purity for challenging genetic engineering applications (i.e. CAR-T cell therapy). We conclude that BB coupled with fluorescence as an in-cell expression readout can be used to produce significantly more, and highly pure (>95%) apo-SpCas9 utilizing conventional purification strategies.

BB-produced SpCas9 is catalytically active in vivo

We validated the activity of BB-expressed SpCas9-siriusGFP by targeting the *ebony* gene in *Drosophila melanogaster*. The *ebony* gene is responsible for body pigmentation and is a widely-used co-CRISPR marker[42]. Ribonucleoprotein complexes (RNPs) were formed with a single guide RNA (sgRNA) sequence targeting the second coding exon in the *ebony* gene (Fig. 6b, S5b)[42]. Syncytial blastoderm embryos of Canton S flies were microinjected with SpCas9/sgRNA complexes targeting the posterior of the embryos to increase the chance of heritable edits in the germ cells (Fig. 6b, S5c). Following microinjection, a subset of embryos was tested for GFP fluorescence, which indicated successful microinjection (Fig. S5d). Remaining embryos were allowed to develop, grown to adulthood, and individually crossed to double balancer ebony flies (*w*; CyO/sp; TM2*, e^s*/TM6b*, e^1), to test for germline transmission of any generated mutations (Fig. S5e). Activity of SpCas9 was assessed by body pigmentation phenotypes within the F_0 and F_1 progeny (Fig. 6c,d). We observed a somatic mosaic phenotype in ~7% of all injected, surviving F_0 animals (Fig. 6c,d). Furthermore, 40% of surviving F_0 were able to transmit the edit in crosses (Fig. 6e). Among the F_1 progeny in 5 crosses with germline transmission of the edit, we observed a consistent percentage of mutants (~30%, Fig. 6f,g). From these data we conclude that production of ≥95% pure SpCas9 can be improved 8-fold using BB. Moreover, we demonstrate that the increased protein yields from BB did not compromise protein activity, as BB-produced SpCas9 was readily capable of catalyzing programmable endonuclease activity in vivo.

Discussion

In this study, we report the development of an auto-expression broth (named BB) that facilitates recombinant protein production without the necessity to add an acute inducer, thus improving upon the classic IPTG induction protocols. The refinement of media formulations for protein production in *E. coli* represents a critical area for enhancing efficiency, particularly in the context of the widely used T7 expression system. This system, developed in 1986, is renowned for its cost-effectiveness and ease of use, but it faces challenges in producing some proteins under standard conditions[2–4]. Our introduction of BB marks an advancement in this field and improves yields for a broad spectrum of recombinant proteins including eGFP and its derivatives and human Cytochrome C. BB achieved superior target protein yields compared to established media including LB, TB, 2xYT, ZYM, and

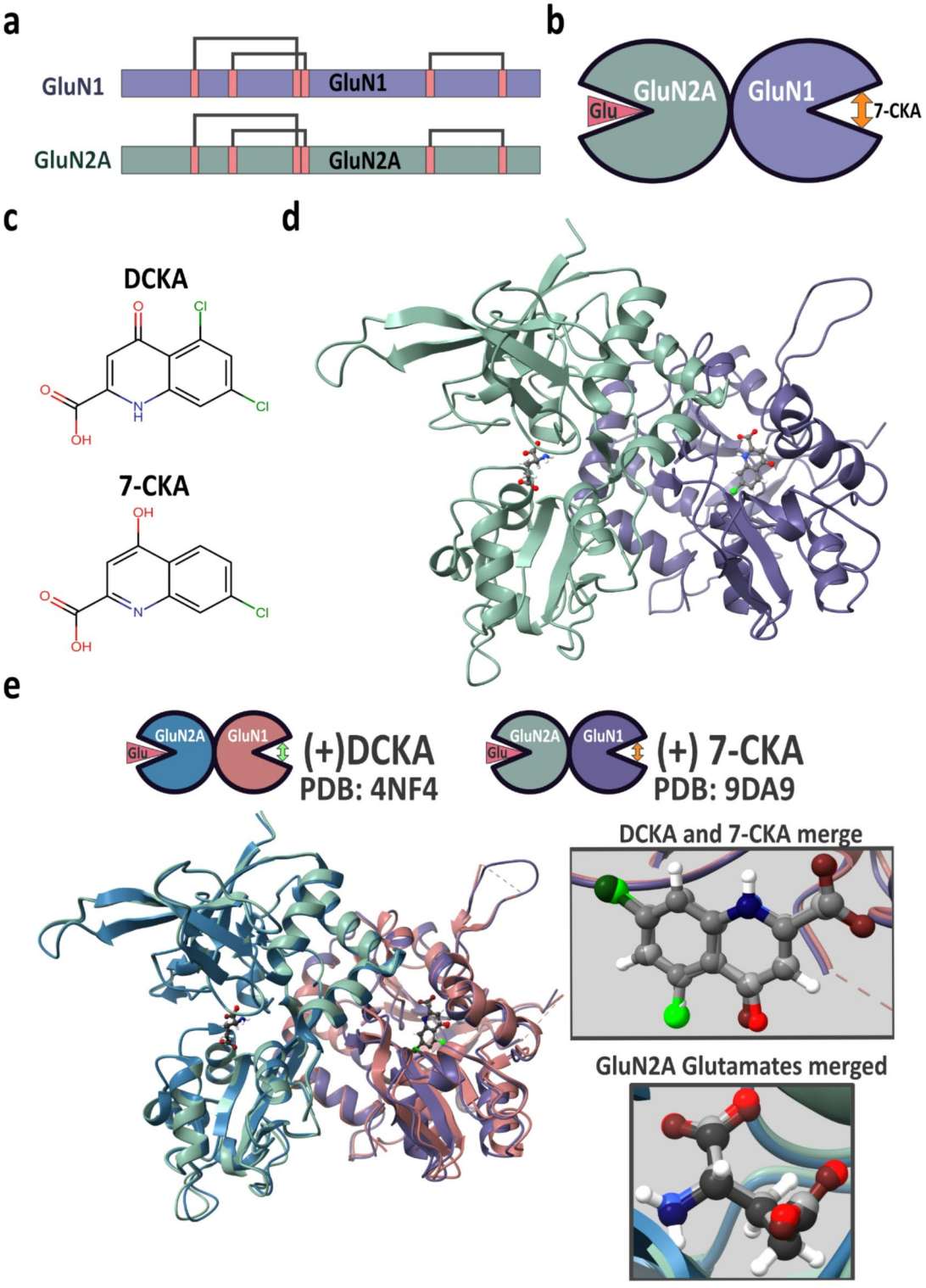

Fig. 5. Production of disulfide bond-containing NMDA receptor agonist binding domains in BB. (**a**) Schematic representing disulfide bonds (solid lines connecting pink bars denoting C420-C454, C436-C455, C744-C798) for GluN1 and GluN2A agonist binding domains (ABDs) as seen in previous structural alignment[34]. (**b**) Cartoon representation of GluN1/GluN2A ABD heterodimer with glutamate bound to GluN2A and 7-chlorokynurenic acid (7-CKA) bound to GluN1. (**c**) Chemical structures of competitive GluN1 antagonists, 5,7-dichlorokynurenic acid (DCKA) and 7-CKA. (**d**) Ribbon diagram of GluN1 (purple)/GluN2A (green) ABD heterodimer in complex with 7-CKA and glutamate (PDB 9DA9) rotated to show ligand-binding pockets. (**e**) Top: schematics of GluN2A/GluN1 ABD heterodimers bound to DCKA/glutamate (red and blue, left) or to 7-CKA/glutamate (grey and purple, right). Bottom: An overlay of DCKA/glutamate bound GluN2A/GluN1 ABD heterodimer (PDB: 4NF4) with 7-CKA/glutamate bound GluN2A/GluN1 ABD heterodimer (PDB 9DA9). Insets: zoomed in agonist binding pockets with bound ligands (also overlaid).

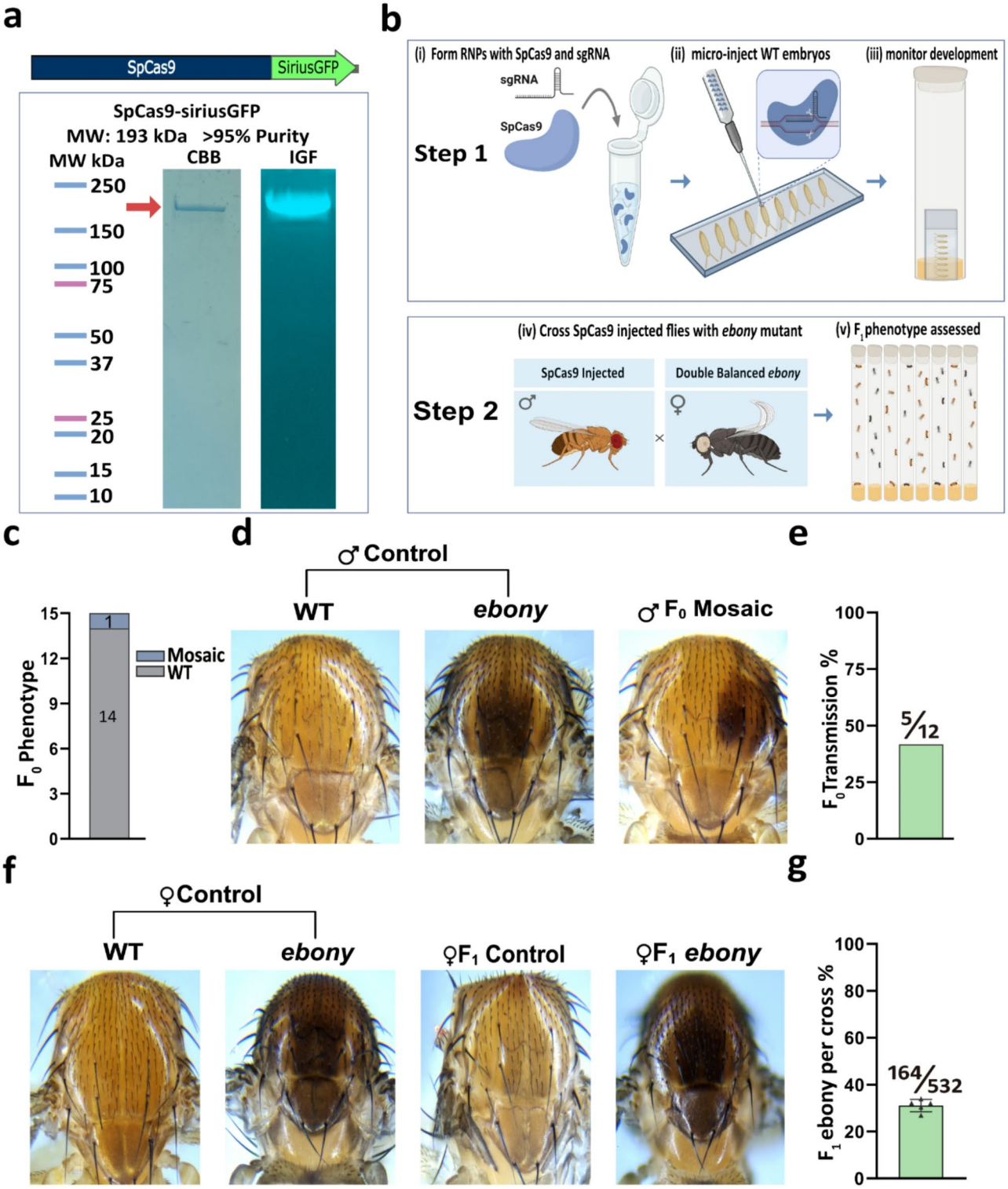

Fig. 6. Production of enzymatically active SpCas9 in BB. (**a**) Top: SpCas9-siriusGFP schematic. Bottom: Total protein gel stained with Coomassie brilliant blue (CBB, left) or in-gel fluorescence (IGF, right) of purified SpCas9-siriusGFP. (**b**) Workflow for the genome editing experiment targeting *ebony* locus in *Drosophila melanogaster*. This panel was generated with BioRender. (**c**) Somatic mosaic frequency based on all survived, SpCas9 injected F_0 flies. (**d**) Pigmentation of WT Canton-S males, double balanced male *ebony* flies, and a mosaic male F_0. (**e**) Germline transmission rate of the Cas9-induced mutation in 12 crosses (see also Fig. S4e). 3/15 surviving injected flies did not produce any progeny. (**f**) Pigmentation of WT or double balanced *ebony* mutant female flies, a representative F_1 female from a control cross (no SpCas9 injection), and a representative phenotype of an F_1 female fly inheriting SpCas9 edit. (**g**) Percentage of mutants in F_1 offspring across the F_0 crosses with germline transmission. Plotted is the average, black triangles represent F_1 mutant percentages in each cross, error bars are SD.

	Final purified protein	Addgene plasmid #	Yield (mg/L)	Purity estimate (%)	Reference
1	SpCas9-NLS-6xHis	#62,731	6.0 $^{\Delta}$	≤ 90*	[38]
2	NLS-SpCas9-6xHis	#62,934	0.7$^{\Delta}$	≥ 95	[39]
3	NLS-SpCas9 – 6xHis	#62,934	1.0	≥ 93*	[40]
4	SpCas9-2xNLS	#69,090	0.16$^{\Delta}$	≥ 90*	[41]
5	SpCas9-siriusGFP-3xNLS	#78,312	8.5	≥ 95	This Study

Table 1. Comparison of purified apo-Cas9 yields per liter of culture. Among numerous Cas9 purification reports, those listed in the table were selected based on approaching ~ 90% apo-SpCas9 purity. *Purity level not suitable for gene therapy applications or high-cost model organism engineering due to inherent editing failure risks associated with contaminants. $^{\Delta}$Publication reports only a single replicate.

MagicMedia™. This increased efficiency is crucial for both reducing operational costs and accelerating research and development in biotechnological applications[4].

The key feature facilitating protein production in BB is its capacity to support high bacterial culture density, while maintaining robust output of the target protein (Fig. 1). BB-expressed eGFP was predominantly found in the soluble fraction, while established auto-induction media appeared to produce copious amounts of insoluble eGFP (Fig. 1). We speculate that this reflects more efficient folding of the POI when produced in BB compared to other high-density cultures. Determining the mechanism mediating improved soluble protein output in BB remains the subject of future studies. BB's effectiveness extends beyond simple expression systems to more complex co-expression processes. For instance, our tests using Hu-Cyt*c* and yeast heme lyase in a co-expression system with BB showed a 100% success rate of POI expression, a marked improvement over the 50–60% success rate observed with traditional media (data not shown). This consistency is particularly advantageous for research where reliable high-yield production of proteins like human cytochrome *c* has been a persistent challenge. Based on the dual-fluorescence reporter expression analyzed by flow cytometry, the increase in effectiveness of dual expression in BB results from a higher fraction of cells that can engage in simultaneous production of two POIs (Fig. 3b,c).

Single-cell analysis of *E. coli* simultaneously expressing two fluorescent proteins by flow cytometry suggested that despite significant increases of the fluorescent protein expression per cell in BB compared to LB and TB, the driving force for higher production levels was the larger fraction of cells producing mStayGold and mTagBFP2 in BB (Fig. 3c–e). Notably, *E. coli* cultured in TB produced a disproportionate amount of mTagBFP2 relative to mStayGold compared to the other two media, suggesting that co-production of multiple POIs is sensitive to medium composition. This differential production of co-expressed proteins might stem from distinct folding efficiency or protein stability in different conditions. Overall, BB was the superior media for dual expression, and simultaneous production of two proteins was optimal when cells were cultured in BB.

Surmounting another protein production challenge, BB in conjunction with *E. coli B*-derived SHuffle T7 Express *E. coli* was found effective for expression of disulfide bonded proteins. We validated correct folding of BB-produced, heterodimerized GluN1/GluN2A NMDA receptor agonist binding domains via X-ray crystallography, and structural analysis (Fig. 5). This holds significant promise in future drug discovery and structural biology endeavors. The compatibility of BB with strains like SHuffle T7 Express, which are essential for producing proteins with correct disulfide bonds, further enhances its utility. We postulate that BB could enable further protein production optimizations for pharmaceutical applications, including therapeutic biologic targets, which often require complex native disulfide bonding patterns for functionality[1].

Our study expands the potential use of BB beyond the commonly utilized BL21(DE3) strain of *E. coli* and provided the first conclusive evidence that galactose is an effective inducer in the SHuffle T7 system independent of contaminant carbon sources found in non-defined media (Fig. 4). Our use of defined M9 medium, which lacks contaminants such as lactose, addressed previous ambiguities caused by undefined media components and enabled a definitive assessment of galactose's role in activating T7 expression[4,30,43,44]. The mechanism through which galactose drives T7 expression system is still unknown. However, these results demonstrate that galactose can be the main inducer carbon source, and in the case of SHuffle T7 Express, functions as a robust auto-expression molecule.

Notably, auto-expression media containing galactose as the inducer (as used in BB) led to a more robust expression in SHuffle T7 Express than that induced by lactose. The genetic background of SHuffle T7 Express, which lacks the *lacZ* gene due to the integration of T7 RNA polymerase at this site, clarifies why lactose was less potent compared to galactose or IPTG. In the absence of β-galactosidase (LacZ), conversion of disaccharide lactose into glucose, galactose, and allolactose is expected to be ineffective[30,45].

The use of galactose as an inducer sets BB apart from most other popular media, which typically use lactose or IPTG. Lactose is cheaper than galactose, and IPTG is more expensive than both. Although BB uses the moderately more expensive galactose as the inducer, the significant improvement in production and broad-spectrum applications make it economically superior. Furthermore, galactose offers greater strain usage diversity than lactose and can be used in Good Manufacturing Practices (cGMP) facilities – unlike lactose. This advantage exists primarily because the market offers very few non-animal derived lactose options, which are more expensive than their animal-free galactose counterparts.

CRISPR-Cas9 and the genome editing process that is enabled by this protein has revolutionized molecular biology and biotechnology[37]. Despite steady progress in genome editing approaches, it is still challenging to produce apo-SpCas9 protein for downstream genetic engineering experiments or gene therapies. Therefore,

commercially available SpCas9 is expensive. Here we have addressed obstacles in SpCas9 production by increasing the final yield of purified, apo-SpCas9 > 8-fold (Table 1 line 3 vs. line 5). Notably, these high yields are achieved without compromising the purity of the enzyme (over 95%). While > 95% purity of SpCas9 is not required for most research laboratory genome editing projects, more demanding gene editing applications in mammalian model systems, or in gene & cell therapies require high-purity SpCas9 to be both cost-effective and feasible[46]. Final FDA approval for producing and purifying SpCas9 is still a notable barrier. Moreover, the production and purification of SpCas9 to be used in a clinical trial is costly and tedious. Generally, FDA approved therapeutic processes are orders of magnitude more complex to execute, and far more expensive. Production improvements on the order of ~ 8x observed for BB-produced SpCas9 in this study, can immediately reduce the effort and costs necessary to generate enough product. As a result, this accelerates pre-clinical gene editing programs, ultimately expediting acquisition of key data necessary for clinical trials.

Using multi-step protocols to improve SpCas9 purity typically leads to significant losses of the protein; therefore, in our yield comparisons we focused on the studies reporting ~ 90% enzyme purity (Table 1). Overall, we find that using BB for protein expression demonstrated higher apo-SpCas9 protein production coupled with optimal purity. Moreover, the BB-expressed SpCas9 shows activity when tested in *Drosophila melanogaster*, thus confirming higher yields of an active enzyme from BB (Fig. 6).

A recent report described an alternative expression and purification protocol for efficient production of SpCas9 in complex with sgRNA (Cas9 RNP)[47]. This system relies on expression of both the SpCas9 ORF and the sgRNA from a single plasmid. Isolation of Cas9 RNP, while more productive, limits the application of the resulting enzyme to a single genetic target. By contrast, apo-SpCas9 allows the user to take advantage of modularity, which is one of the most powerful attributes of CRISPR-Cas9 genome editing technology. Apo-SpCas9 can be assembled with various specific sgRNAs or entire libraries to facilitate diverse gene editing applications or screening. By comparison, utilizing Cas9 RNP requires the full workflow of cloning, plasmid validation, expression, and purification for each unique sgRNA. This will likely not provide a long-term, cost-effective solution for scalable gene cell therapy production, and is less efficient in research laboratory settings where scientists could design, order, and usually receive a new sgRNA reagent within a few days of the order being placed.

Overall, we report that Bosco Broth not only addresses longstanding challenges of protein production in *E. coli* but also sets a new benchmark for the field, offering a versatile, efficient solution that can significantly impact both academic research and industrial biotechnology.

Materials and methods

Molecular cloning

The DNA sequence harboring the 6x Histidine tag, the T7 peptide leader, a TEV proteolytic cleavage site, and eGFP (6xHis-T7Leader-TEV-eGFP) was isolated from Addgene construct (#54762). eGFP was cloned into a pET21a(+) vector (EMD Millipore #69740) backbone by restriction enzyme digest followed by T4 DNA ligation. DNA sequence was confirmed via Sanger Sequencing (Eurofins).

The DNA sequence harboring the 6x Histidine tag, the T7 peptide leader, an X-press affinity tag, and mStayGold (6xHis-T7Leader-XpressTag-mStayGold) was amplified from Addgene construct (#212017). The DNA sequence harboring the 6x Histidine tag, the T7 peptide leader, a TEV proteolytic cleavage site, and mTagBFP2 was amplified from Addgene (#54572). The fragments were assembled via Gibson assembly into the pRSETB vector backbone (Addgene #212017)[28,29]. DNA sequences were confirmed via Sanger Sequencing (Eurofins).

BB recipe

The BB recipe used for all experiments in this study consists of the following:

12 g/L tryptone (Millipore Sigma T9410)
24 g/L yeast extract (Millipore Sigma Y1625)
KH_2PO_4 (0.05–0.1 M)
Na_2HPO_4 (0.05–0.1 M)
NH_4Cl (0.05 M)
Na_2SO_4 (0.005 M)
$MgSO_4$ (0.002 M)
Glycerol (0.54 M)
D-glucose (0.03 M)
D-galactose (0.055 M)
$Na_2C_6H_6O_7$ (0.025 M)
$FeCl_2{\cdot}6H_2O$ (0.1 M)
$CaCl_2$ (0.02 M)
$MnCl_2{\cdot}4H_2O$ (0.035 M)
$ZnSO_4{\cdot}7H_2O$ (0.087 M)
$CoCl_2{\cdot}6H_2O$ (0.0087 M)
$CuCl_2{\cdot}2H_2O$ (0.002 M)
$NiCl_2{\cdot}6H_2O$ (0.002 M)
$Na_2MoO_4{\cdot}5H_2O$ (0.002 M)
$FeSO_4{\cdot}7\ H_2O$ (0.0179 M)
$(NH4)_6Mo_7O_{24}{\cdot}4\ H_2O$ (0.0009 M)
H_3BO_3 (0.186 M)
EDTA (0.05 M)

$CuSO_4 \cdot 5\ H_2O$ (0.0062 M).
The pH of the media is adjusted to the neutral range (7.3–7.6).

eGFP induction and in-cell eGFP measurements

In-cell eGFP fluorescence intensity measurements were conducted as previously described with minor changes[14]. eGFP expression construct was transformed into BL21(DE3) competent cells (New England Biolabs). An overnight starter culture was grown in LB at 37 °C with 100 ug/mL carbenicillin. 50 mL production cultures in non-baffled disposable cell culture Fernbach flasks (Corning 431143) were started in each media (LB, TB, MM, ZYM, BB) with 100 ug/mL carbenicillin by 1:100 dilution of the starter culture and were grown at 30 °C at 250 rpm. 1 mM IPTG was added to LB and TB media at 6 h. For each biological replicate, and each respective time point across all five tested media, 1 mL of culture was collected and pelleted at 10,000*g* for 10 min. Media was decanted, and pellets were resuspended with 1 mL of 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4 (1x PBS) each. Samples were diluted with 1x PBS into range for accurate OD_{600} measurements. Absorbance at OD_{600} and eGFP fluorescence intensity (Ex485/Em516) were then measured by adding 350 µL of diluted sample into a 96-well plate (Grenier 655180) and reading in a Synergy H1 plate reader (Agilent). Blank measurements were subtracted prior to correcting for dilution. 20 mL from all biological replicates, for each type of media were individually spun down at 4,000 rpm at the last time point, 36 hours (Fig. 1a). Media was decanted, and tubes were stored at − 80 °C until eGFP purification and solubility experiments were executed.

eGFP purification and yield quantification

Cells from a single replicate each of the five *E. coli* eGFP liquid cultures (BB, LB, TB, MM and ZYM) were resuspended in lysis buffer: 1 mg mL^{-1} lysozyme (Thermo Fisher) and 1 mg mL^{-1} DNase I (Thermo Fisher) in 1x PBS containing 5 mM imidazole. Cells were incubated for 20 min on a nutator, spun down at 17,000 *g*, 4 °C for 10 min, and clarified protein-containing lysate was collected. Five cycles were then carried out in which the remaining cell pellet was refrozen at − 80 °C, thawed, resuspended in lysis buffer, spun down and the soluble fraction collected. All soluble fractions were pooled. The pooled lysates from each liquid culture were applied to individual gravity columns containing 10 mL Ni-NTA resin (Cytiva) equilibrated in 1x PBS. Clarified cell lysates were flown over each respective gravity column 5 total times at 3 mL min^{-1}. Gravity columns were then washed with 1x PBS for > 50 column volumes at 4 °C. Protein elution during wash steps was monitored using Bradford reagent (Bio-Rad) until no detectable protein was eluted. eGFP was eluted off Ni-NTA resin with 1x PBS containing 300 mM imidazole pH 7.4. Protein elution from the Ni-NTA resin was monitored by Bradford reagent (Bio-Rad) throughout elution. eGFP samples were buffer exchanged into 1x PBS and concentrated by centrifuge ultrafiltration (Amicon Ultra-15, 10k MWCO, Millipore) to a final volume of 10 mL each. An aliquot from each sample was mixed 1:1 with Laemmli Sample Buffer (Bio-Rad), boiled for 5 min, and loaded into a 4–20% Mini-PROTEAN® TGX Stain-Free™ Gel (Bio-Rad). Stain-free chemistry incorporates a trihalo compound that covalently binds to tryptophan residues and allows the visualization of proteins following electrophoresis by making the proteins fluorescent when exposed to UV light[48]. The rest of the samples were flash frozen with liquid nitrogen for further analysis. The stain-free gel was imaged on a BioRadChemiDoc™ MP imaging system and analyzed using ImageLab (v 5.1) software. Total intensities for all bands in each lane were quantified and the fraction of total protein corresponding to the eGFP band was determined. Freshly thawed remaining samples were used in a BCA assay[49] (Thermo Fisher) to measure the total protein concentration using a Synergy H1 plate reader (Agilent) to measure A_{562} of 350 µL samples in wells of a clear 96-well plate (Greiner 655180). The percentage of eGFP signal measured above was applied to the total protein yields measured by the BCA assay. Final corrected eGFP yields were scaled up from mg/20 mL to mg L^{-1} for all 5 media.

Immunoblotting

To detect 6xHis-eGFP in the soluble and insoluble fractions, bacterial lysates either from the soluble fraction of clarified lysate, or from resolubilized insoluble fractions from the cell pellets (8 M urea with buffer exchange back to 1x PBS) were separated on a 4–20% SDS-PAGE gel (Bio-Rad), and proteins were transferred to Immobilon-P PVDF membrane (EMD Millipore). After blocking in TBS/0.1% Tween 20/5% non-fat dry milk, the blots were probed with mouse anti-6xHistidine primary antibody (Sigma) followed by the secondary goat anti-mouse IgG2a labeled with HRP (Southern Biotech) diluted in blocking solution. Blots were developed using Luminata Crescendo Western HRP substrate (EMD Millipore) and recorded on ChemiDoc MP Imaging System (Bio-Rad).

Human cytochrome c expression and purification

Human cytochrome *c* (Hu-Cyt*c*) was expressed and purified as described previously[19]. The pBTR (Human Cyt*c*) vector, a derivative of the pBTR1 expression vector, with the yeast iso-1-Cyt*c* gene replaced with a synthetic Hu-Cyt*c* gene, was used for expression[17,50,51]. This vector co-expresses yeast heme lyase, which covalently attaches heme to the Cys-Ser-Glu-Cys-His heme recognition sequence of Hu-Cyt*c* in the cytoplasm of *E. coli*. Competent *E. coli* BL21(DE3) cells were transformed with the pBTR(Human Cyt*c*) vector. The colonies were washed off the selection plate and grown in 1 L of BB for 24 to 36 h at 37 °C with 100 ug/mL carbenicillin. Cells were then pelleted and suspended in a lysis buffer containing 10 mM Tris pH 7.5, 500 mM NaCl, 2 mM PMSF, and a small quantity of DNase. Lysis was carried out via sonication with a Qsonica Q700 sonicator. The lysate was cleared by centrifugation and then incubated overnight while stirring with 12% ammonium sulfate to precipitate impurities. Centrifugation separated out the precipitate, and the supernatant was dialyzed overnight twice at 4 °C, both times in 3 L of buffer containing 12.5 mM sodium phosphate, pH 7.2, 1 mM disodium EDTA, and 2 mM β-mercaptoethanol (BME). The dialyzed protein solution was loaded onto a CM Sepharose column that had been equilibrated in 50 mM sodium phosphate buffer, pH 7.0, 1 mM EDTA, and 2 mM BME. The column

was washed with 5 column volumes of this same buffer to remove any additional protein impurities, and the final protein was eluted with a linear gradient from 0 to 0.8 M NaCl in the same buffer. The eluate was concentrated and exchanged into a low salt buffer, 50 mM sodium phosphate, pH 7.0, by centrifuge ultrafiltration (Amicon Ultra-15, 3k MWCO, Millipore) before being flash frozen and stored at -80 °C. On the day of the experiment, protein was thawed and further purified with an AKTA-prime plus chromatography system (GE Healthcare Life Sciences) equipped with a HiTrap SP HP 5 mL cation exchange column and a gradient from 0 to 0.3 M NaCl in 50 mM sodium phosphate buffer, pH 7.0, over 30 min at a flow rate of 1 mL/min. The purified protein was concentrated via centrifuge ultrafiltration (Amicon Ultra-15, 3k MWCO, Millipore). Protein yields were calculated from the final purified, concentrated protein by UV–Vis Spectroscopy prior to peroxidase activity assays[52].

Peroxidase activity measurements

The peroxidase activity of human cytochrome c (Hu-Cyt*c*) was measured via conversion of guaiacol to tetraguaiacol which strongly absorbs at 470 nm[19,27]. Hu-Cyt*c* was oxidized with potassium ferricyanide, $K_3[Fe(CN)_6]$, at room temperature for 7 to 10 minutes, and then the $K_3[Fe(CN)_6]$ was separated from the protein using a G25 Sephadex desalting column (Cytiva) equilibrated with 50 mM sodium phosphate buffer at pH 7.0. The buffer solution was degassed under argon for 30 to 45 minutes to inhibit non-enzymatic oxidation of guaiacol by dissolved O_2. Solutions of 4 µM protein, 400 µM guaiacol, and 100 ± 4 mM hydrogen peroxide were made in the degassed buffer. The H_2O_2 concentration used was selected so that it was above the Michaelis constant, K_m, of H_2O_2 for peroxidase activity which is about 37 mM for human Cytc[21–24,27,53]. Guaiacol and H_2O_2 concentrations were measured by absorbance using extinction coefficients, $e_{274} = 2150\ M^{-1}cm^{-1}$, and $e_{240} = 41.5\ M^{-1}cm^{-1}$, respectively[25–27]. The protein solution, degassed buffer, and guaiacol solutions in different concentrations were mixed immediately prior to taking measurements. The protein/guaiacol solution was then automatically mixed in a 1:1 ratio with the 100 mM peroxide solution using an Applied Photophysics SX20 stopped-flow spectrometer. Time-dependent changes were recorded at 470 nm (A470). Final protein concentration was 1 mM, and final H_2O_2 concentration was 50 mM. All data measurements were taken at 25 ± 0.1°C. Five traces were recorded per guaiacol concentration. A470 was plotted as a function of time for each trial, and the highest slope of the linear region was used to determine initial enzyme velocity, v. The slopes were averaged, then multiplied by 4 since oxidation of guaiacol removes 4 electrons, divided by final protein concentration and multiplied by 26.6 $mM^{-1}cm^{-1}$, the extinction coefficient for tetraguaiacol (3,3'-dimethoxy-4,4'-biphenoquinone)[19,21,53]. This gives v/[Cyt*c*] values as a function of guaiacol concentration. The Michaelis-Menten equation was fit to the data using SigmaPlot v. 13 to extract the Michaelis constant, K_m, as well as the catalytic rate constant, k_{cat}.

Dual expression of fluorescent proteins in *E. coli*

BL21(DE3) competent cells were transformed and cultivated with carbenicillin resistance[28]. Each flask for each replicate was inoculated with 1:100 starter culture from overnight cultivation (14 h) for BB, TB, and LB media. 50 mL total cultures were used for each replicate and cells were grown at 37 °C until reaching OD_{600} 0.6. Then, TB and LB cultures were induced with 0.1 mM IPTG, and all cultures were shaken at 25 °C at 250 rpm for five days. Cells were harvested at 4000 rpm, media was decanted, and samples were flash frozen with liquid nitrogen, then stored at − 80 °C until flow cytometry experiments.

Flow cytometry

Samples were analyzed with the Attune NxT (Life Technologies, Grand Island, NY, USA) flow cytometer. Forward and side scatter properties were displayed on a logarithmic scale with the forward scatter threshold set to 1000 for all samples analyzed. mStayGold green fluorescent protein was measured with the BL1 laser line, excitation at 488 nm with emission detection at 530 nm with a 30 nm bandpass filter. mTagBFP2 was measured by the VL1 laser line excitation at 405 nm with emission detection at 440 nm with a 50 nm bandpass filter. Thawed *E. coli* cells were resuspended in 1x PBS supplemented with 1mM $MgSO_4$ and 0.1 mM $CaCl_2$, and further diluted 1:50 in PBS. 100 µL of the cell suspension was analyzed at a flow rate of 25 µL/minute for each replicate. Data was analyzed using FlowJo V10.10 (BD Life Sciences, Franklin Lakes, NJ, USA). Cells of all replicates were gated for intact single cells based on the forward and side scatter (Fig. S3) and universal gates for ± mStayGold and ± mTagBFP2 were set based on the negative control cells cultured in BB, but not expressing fluorescent proteins (Fig. S3).

Imaging *E. coli*

Live-cell imaging of the bacterial cells was performed on a Zeiss LSM 880 laser scanning confocal microscope with a Plan-Apochromat 63x/1.4 NA DIC M27 oil immersion objective. A 405 nm laser was used to excite mTagBFP2 with fluorescence detected from 410 to 498 nm. The excitation for mStayGold was a 488 nm laser with fluorescence being detected from 498 to 598 nm. All cultures were imaged using the same settings in ZenBlack 2.3. Maximum intensity projections were generated from z-stacks of *E. coli* in each respective medium using ZenBlue 3.6.

eGFP expression measurements using M9 agar plates

All M9 minimal media combinations were prepared following the base recipe[20]. Each treatment was supplemented with glucose to a final concentration of 0.5% (w/v). Where appropriate, additional carbon sources (glycerol, glucose, lactose, and galactose) were also added to a final concentration of 1% (w/v); for the positive control, IPTG was supplemented to a final concentration of 0.1 mM. Plates meant for BL21(DE3) additionally contained carbenicillin at a final concentration of 100 µg mL^{-1}, and plates meant for SHuffle® T7 Express (New England Biolabs C3029J) contained carbenicillin and streptomycin at final concentrations of 100 µg mL^{-1} and 25 µg

mL^{-1}, respectively. Starter cultures of SHuffle® T7 Express and BL21(DE3) *E. coli* were prepared by inoculating eGFP transformants into 30 mL M9 medium (0.5% w/v glucose) containing the appropriate antibiotics described above. Flasks were incubated in rotary shakers set to 250 rpm at 30–37 °C for SHuffle® T7 Express or BL21(DE3) *E. coli*, respectively. The BL21(DE3) culture was grown for 24 h under these conditions, whereas SHuffle® T7 Express needed 72 h to achieve the same qualitative density. Following this initial growth, cell suspensions were plated in parallel on six M9 modifications with the appropriate antibiotic selection: M9 (0.5% w/v glucose), M9 + glucose (1% w/v), M9 + glycerol (1% w/v), M9 + lactose (1% w/v), M9 + galactose (1% w/v), and M9 + IPTG (0.1 mM). For each plate, sixteen 10 µL aliquots of homogenous culture were dotted in a 4 × 4 grid and allowed to dry under a laminar flow hood. All plates were incubated at 30 °C and were imaged after 6 days (SHuffle T7 Express) or 8 days (BL21(DE3)) of culture using a BioRad ChemiDoc™ MP imaging system. To assess in-colony eGFP fluorescence, images of each plate were captured using a 530/28 filter under UV illumination using the same exposure settings and image area. Images were analyzed using the companion ImageLab (v5.1) software, performing automatic background subtraction and quantifying integrated colony fluorescence intensity for each colony without normalization for colony size. Non-uniform or fused colonies were excluded from the analysis, resulting in 12–16 values per condition. Average fluorescence intensity was calculated for each condition.

Expression and purification of NMDA receptor ABDs

The GluN1 and GluN2A agonist binding domains (ABDs) were purified using a similar protocol to previously described with several adjustments[54]. Briefly, the GluN1 S1S2 ABD contains residues 394–544 and 663–800 from the full-length GluN1 joined by a Gly-Thr dipeptide linker and the GluN2A ABD contains residues 402–539 and 661–802 from the full-length GluN2A joined by a Gly-Thr dipeptide linker as previously described[34,54,55]. The GluN1 ABD was fused with 6x-Histidine tag and thrombin proteolytic cleavage site (6xHis-Thrombin-GluN1 ABD)[34]. The GluN2A ABD consisted of a fusion between 6xHis-small ubiquitin-like modifier and GluN2A ABD (6xHis-SUMO-GluN2A ABD). Each ABD protein was expressed using SHuffle® T7 Express *E. coli* in BB at 25 °C–20 °C for 24 h respectively. Each construct contained a 6xHis-tag at the amino-terminus and was captured by applying clarified cell lysate to a 5 mL HiTrap IMAC HP column (Cytiva). Following loading of the GluN1 and GluN2A ABDs onto the IMAC, proteins were eluted with a linear gradient of imidazole concentrations, and peak fractions containing proteins were pooled and digested with either thrombin protease or ubiquitin ligase protease-1 (ULP-1) to remove N-terminal-tag sequences and isolate GluN1 or GluN2A ABD protein fragments respectively. Post-cleavage, GluN1/GluN2A ABD proteins were further applied to a 5 mL Hi-Trap SP column (Cytiva) to remove major contaminants, and GluN1/GluN2A ABD protein fragments were further purified with Superdex 200 10/300 GL (Cytiva) via size-exclusion chromatography in a buffer consisting of in 10 mM HEPES, 100 mM NaCl, 1 mM glycine, 1 mM L-glutamate, pH 7.0.

Crystallography for GluN1/GluN2A ABDs

Purified GluN1A and GluN2A ABD proteins were mixed at a 1:1 ratio and incubated O/N at 4 °C before the GluN1/N2A complex was subjected to another round of purification using a Superdex 200 10/300 GL size exclusion column (Cytiva) in 10 mM HEPES-NaOH, 100 mM NaCl, 1 mM glycine, 1 mM L-glutamate, pH 7.0. Complexed proteins were concentrated to 4–7 mg/mL and used to set 1:1 drops against a reservoir consisting of 0.2 M ammonium acetate, pH 6.8 with 14–18% PEG 4000 using the hanging drop diffusion method. Following 48 h of crystal growth, two rounds of crystal soaking were conducted. In the first round of soaking, drops were soaked with reservoir solution including 100 µM glutamate, and 100 µM 7-Chlorokynurenic acid (7-CKA). After at least 24 h of soaking, the drops were then soaked for a second time with reservoir solution including 100 µM glutamate, and 300 µM 7-CKA for a minimum of 24 h prior to harvesting. Crystals were cryoprotected in reservoir solution containing 20% (v/v) glycerol and flash-frozen in a 100 K nitrogen gas stream. Diffraction data were collected at the Stanford Synchrotron Radiation Lightsource (SSRL-SMB) on the 12 – 2 beamline. Images were processed using autoPROC[56]. Initial phases were determined by molecular replacement in PHASER[57] using a published glycine/glutamate bound GluN1/2A ABD structure (PDB ID 4NF4)[34] as search model. The initial models were fit into a $2mF_o$-DF_c map with COOT[58] and subjected to one cycle of rigid body refinement using procedures in the PHENIX program suite[59]. Subsequently, the model was refined by iterative model rebuilding in COOT and refinement with PHENIX. L-glutamate and the 7-CKA ligand were located in the structure using the mF_o-DF_c map[60]. Data collection and refinement statistics are shown in Table S1. Coordinates and diffraction data have been submitted to the Protein Databank with ID 9DA9.

SpCas9 expression and purification

Recombinant nuclease was purified according to a previously published protocol with minor modifications[35]. Briefly, pET-28b-6xHis-MBP-TEV-Sp-Cas9-siriusGFP-3xNLS (Addgene #78312) was transformed in Rosetta (DE3) pLysS *E. coli* competent cells (Novagen), and single colonies were inoculated and grown overnight. 500mL BB cultures were inoculated by 1:100 dilution of the starter cultures and grown at 18 °C while siriusGFP fluorescence was monitored through time (Fig. S4a). Bacterial pellets were then re-suspended in the lysis buffer (20 mM Tris [pH 8], 500 mM NaCl, 5 mM $MgCl_2$, and 5 mM imidazole) supplemented with lysozyme (0.5 mg/mL), incubated for 20 min at 4 °C while shaking and lysed by high-pressure cell disruption. After clarification by centrifugation (40,000×*g*, 45 min, 4 °C), the supernatants were flown over an IMAC 5mL HiTrap 3 times in total for initial affinity purification step. Following washes with 50 mL of PBS, SpCas9 was eluted with 400 mM imidazole step gradient, and dialyzed into 20 mM Tris [pH 8], 20 mM NaCl, 5 mM $MgCl_2$ in the presence of TEV protease overnight at 4 °C. The rest of the protocol adheres faithfully to[35]. According to protocol, two further steps of purification were performed: ion-exchange (IEX) chromatography (HiTrap SP FF column, GE Healthcare) followed by size-exclusion chromatography (HiLoad 16/600 Superdex 200 PG, GE Healthcare) using an AKTA Pure 25 FPLC system (GE Healthcare). After the final elution, fractions containing SpCas9 were

pooled, concentrated using a centrifugal concentrator (100,000 MWCO) to reach at least 10 mg/mL, and stored in aliquots at − 80 °C until use. The expression and purification was performed 3 times, with the average yield of 8.5 mg/ml.

Preparation of SpCas9-sgRNA RNP for microinjection

For an injection preparation, 11.9 µg Cas9-siriusGFP in 25mM Tris, 300mM NaCl, 2 mM $MgCl_2$, and 1 µm DTT was mixed with 2.36 µg sgRNA in TE buffer (1:1 molar ratio), and KCl was added to a final concentration of 150mM in a final reaction volume of 10 µL. RNP assembly reaction was incubated while being centrifuged (17,000 RCF) for 30 min at 4 °C. For all genome editing experiments, ribonucleoprotein complexes (RNPs) were formed between SpCas9 and sgRNA targeting coding exon 2 of the *ebony* (CG3331) locus (Fig. S6a).

Full sgRNA sequence (IDT) with complementary crRNA underlined:

5′CCACAAUUGUCGAUCGUCAGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU-3′.

Drosophila melanogaster microinjection

Wildtype Canton-S flies were maintained in 6 oz bottles (Flystuff #32-130) with 50 mL standard fly food. Approximately 300 adults aged 2–5 days were transferred to embryo collection chambers (Flystuff #59-105) with 2% apple juice agar collection plates (35 mm petri dish) topped with yeast paste for 30 min. Embryos were collected, washed with ddH_2O, and arranged on a 18 × 18 mm coverslip in olive oil. Sutter P-1000 horizontal micropipette puller and Sutter BV-10 Microelectrode Beveler were used to pull and bevel aluminosilicate glass capillary tubes (Sutter, #AF100-64-10), respectively. Pipettes were back loaded with injection mix. Syncytial stage embryos were injected in the posterior end (Fig. S4C) at 20 °C with a trinocular microscope (ACCUSCOPE EXI-310) and PM1000 Microinjector (MicroData Instrument Inc.) on manual setting, facilitated by an N_2 gas compressor and controlled by a hydraulic micromanipulator set up (Narishige #MMO-4, #MMN-1, #P-12, #GJ-8, #IP). Coverslip with injected embryos was then washed with 95% ethanol and quickly rehydrated with ddH_2O. Coverslips were transferred to food vials (Flystuff, #32–117) and incubated at 26 ± 2 °C, 65% humidity with a 12 h light/dark cycle.

Drosophila crosses and phenotypic characterization

Individual injected surviving adults were crossed to 3 double balanced *ebony* flies (w[*]; CyO/sp; TM2, e^s/TM6b, e^1) in 16 × 100 mm glass isolation vials (Thermo Fisher #14-961-29) with 1.5mL standard fly food. Canton-S controls were isolated as pupae and subject to the same cross set-up in biological triplicate. Cross vials were flipped daily for 10 days; incubated at 26 ± 2 °C, 65% humidity with a 12 h light/dark cycle. F_1 adults were screened for the ebony phenotype (e^{edited}/TM2, e^s or e^{edited}/TM6, e^1; Fig. S4E).

Imaging fly embryos

Canton-S embryos were injected with Cas9-siriusGFP. Injected embryos were mounted on a glass slide (VWR #16005-106) in halocarbon 27 oil, overlaid with a coverslip sealed with nail polish and imaged using a Zeiss LSM 880 laser scanning confocal microscope with a Pln Apo 20x/0.8 NA air objective. Excitation wavelength of 488 nm was used to image with detection wavelength range of 493–530 nm. Images were acquired using ZenBlack 2.3 software and pseudo colored turquoise in FIJI.

Imaging adult flies

Adults were imaged using a Sony a7R III camera mounted on an Olympus BH-2 microscope with Olympus 4xDPlan 0.10 160/0.17 objective and M2XNEX 2x Sony E-mount adapter. The specimens were lit bilaterally with Hera AKOD20/CW 5-Watt LED Luminaire lamps. Flies were decapitated and mounted on glass slides (VWR #16005-106) with clear Gorilla Super Glue Gel. Images were stacked using manual fine-focus increments and combined in Helicon Focus 7.6.1 Pro using Method C with smoothing set to 2.

Data availability

Macromolecular structure: the atomic coordinates and structure factors have been deposited in the Protein Data Bank, PDB ID code 9DA9, https://doi.org/10.2210/pdb9DA9/pdb. Remaining data are provided within the manuscript and supplementary information files.

Received: 10 October 2024; Accepted: 24 February 2025
Published online: 15 March 2025

References

1. Goeddel, D. Expression in *Escherichia coli* of chemically synthesized genes for human insulin. *PNAS* **76**(1), 106–110 (1979).
2. Studier, W. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189** (1), 113–130 (1985).
3. Zulkifly, N. A. H. Optimisation of recombinant TNFα production in *Escherichia coli* using GFP fusions and flow cytometry. *Front. Bioeng. Biotechnol.* **11**, 1171823 (2023).
4. Studier, W. Protein production by auto-induction in high-density shaking cultures. *Protein Exp. Purif.* **41** (1), 207–234 (2005).
5. Du, F. Regulating the T7 RNA polymerase expression in *E. coli* BL21 (DE3) to provide more host options for recombinant protein production. *Microb. Cell. Fact.* **20**, 189 (2021).
6. Kolb, A. Transcriptional regulation by cAMP and its receptor protein. *Annu. Rev. Biochem.* **62**, 749–797 (1993).
7. Busby, S. Transcription activation by catabolite activator protein (CAP). *J. Mol. Biol.* **293** (2), 199–213 (1999).
8. Kataoka, K. IPTG-independent autoinduction of extracellular matrix proteins using recombinant *E. coli* as the expression host. *Polym. J.* **53**, 385–391 (2021).

9. Xu, J. Galactose can be an inducer for production of therapeutic proteins by auto-induction using *E. coli* BL21 strains. *Protein Exp. Purif.* **83** (1), 30–36 (2012).
10. Beckwith, J. Regulation of the Lac Operon: recent studies on the regulation of lactose metabolism in *Escherichia coli* support the Operon model. *Science* **156**, 597–604 (1967).
11. Hoffman, F. Stress induced by recombinant protein production in *Escherichia coli*. *Adv. Biochem. Eng.* **89**, 73–92 (2004).
12. Wagner, S. *Consequences of Membrane Protein Overexpression in Escherichia coli* (Molecular & Cellular Proteomics, 2007).
13. Waldo, G. S. Rapid protein-folding assay using green fluorescent protein. *Nat. Biotechnol.* **17**, 691–695 (1999).
14. Schilling, P. J. Improved designs for pET expression plasmids increase protein production yield in *Escherichia coli*. *Commun. Biol.* (2020).
15. Santillan, M. Quantitative approaches to the study of bistability in the *lac* Operon of *Escherichia coli. J. R Soc. Interface* **5**(Suppl 1), S29–39 (2008).
16. Valbuena, M. F. A photostable monomeric superfolder green fluorescent protein. *Traffic: Mov. Front. Cell. Biol.* **21**(8), 534–544 (2020).
17. Olteanu, A. Stability and apoptotic activity of recombinant human cytochrome C. *Biochem. Biophys. Res. Commun.* **312**, 733–740 (2003).
18. Rumbley, J. R. Recombinant equine cytochrome c in *Escherichia coli*: High-level expression, characterization, and folding and assembly mutants. *Biochemistry* **41**(47), (2002).
19. Nold, S. M. Effect of a K72A mutation on the structure, stability, dynamics, and peroxidase activity of human cytochrome C. *Biochemistry* **56** (26), 3358–3368 (2017).
20. Sambrook, J. *Molecular Cloning: A Laboratory Manual* 2nd edn, Vol. 3 (Cold Spring Harbor Laboratory Press, 1989).
21. Diederix, R. E. M. Peroxidase activity as a tool for studying the folding of c-type cytochromes. *Biochemistry* **41** (43), 13067–13077 (2002).
22. Diederix, R. E. M. Kinetic stability of the peroxidase activitity of unfoled cytochrome C: Heme degradation and catalyst inactivation by hydrogen peroxide. *Inorg. Chem.* **42**, 7249–7257 (2003).
23. Radi, R. Cytochrome c-catalyzed oxidation of organic molecules by hydrogen peroxide. *Arch. Biochem. Biophys.* **288** (1), 112–117 (1991).
24. Samsri, S. *Influence of Cysteine-directed Mutations at the Ω-loops on Peroxidase Activity of Human Cytochrome C 709* (Archives of Biochemistry and Biophysics, 2021).
25. Noble, R. The reaction of ferrous horseradish peroxidase with hydrogen peroxide. *J. Biol. Chem.* **245** (9), 2409–2413 (1969).
26. Nelson, D. P. Enthalpy of decomposition of hydrogen peroxide by catalase at 25 °C (with molar extinction coefficients of H_2O_2 solutions in the UV). *Anal. Chem.* **49** (2), 474–478 (1972).
27. Doerge, D. R. Identification of the colored guaiacol oxidation product produced by peroxidases. *Anal. Biochem.* **250** (1), 10–17 (1997).
28. Ando, R. StayGold variants for molecular fusion and membrane-targeting applications. *Nat. Methods*. **21**, 648–656 (2024).
29. Subach, O. M. An enhanced monomeric blue fluorescent protein with the high chemical stability of the chromophore. *PLoS One* **6**(12), e28674 (2011).
30. Lobstein, J. SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microb. Cell Factor.* **11**, 56 (2012).
31. Lind, G. E. Structural basis of subunit selectivity for competitive NMDA receptor antagonists with preference for GluN2A over GluN2B subunits. *PNAS* **114**(33), E6942–E6951 (2017).
32. Furukawa, H. Subunit arrangement and function in NMDA receptors. *Nat. Biotechnol.* **438**, 185–192 (2005).
33. Hansen, K. Structure, function, and pharmacology of glutamate receptor ion channels. *Pharmacol. Rev.* **73**, 298–487 (2021).
34. Jespersen, A. Structural insights into competitive antagonism in NMDA receptors. *Neuron* **81** (2), 366–378 (2014).
35. Fleitas, A. L. A robust expression and purification method for production of Sp Cas9-GFP-MBP fusion protein for in vitro applications. *Methods Protocols*. **5** (3), 44 (2022).
36. Burger, A. Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development* **143**(11), 2025–2037 (2016).
37. Jinek, M. A programmable Dual-RN: Guided DNA endonuclease in adaptive bacterial immunity. *Science* **337** (6096), 816–821 (2012).
38. D'Astolfo, D. Efficient intracellular delivery of native proteins. *Cell* **161** (3), 674–690 (2015).
39. Martin-Martin, I. Optimization of sand fly embryo microinjection for gene editing by CRISPR/Cas9. *PLoS Negl. Trop. Dis.* **12**(9), e0006769 (2018).
40. Rajagopalan, N. A Two-Step method for obtaining highly pure Cas9 nuclease for genome editing, biophysical, and structural studies. *Methods Protocols* **1**(17). (2018).
41. Dilworth, D. *Purification of Recombinant Cas9* (Open Lab Notebooks, 2018).
42. Kane, N. S. Efficient screening of CRISPR/Cas9-Induced events in Drosophila using a Co-CRISPR strategy. *G3 Genes Genomes Genet.* **7**, 87–93 (2017).
43. Pinto, C. The selective advantage of the *lac* operon for *Escherichia coli* is conditional on diet and microbiota composition. *Front. Microbiol.* **12**, 709259 (2021).
44. Fethi-Roudsari, M. Auto-induction for high level production of biologically active reteplase in *Escherichia coli*. *Protein Exp. Purif.* **151**, 18–22 (2018).
45. Hall, B. G. Transgalactosylation activity of Ebg beta-galactosidase synthesizes allolactose from lactose. *J. OfBacteriol.* **150** (1), 132–140 (1982).
46. Bloomer, H. CRISPR/Cas9 ribonucleoprotein-mediated genome and epigenome editing in mammalian cells. *Adv. Drug Deliv. Rev.* **181**, 114087 (2022).
47. Qiao, J. Co-expression of Cas9 and single-guided RNAs in *Escherichia coli* streamlines production of Cas9 ribonucleoproteins. *Commun. Biol.* (2019).
48. Taylor, S. *The Design of a Quantitative Western Blot Experiment* (BioMed Reseasrch International, 2014).
49. Smith, P. K. Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **150** (1), 76–85 (1985).
50. Pollock, W. B. R. Bacterial expression of a mitochondrial cytochrome c. Trimethylation of Lys72 in yeast iso-1-cytochrome c and the alkaline conformational transition. *Biochemistry* (1998).
51. Mauk, A. G. Spectroscopic properties of a mitochondrial cytochrome c with a single thioether bond to the heme prosthetic group. *Biochemistry* (2002).
52. Margoloash, E. Spectrum of horse-heart cytochrome C. *Biochem. J.* **71** (3), 570–572 (1959).
53. Diederix, R. E. M. *The Peroxidase Activity of Cytochrome c-550 from Paracoccus versutus* (European Journal of Biochemistry, 2001).
54. Yi, F. Structural basis for negative allosteric modulation of GluN2A-Containing NMDA receptors. *Neuron* **91** (6), 1316–1329 (2016).
55. Furukawa, H. Mechanisms of activation, Inhibition and specificity: Crystal structures of the NMDA receptor NR1 ligand-binding core. *EMBO J.* **22**, 2873–2885 (2003).
56. Vonrhein, C. Data processing and analysis with the autoproc toolbox. *ActaChrystallogr. Biol. Crystallogr.* **67**, 293–302 (2011).
57. McCoy, A. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
58. Emsley, P. Features and development of coot. *ActaChrystallogr. Struct. Biol.* **66**, 486–501 (2010).

59. Liebschner, D. Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in phenix. *ActaChrystallogr. Struct. Biol.* **75**, 861–877 (2019).
60. Read, R. J. Improved fourier coefficients for maps using phases from partial structures with errors. *ActaChrystallogr. Biol. Crystallogr.* **42** (3), 140–149 (1986).

Acknowledgements
We thank the members of Integrated Structural Biology Core and Voronina lab for helpful discussions. We thank the New England Biolabs production team Bryce Causey, Andrew Schumaker, and Aubrey Lewis for their protein production insights and suggestions on experimental design. We thank Thomas G.W. Graham, John J. Ferrie, and Michael Layeux for critical reading of this manuscript, and advice with experimental design. We thank William and Susan Studier for their support and reading of this manuscript. We thank Travis Hughes and Michelle Nemetchek for suggestions related to experimental designs, critical reading of this manuscript, and generous access to lab equipment. Lab equipment provided for plate reader experiments was supported by P20GM103546 (T. Hughes). Flow cytometry was performed in the University of Montana Flow Cytometry Core supported by the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (NIH) under Award Number P30GM103338, the M.J. Murdoch Charitable Trust and the University of Montana. Confocal microscopy was performed in the University of Montana BioSpectroscopy Core Research Laboratory operating with support from NIH awards S10OD021806 and P30GM140963. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the NIH/NIGMS (P30GM133894). This work was supported by the NIH grants R01GM109053 (E. Voronina), R01NS097536 and R01NS116055 (K. Hansen) and the National Science Foundation grant CHE-1904895 (B. Bowler). The contents of the manuscript are solely the responsibility of the authors and do not necessarily represent the official views of the funders.

Author contributions
J.B. contributed to all figures, participated in study design, manuscript drafting and editing. E.G., K.B., J.S., T.W., D.C., A.S., C.N., B.L., A.D., C.L.P., C.Y.-H., M.S., M.E.F., and K.C.K. contributed materials, experimental data, and figure preparation. M.E.F., K.B.H., S.S., and L.McC. contributed to the analysis of results. B.C., B.J.P., S.J.C., K.B.H., S.S., B.B., L.McC., M.B., and E.V. contributed to study design and supervised the work. E.V. contributed to manuscript drafting and editing. All authors reviewed and approved the manuscript.

Declarations

Competing interests
J.B. is the inventor for provisional patent application #63546333 filed on 10/30/2023 held by the University of Montana that covers the use of Bosco Broth in recombinant protein overexpression. All the remaining authors declare no conflict of interest.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-91954-5.

Correspondence and requests for materials should be addressed to J.B. or E.V.