BMC
Biotechnology

RESEARCH ARTICLE   **Open Access**

Codon optimization and factorial screening for enhanced soluble expression of human ciliary neurotrophic factor in *Escherichia coli*

Jaakko M Itkonen[1], Arto Urtti[1,2*], Louise E Bird[3,4] and Sanjay Sarkhel[1*]

Abstract

Background: Neurotrophic factors influence survival, differentiation, proliferation and death of neuronal cells within the central nervous system. Human ciliary neurotrophic factor (hCNTF) has neuroprotective properties and is also known to influence energy balance. Consequently, hCNTF has potential therapeutic applications in neurodegenerative, obesity and diabetes related disorders. Clinical and biological applications of hCNTF necessitate a recombinant expression system to produce large amounts of functional protein in soluble form. Earlier attempts to express hCNTF in Escherichia coli (*E. coli*) were limited by low amounts and the need to refold from inclusion bodies.

Results: In this report, we describe a strategy to effectively identify constructs and conditions for soluble expression of hCNTF in *E. coli*. Small-scale expression screening with soluble fusion tags identified many conditions that yielded soluble expression. Codon optimized 6-His-hCNTF construct showed soluble expression in all the conditions tested. Large-scale culture of the 6-His-hCNTF construct yielded high (10 – 20 fold) soluble expression (8 – 9 fold) as compared to earlier published reports. Functional activity of recombinant 6-His-hCNTF produced was confirmed by its binding to hCNTF receptor (hCNTFRα) with an $EC_{50} = 36$ nM.

Conclusion: Our results highlight the combination of codon optimization and screening soluble fusion tags as a successful strategy for high yielding soluble expression of hCNTF in *E. coli*. Codon optimization of the hCNTF sequence seems to be sufficient for soluble expression of hCNTF. The combined approach of codon optimization and soluble fusion tag screen can be an effective strategy for soluble expression of pharmaceutical proteins in *E. coli*.

Keywords: Neurotrophic factors, Human CNTF, Codon optimization, Recombinant soluble expression, *E. coli*

Background

Neurotrophic factors play central role within the nervous system. They regulate cell survival, proliferation, differentiation, migration, dendrite and axonal growth, synaptic plasticity and interactions of neuronal and glial cells [1-4]. The neuropoietic cytokine, ciliary neurotrophic factor (CNTF), first identified in chick ciliary neurons [5,6], belongs to the interleukin-6 and leukemia inhibitory factor (LIF) family of four helix bundle cytokines [7,8]. Human CNTF (hCNTF), a polypeptide of approximate 23 kDa, is involved in the maintenance of ciliary and motor neurons. hCNTF exerts its signaling effect by binding to a heterotrimeric complex of hCNTF receptor (hCNTFRα), gp130 and LIF receptor (LIFRβ) [9-11]. Apart from neuro-protection [12,13], hCNTF is also known to influence energy balance [14-16]. Consequently, hCNTF has potential therapeutic applications not only in neurodegenerative diseases such as Amytropic Lateral Sclerosis (ALS) and Huntington's disease (HD), but also in obesity and related type II diabetes. hCNTF secreting implant has completed phase II studies as an effective therapy for dry age related macular degeneration (AMD) in the eye [17].

The therapeutic potential of CNTF necessitates a recombinant production system for large amounts of protein for clinical and biological applications. Bacterial expression systems are known to be rapid and economical for obtaining recombinant proteins [18-22]. However, attempts to express hCNTF in *Escherichia coli* (*E. coli*) have either yielded low soluble levels (approx. 13%) or insoluble

* Correspondence: arto.urtti@helsinki.fi; sanjay.sarkhel@helsinki.fi
[1]Centre for Drug Research, Division of Pharmaceutical Biosciences, Faculty of Pharmacy, University of Helsinki, P.O. Box 56, Viikinkaari 5E, Helsinki 00014, Finland
[2]School of Pharmacy, University of Eastern Finland, Kuopio, Finland
Full list of author information is available at the end of the article

inclusion bodies [8,23-26][a]. Soluble recombinant expression of pharmaceutical proteins is of utmost importance not only for direct applications but to be able to employ advanced molecular engineering methods to alter their physicochemical properties. Bacterial expression systems offer several advantages; they are relatively inexpensive to culture, can be easily modified genetically and have fast turnaround time. *E. coli*, a preferred bacterial host, is considered to be a 'laboratory workhorse' for recombinant protein expression. However, the advantages are easily offset by the challenges poised by recombinant production of eukaryotic proteins [27,28]. Bacterial hosts such as *E. coli* are often limited in amounts of tRNA for the codons that are used rather less frequently. Instances where codon bias of the gene is significantly different from that of the expression host *E. coli*, paucity of the tRNAs severely limits the translation process. This results in non-optimal translation of the RNA including stalling, termination, possible frame-shifting and low levels of protein expression [29,30]. Apart from codon usage being a determining factor in protein expression levels, structural features of mRNA that interfere with the initiation of translation is also believed to adversely affect protein expression [31,32]. In contrast, a recent study found that there was very less influence of mRNA structure on protein expression and that depletion of charged tRNAs was the main limiting factor for protein expression [33]. Codon optimization of the target gene and/or use of tRNA enhanced strains have become an attractive starting point for heterologous protein expression in *E. coli*. Codon optimization has been successfully utilized to express human pigment epithelium derived factor in *E. coli* [34] and thirty human short chain dehydrogenase/reductive genes showed vast improvement in expression with optimized codon and use of *E. coli* strains containing rare codon tRNAs [35]. The other relevant issue with heterologous expression in *E. coli* is the formation of unordered aggregates (inclusion bodies) due to improper folding of the polypeptide chain. High rates of protein expression or unfavorable conditions for protein folding cause inclusion bodies formation. Successful strategies to mitigate improper folding include decreasing the cultivation temperature [36], co-expression of folding modulators [37] and reducing the rate of gene expression [38]. An alternative strategy to improve the solubility has been to express the protein of interest fused to a solubility enhancing polypeptide 'tag' [39-42]. The soluble 'tag' also serves to provide a reliable translation initiation [43]. Although useful, the fusion 'tag' strategy is limited by the lack of a rational design for choice and a given 'tag' might be effective with only certain targets [40,44-46]. In general, production of soluble proteins in *E. coli* remains largely a trial-and-error process.

In this study, we report an effective strategy towards screening and optimization for soluble expression of hCNTF in *E. coli*. We address the twin issues of codon bias and possible improper folding by optimizing the hCNTF gene sequence for use in *E. coli* and adopting a factorial screening strategy. Small-scale screening methods have been successful in structural genomics initiatives [47-49] in providing reproducible qualitative and quantitative leads for large-scale expression. hCNTF gene was codon optimized and cloned into a suite of nine vectors harboring different fusion (soluble/affinity) tags. The constructs were screened for expression in two different *E. coli* strains and culture medias at different temperatures.

Results and discussion

The objective of our work was to achieve soluble, high yielding production of hCNTF in *E. coli*. Earlier attempts of hCNTF production in *E. coli* resulted in low amounts of soluble protein or required purification from insoluble inclusion bodies [8,23-26]. McDonald and co-workers have reported the presence of only 13% of soluble hCNTF in *E. coli* (BL21) cell extracts [23]. In another report, the insoluble fraction from bacterial cell lysate was found to contain 80% of recombinant hCNTF [24]. Masiakowski et al. have also reported purifying hCNTF from inclusion bodies even though the translation of recombinant hCNTF was 20 – 40% of total protein [25]. Purification of hCNTF from inclusion bodies requires an additional refolding step and such processes are usually not desirable since they are cumbersome and in many cases might result in non-native state of the target protein. Non-native state of proteins often leads to aggregation that can cause fatal immunogenic reactions from protein therapeutics [50]. Thus, for proteins of pharmaceutical relevance, such as hCNTF, it is highly desirable to have high yielding, soluble and cost effective production of protein. An effective strategy to express soluble proteins requires cost effective screening system where several factors could be tested in parallel. Structural genomics initiatives have effectively utilized such approaches to produce recombinant proteins in *E. coli* [47-49].

In our present work, we adopted the strategy of fusing the codon optimized sequence of hCNTF to nine tags (eight soluble tags and the 6–His tag) and carrying out a factorial screening for protein expression in different conditions (temperature/culture media/expression strains). The overall scheme is depicted in Figure 1. The codon optimized hCNTF sequence is shown in Figure 2. A proprietary algorithm (OptimumGene™, Genscript USA) from the vendor was used to increase the codon usage bias in *E. coli* by improving the codon adaptation index (CAI) to 0.87. GC content and unfavorable peaks were optimized to prolong the half-life of the mRNA. Stem-loop structures that have an impact on ribosomal binding and stability of mRNA, were disrupted. In addition, the codon optimization process screened and successfully modified
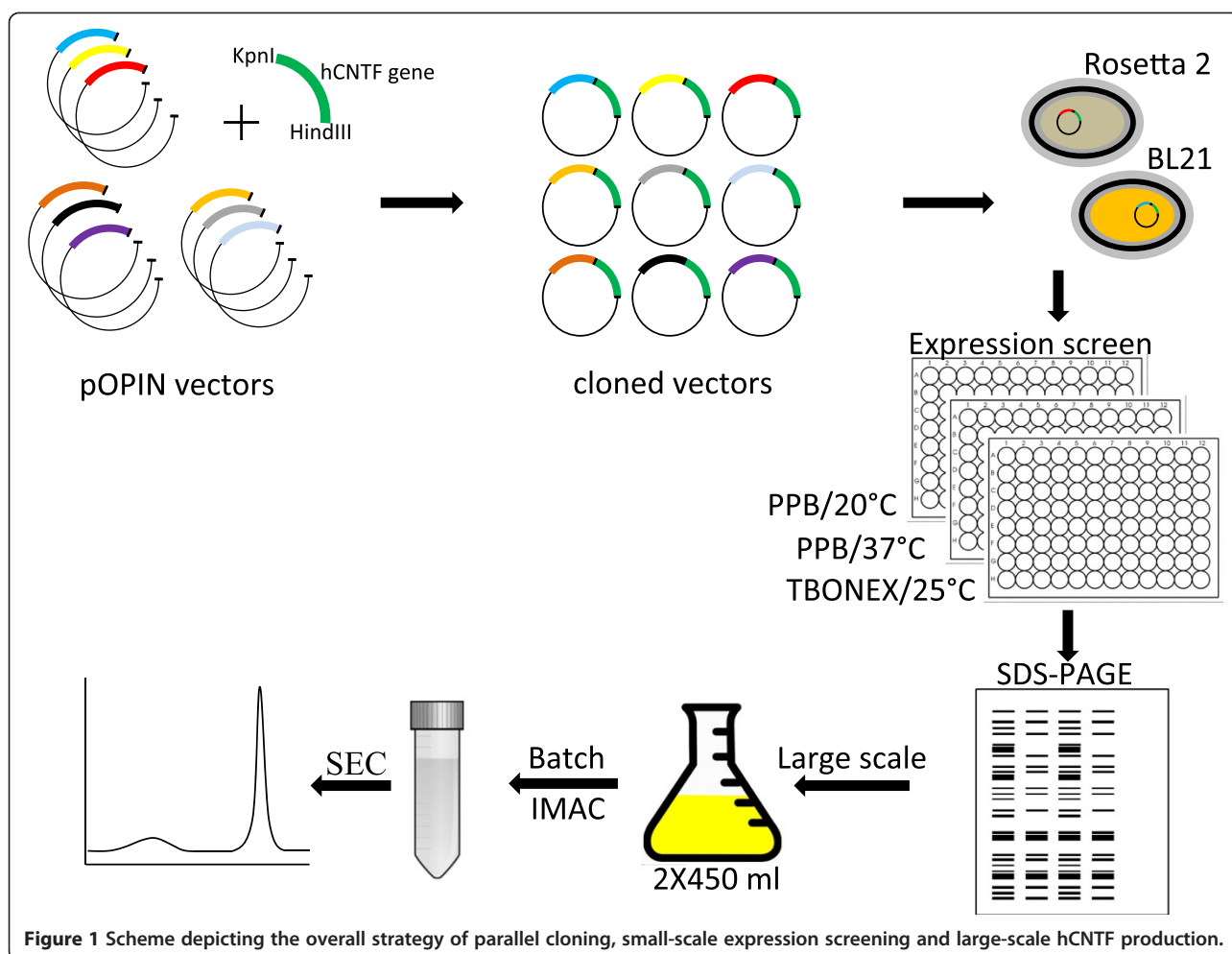
Figure 1 Scheme depicting the overall strategy of parallel cloning, small-scale expression screening and large-scale hCNTF production.

negative *cis*-acting sites. The construction of the pOPIN family of vectors have been described elsewhere [51-53]. Each vector has a N-terminal fusion 'tag' (Table 1) followed by Rhinovirus 3C protease cleavage site between the 'tag' and the inserted target sequence to facilitate 'tag' removal. The set of vectors have common cloning sites (KpnI and HindIII) that facilitate parallel cloning and construction of expression vectors. A cloning efficiency of 100% was achieved by picking three clones for each vector construction. The constructs were transformed into *E. coli* strains, BL21(DE3)pLysS and Rosetta 2(DE3)pLysS (a derivative of BL21 supplying tRNAs for 7 rare codons; AGA, AGG, AUA, CUA, GGA, CCC, and CGG) and small-scale expressions were carried out with culture medias Power Prime Broth (PPB) at 20℃/37℃ and Overnight Expression Terrifc Broth (TBONEX) at 25℃.

The soluble fusion tags used in expression screening are listed in Table 1. The choice of soluble tags has been dictated by literature precedence on the success rate of soluble expression of a diverse set of proteins. No single soluble tag can increase the expression and solubility of all target proteins. However, some general trends on

effectiveness of a limited set of tags (amongst those used in present study) have been reported in some studies. SUMO and NusA were found to be equally efficient in terms of total and soluble protein expression in studies on three model proteins [41,45]. The success of a particular soluble fusion tag is dependent on the target protein of interest and hence the need to screen a diverse set. Both the culture medias used, PPB and TBONEX, are rich and support high cell densities. TBONEX, being an auto-induction media, also offers the advantage of a simplified expression protocol by eliminating the monitoring and induction steps. Induction free protocol is particularly useful for high-throughput protein expression approaches. Culture growth in PPB was screened at temperatures 20℃ and the normal 37℃. Lower culture temperature is known to facilitate proper folding and soluble protein expression [36].

The soluble fractions were treated with Ni-NTA magnetic beads to capture recombinant 6-His-(fusion tag)-hCNTF proteins. The eluted soluble, Ni-NTA affinity purified fractions were analyzed on a SDS-PAGE gel (Figure 3).

```
                   10                    30                              60
wt    ATGGCGTTTACCGAACATAGCCCGCTGACCCCGCATCGCCGCGATCTGTGCAGCCGCAGC
co    ATGGCGTTTACCGAACATTCCCCGCTGACCCCGCATCGTCGTGACCTGTGTAGTCGCTCA
aa     M  A  F  T  E  H  S  P  L  T  P  H  R  R  D  L  C  S  R  S

                   70                    90                            120
wt    ATTTGGCTGGCGCGCAAAATTCGCAGCGATCTGACCGCGCTGACCGAAAGCTATGTGAAA
co    ATCTGGCTGGCTCGCAAAATCCGCTCCGATCTGACCGCCCTGACGGAATCATATGTTAAA
aa     I  W  L  A  R  K  I  R  S  D  L  T  A  L  T  E  S  Y  V  K

                  130                   150                            180
wt    CATCAGGGCCTGAACAAAAACATTAACCTGGATAGCGCGGATGGCATGCCGGTGGCGAGC
co    CATCAGGGCCTGAACAAAAATATTAACCTGGATTCGGCAGACGGTATGCCGGTCGCTAGC
aa     H  Q  G  L  N  K  N  I  N  L  D  S  A  D  G  M  P  V  A  S

                  190                   210                            240
wt    ACCGATCAGTGGGAGCGAACTGACCGAACGCGGAACGCCTGCAGGAAAAACCTGCAGGCGTAT
co    ACCGACCAGTGGTCTGAACTGACGGAAGCGGAACGTCTGCAGGAAAAATCTGCAAGCCTAC
aa     T  D  Q  W  S  E  L  T  E  A  E  R  L  Q  E  N  L  Q  A  Y

                  250                   270                            300
wt    CGCACCTTTCATGTGCTGCTGGCGCGCCTGCTGGAAGATCAGCAGGTGCATTTTACCCCG
co    CGTACCTTTCATGTCCTGCTGGCACGCCTGCTGGAAGATCAGCAAGTGCACTTTACCCCG
aa     R  T  F  H  V  L  L  A  R  L  L  E  D  Q  Q  V  H  F  T  P

                  310                   330                            360
wt    ACCGAAGGCGATTTTCATCAGGCGATTCATACCCTGCTGCTGCAGGTGGCGGCGTTTGCG
co    ACGGAAGGCGACTTCCATCAGGCTATCCACACGCTGCTGCTGCAGGTGGCAGCATTCGCG
aa     T  E  G  D  F  H  Q  A  I  H  T  L  L  L  Q  V  A  A  F  A

                  370                   390                            420
wt    TATCAGATTGAAGAACTGATGATTCTGCTGGAATATAAAATTCCGCGCAACGAAGCGGAT
co    TATCAAATTGAAGAACTGATGATCCTGCTGGAATACAAAATTCCGCGCAATGAAGCGGAT
aa     Y  Q  I  E  E  L  M  I  L  L  E  Y  K  I  P  R  N  E  A  D

                  430                   450                            480
wt    GGCATGCCGATTAACGTGGGCGATGGCGGCCTGTTTGAAAAAAAAACTGTGGGGCCTGAAA
co    GGCATGCCGATCAACGTTGGTGACGGCGGTCTGTTTGAGAAAAAACTGTGGGGTCTGAAA
aa     G  M  P  I  N  V  G  D  G  G  L  F  E  K  K  L  W  G  L  K

                  490                   510                            540
wt    GTGCTGCAGGAACTGAGCCAGTGGACCGTGCGCAGCATTCATGATCTGCGCTTTATTAGC
co    GTGCTGCAGGAACTGAGTCAATGGACCGTTCGTTCCATTCACGATCTGCGCTTCATCTCT
aa     V  L  Q  E  L  S  Q  W  T  V  R  S  I  H  D  L  R  F  I  S

                  550                   570                            600
wt    AGCCATCAGACCGGCATTCCGGCGCGCGGGCAGCCATTATATTGCGAACAACAAAAAAATG
co    TCACATCAGACGGGTATTCCGGCTCGTGGCTCGCACTACATCGCAAACAACAAGAAGATG
aa     S  H  Q  T  G  I  P  A  R  G  S  H  Y  I  A  N  N  K  K  M
```

Figure 2 Comparison of wild type (wt) and codon optimized (co) hCNTF sequence.

Table 1 *Expression vectors and their fusion tags*

| Lane	Vector	Fusion tag	BL21			Rosetta 2		
			20°C PPB	37°C PPB	25°C TBONEX	20°C PPB	37°C PPB	25°C TBONEX
1	pOPIN**F**	6-His-	1	1	1	1	1	1
2	pOPIN**S3C**	6-His-SUMO-	0.7	1.5	2.1	1.3	2.7	1.1
3	pOPIN**TRX**	6-His-Thioredoxin-	1.2	1.7	4.9	1.1	0.7	0.1
4	pOPIN**MSYB**	6-His-MsyB-	0.5	0.8	2.9	1.8	3.6	2.0
5	pOPIN**J**	6-His-GST-	0.3	1.3	0.4	0.1	0.5	0.1
6	pOPIN**HALO7**	6-His-Halo-	1.5	1.6	1.7	1.6	3.0	2.0
7	pOPIN**M**	6-His-MBP-	0.8	1.4	2.5	0.9	1.1	0.7
8	pOPIN**TF**	6-His-Trigger Factor-	0.5	1.1	1.7	0.3	0.6	0.5
9	pOPIN**NusA**	6-His-NusA-	0.6	1.2	1.1	0.7	0.8	0.1

Empirical scores of soluble fraction yield in small-scale expression relative to 6-His-construct. Highlighted text in bold within the vector name symbolize the fusion tag.
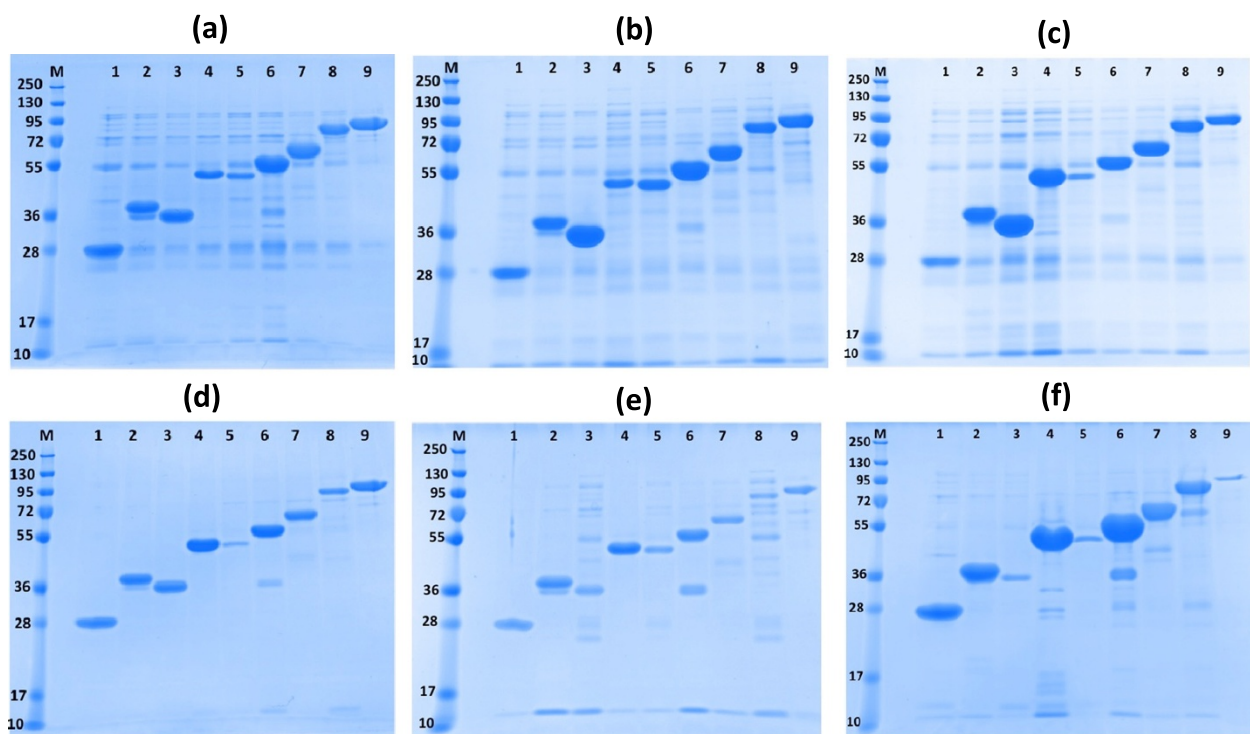
Figure 3 SDS-PAGE analysis of small-scale expression screen. (a) BL21/PPB/20℃; **(b)** BL21/PPB/37℃; **(c)** BL21/TBONEX/25℃; **(d)** Rosetta 2/PPB/20℃; **(e)** Rosetta 2/PPB/37℃; **(f)** Rosetta/TBONEX/25℃. Lanes M, 1–9 denote protein marker and Ni-NTA purified soluble hCNTF fusion constructs from expression vectors as listed in Table 1, respectively. The gel was stained with Coomassie Blue dye.

Most conditions (constructs/temperature/media) screened for expression showed the presence of soluble hCNTF as fusions (Figures 3 and 4). On an average, the percentage of hCNTF in the soluble fraction was close to 80% or more in majority of the conditions screened. Relative amounts of soluble hCNTF as fusions with soluble tags were empirically estimated (comparing intensity of bands) with respect to the 6-His-construct in each experimental condition (Table 1). The intensity of the bands depends on protein concentration and also on the size and content
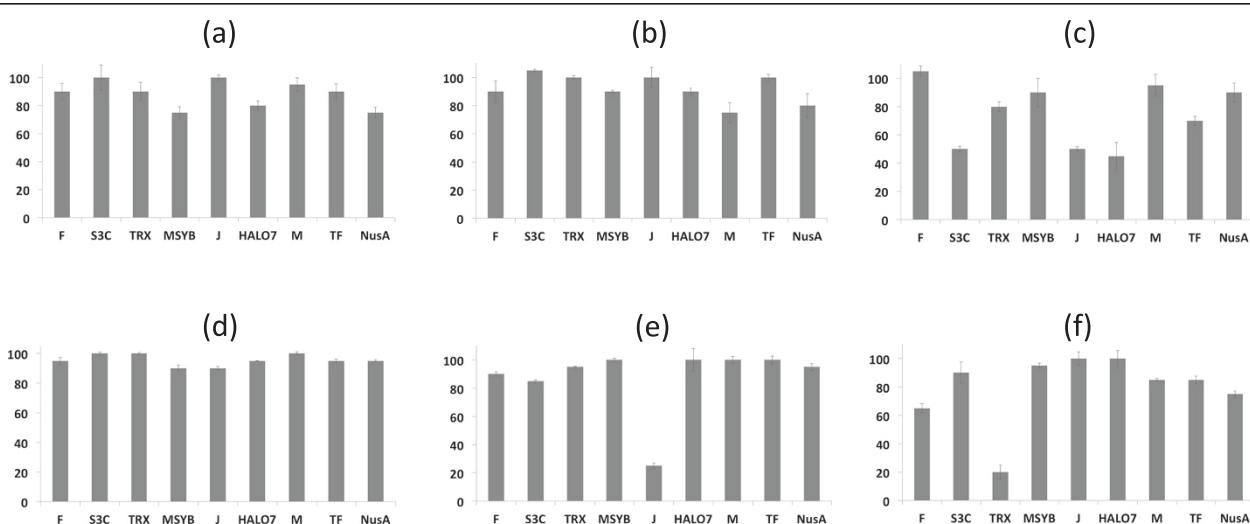
Figure 4 Histograms depicting percentage soluble fraction of hCNTF in small-scale expression screen. (a) BL21/PPB/20℃; **(b)** BL21/PPB/37℃; **(c)** BL21/TBONEX/25℃; **(d)** Rosetta 2/PPB/20℃; **(e)** Rosetta 2/PPB/37℃; **(f)** Rosetta/TBONEX/25℃. The estimates are based on measurements from three technical repeats.

of residue types of the construct. The empirical scores are not normalized against molecular size and residue content and as such the scores should not be used for direct comparisons but for easy visual summarization of relative trends. The efficacies of the fusion tags (scores averaged over the six expression conditions) were ranked as compared to the 6-His- tag. The trend observed was in the order of MsyB/Halo (1.9-fold); SUMO/Trx (1.6-fold); MBP/TF/NusA (approx. 1-fold) and GST (0.5 fold). Significant proteolysis seems to have taken place for some of the larger fusions (Figure 3; Halo, MBP, TF). This could possibly be due to improper folding of the larger fusion constructs. In a recent publication [53], Bird has reported similar observations with the larger fusion constructs. The most significant observation of our small scale screening experiment was that under all conditions, 6-His-hCNTF constructs showed appreciable amounts of expression. It is highly desirable to obtain soluble expression of proteins without a solubility enhancing fusion partner as the downstream purification process can avoid additional steps of tag cleavage and protease removal. Also, some proteins might exhibit markedly low solubility after the fusion tag has been cleaved off; limiting the utility of soluble fusion tags in such cases. The 6-His-tag is merely an affinity tag for purification and in all probability does not improve the solubility of recombinant hCNTF as such, unlike the other fusion tags. In fact, in a study on 20 human proteins, the 6-His-tag had a negative impact on protein solubility when present at either the N- or C-terminus [54]. In this context, our results for soluble expression of 6-His-hCNTF assume significant importance. Assuming that the 6-His-tag does not influence the solubility of recombinant hCNTF, or at least in a positive way, the soluble expression of hCNTF in the conditions tested can be attributed to codon optimization of the hCNTF sequence. Our experimental design was aimed at maximizing the chances of obtaining soluble protein in the backdrop of literature precedence [23-26]. However, we do note that in earlier published reports [23-26] of recombinant expression in *E. coli* with the wild type hCNTF; conditions of expression as varied as in our experimental set up (soluble fusion tags, lower temperatures and expression strain rich in tRNAs for rare codons) were not explored. Results from the earlier studies hint at the fact that most probably both translation (low yields of 5–10 mg/L) and folding (inclusion body formation) were the limiting factors in expression. It might thus be of interest to see the outcome of recombinant expression of the wild type sequence in a strain rich in tRNAs for rare codons, such as Rosetta 2 (DE3) or as a fusion construct with soluble tags in a strain, such as BL21(DE3). Even if the results of such experiments were positive, the codon optimized sequence used in the present study clearly results in a vast improvement since it allows high yielding soluble expression in BL21

at 37°C (approx. 90% or more in soluble fraction) without the aid of soluble fusion tags in contrast to the wild type sequence under similar conditions in previous studies [23-26]. With easy and cheap availability of synthetic genes, the optimized hCNTF codon sequence presents an ideal option for recombinant production in BL21 (a common laboratory strain) without the need for a soluble fusion tag.

Results from small-scale expression show varying efficacies of the fusion constructs in different conditions of growth as determined empirically (Table 1). However, since the overall yield after the cleavage of the fusion tag and purification might be very different from that of the fusion construct and also the need for additional tag cleavage and removal steps; 6-His-hCNTF was chosen for large scale expression with Rosetta 2(pLysS) and the auto inducing media, TBONEX at 25°C as a representative example. The analysis of the overall purification scheme is shown in Figure 5.

The amount of hCNTF protein obtained after final purification was 112 mg/L of culture. This is a vast improvement (greater than 10 – 20 fold) in overall yield from earlier reports of 5 – 10 mg/L [23,24]. Functional activity of recombinant hCNTF was confirmed by binding to hCNTFRα (Figure 6). The binding showed an EC_{50} value of 36 nM in accordance with an earlier published report [55]. A very recent report [56] published during the course of this manuscript preparation, has highlighted an attractive alternate strategy to express high yielding soluble hCNTF in *E. coli*. The study reports the usage of a single protein production (SPP) system that co-expresses *E. coli* toxin mRNA interferase, MazF and hCNTF. The strategy targets down-regulating the expression of endogenous bacterial proteins by selectively cleaving the ACA sites in host mRNAs by MazF. The hCNTF gene sequence was optimized for codon usage in *E. coli* and removing the ACA sites. So, whereas in the recently published study, a codon optimization strategy was directed at more specific use with the expression system, in our present study, a more sophisticated proprietary codon optimization algorithm OptimumGene™ (GenScript, NJ, USA) was used. hCNTF codons in our present study were not only optimized for codon usage in *E. coli* but parameters such as GC content, CpG dinucleotides content, mRNA secondary structure, cryptic splicing sites, premature PolyA sites, internal chi sites and ribosomal binding sites, negative CpG islands, RNA instability motif, repeat sequences were also optimized. The SPP system seems to be a promising approach, however, its robustness as a general strategy needs to be validated for a number of proteins. It appears that the SPP system might be useful in obtaining high yields of recombinant proteins indirectly by suppressing endogenous protein expression. However, protein sequences that might not fold properly once translated might still need soluble

Figure 5 SDS-PAGE analysis of recombinant 6-His-hCNTF purification. M: protein marker; 1: crude cell lysate; 2: soluble fraction; 3: Ni-IDA purified hCNTF; 4: size exclusion chromatography purified hCNTF. The gel was stained with Coomassie Blue dye.
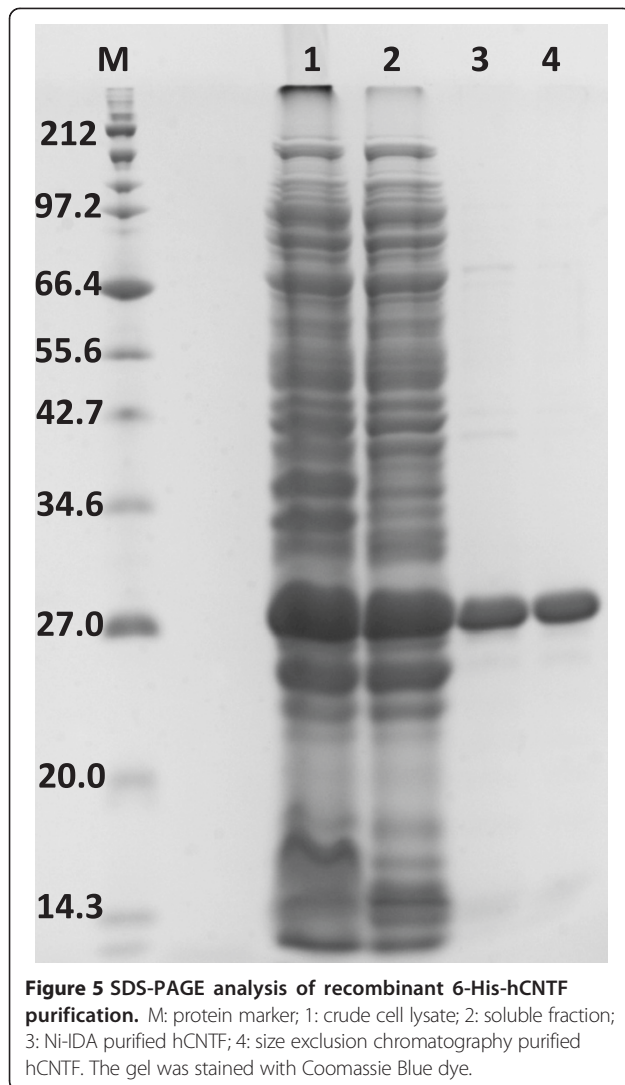
tags for proper folding. From such a perspective, our approach to screen with a suite of soluble tags might have a more versatile utility. It would be interesting to integrate our approach (a sophisticated codon optimization and screening with soluble tags in different strains of *E. coli*) with the SPP system (down-regulating endogenous bacterial proteins) for enhanced expression of difficult eukaryotic targets.

Conclusions

We report here high yielding soluble expression of functional hCNTF in *E. coli*. Our combined approach of codon optimization and screening for soluble expression with nine fusion tags under varying culture conditions identified several constructs producing soluble hCNTF expression (8 – 9 fold higher soluble fraction) and the overall yield of purified recombinant hCNTF achieved was significantly higher (greater than 10 – 20 fold) as compared to previous published reports. Soluble expression of the 6-His-hCNTF in all conditions tested points to the fact that codon optimization is sufficient for soluble expression of hCNTF in *E. coli*. However, for maximal yields, some fusion constructs might be more useful than 6-His-hCNTF. The strategy of combining codon optimization and fusion tag screening might be useful for soluble expression of pharmaceutical proteins in *E. coli*.

Methods

Codon optimization

The codons of the hCNTF gene were optimized for optimal usage in *E. coli* using a commercial proprietary algorithm, OptimumGene™ (GenScript, NJ, USA). Variety of parameters were optimized, including codon usage bias, GC content, CpG dinucleotides content, mRNA secondary structure, cryptic splicing sites, premature PolyA sites, internal chi sites and ribosomal binding sites, negative

Figure 6 Binding of biotinylated hCNTF to hCNTFRα.
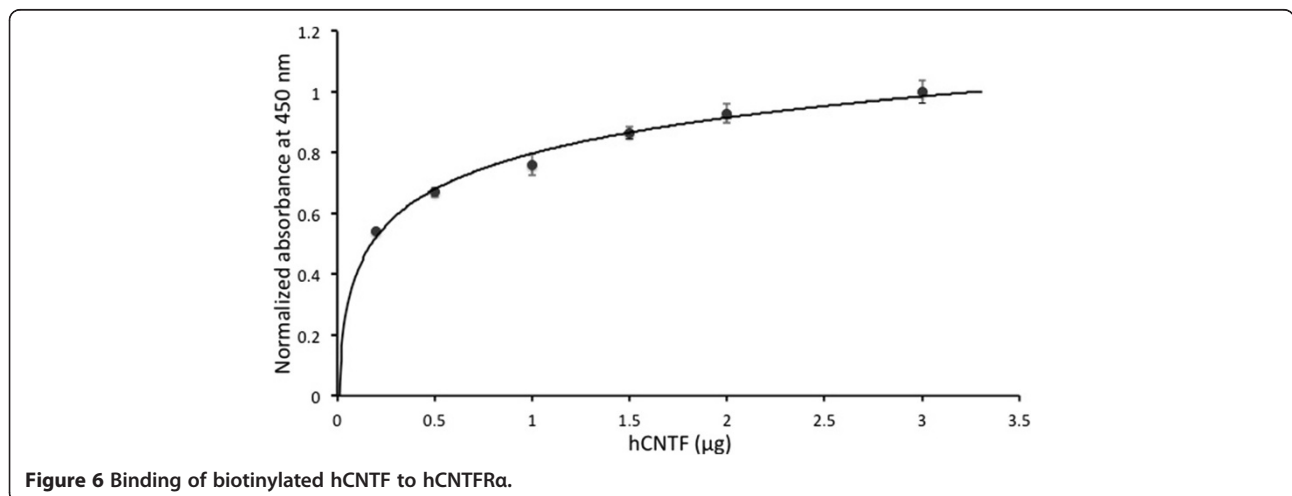
CpG islands, RNA instability motif, repeat sequences. The synthetic gene was purchased from GenScript, US.

Construction of expression vectors

The construction of pOPIN-fusion vector backbones has been described elsewhere [51-53]. The complete coding sequence of hCNTF was synthesized by GenScript, NJ, USA and provided as an insert within pUC57 plasmid. hCNTF gene was amplified by polymerase chain reaction (PCR) using KOD Hot Start (Novagen) with primers that had 5'-extensions of AAGTTCTGTTTCAGGGCCCG and ATGGTCTAGAAAGCTTTA on the forward and reverse primers, respectively (Additional file 1: Figure S1). The set of nine vectors harboring various fusion tags is listed in Table 1. The purified PCR product was inserted into digested pOPIN vectors between KpnI and HindIII restriction sites using Gibson Assembly™ cloning kit (New England Biolabs, MA, USA) and the reaction mixture transformed into NEB 5α (New England Biolabs, MA, USA) competent cells. Three colonies were screened for each construct and the presence of insert verified by PCR using a vector specific forward primer and hCNTF gene specific reverse primer (Additional file 1: Figure S2). pOPIN expression vectors 6-His-(fusion tag)-hCNTFs obtained were (6-His)-hCNTF, 6-His-(S3C)-hCNTF, 6-His-(Trx)-hCNTF, 6-His-(MSYB)-hCNTF, 6-His-(J)-hCNTF, 6-His-(HALO)-hCNTF, 6-His-(M)-hCNTF, 6-His-(TF)-hCNTF, 6-His-(NusA)-hCNTF.

Small-scale expression, affinity purification for soluble protein characterization

Expression strains BL21 (DE3)pLysS (Novagen) and Rosetta 2(DE3)pLysS (Novagen) were transformed with expression plasmids. Single colonies were used to inoculate 0.7 ml Power Prime Broth, PPB (AthenaEs) supplemented with ampicillin (100 µg/ml) and chloramphenicol (34 µg/ml) taken in a 2 ml 96-well plate. The plates were sealed with gas permeable adhesive seals and the seed cultures grown overnight at 37°C with shaking at 250 rpm. For the small-scale expression screen, 150 µl of BL21 (DE3)pLysS and 250 µl of Rosetta 2(DE3)pLysS seed cultures were used to inoculate 3 ml of PPB and Overnight Express™ Instant TB media, TBONEX (Novagen) taken in a 10 ml 24-deepwell plate and supplemented with ampicillin (100 µg/ml) and chloramphenicol (34 µg/ml). Cultures in PPB were grown at 37°C with shaking at 250 rpm. For expression at 37°C, 1 mM IPTG was added when the OD_{600} reached between 0.5 - 0.6. For expression at 20°C, the cultures were cooled for 15 minutes on a bench top before adding 1 mM IPTG and grown at 20°C overnight (18 h) with shaking at 250 rpm. TBONEX auto-inducing cultures were grown at 25°C for 24 h with shaking at 225 rpm. The cultures were harvested by centrifugation and the supernatant removed from the pellets. The pellets

were stored at −80°C. The frozen pellets were resuspended in 210 µL of 50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0; 10 mM imidazole, 1 mM dithiothreitol (DTT), protease inhibitor cocktail, PIC (Sigma). The pellets were freeze-thawed three times; resuspending each time in between the freeze thaw cycle. 1 µL of Lysonase (EMD Millipore) was added to each pellet suspension and incubated at room temperature with shaking at 270 rpm for 45 minutes. The lysates were then centrifuged at 21,100 g at 4°C for 40 minutes. 20 µL of suspended nickel nitriloacetic (Ni-NTA) magnetic beads (GenScript, NJ, USA) was dispensed into each well of a 96-well flat bottom plate. The beads were conditioned with 50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0; 10 mM imidazole, 1% Tween 20. Supernatants from cleared lysates were transferred to the wells containing the magnetic beads and incubated at room temperature for 50 minutes with shaking at 270 rpm. The plate was then placed on a magnetic rack (New England Biolabs, MA, USA) for 5 minutes and the supernatant was removed. The magnetic beads were washed four times with 50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0; 20 mM imidazole, 0.05% Tween 20 and 1 mM DTT. Proteins bound to the magnetic beads were eluted by adding 50 µL of 50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0; 250 mM imidazole, 0.05% Tween 20 and 1 mM DTT and incubating with shaking (270 rpm) at room temperature for 10 minutes. The plate was placed on the magnetic rack for 7 minutes and 45 µL of the supernatant was removed. Protein samples were heated at 95°C with SDS-PAGE sample buffer and analyzed on a SDS-PAGE gel. The 6-His-fusion constructs of hCNTF were verified by specific detection of the 6-His tag using Pierce 6XHis protein tag stain reagent set (Thermo Fisher Scientific Inc., USA).

Estimation of soluble fraction of total recombinant protein expressed

Cell pellets from 3 ml expression culture were thawed and resuspended in 250 µL of 50 mM NaH_2PO_4, 300 mM NaCl pH 8.0; 0.2% Tween 20. The pellets were freeze-thawed three cycles and 50 µL of the suspended pellets were each treated with 1 µL of Lysonase. The pellets were incubated at room temperature for 45 minutes with shaking at 270 rpm. The mixture was divided into two fractions of 25 µL each. For total protein samples, 25 µL of the cell suspension was diluted with an equal volume of water followed by 10 µL of 6X SDS-PAGE sample buffer and boiled for 15 minutes at 95°C. The samples were centrifuged at high speed for 15 minutes and the supernatant run as the 'total protein' sample on a SDS-PAGE gel. For the soluble fraction, the 25 µL cell suspension was diluted with an equal volume of water and centrifuged at high speed for 15 minutes. The supernatant was removed and mixed with 10 µL of 6X denaturing SDS-PAGE sample buffer and boiled for 15 minutes at 95°C. The samples

were analyzed as 'soluble fraction' on a SDS-PAGE gel. GelQuant.NET software (V 1.7.8) from biochemlabsolutions.com was used to quantitate band intensities to estimate the ratio of soluble to total protein. The values were averaged over three measurements. The intensity estimates for comparisons between constructs should be restricted to samples run within the same gel.

Large-scale expression and purification of hCNTF

6-His-hCNTF construct was chosen for large scale expression in Rosetta 2 (DE3)pLysS with the auto-inducing media TBONEX. 2 X 450 ml of cultures were inoculated with overnight seed cultures (30:1) and grown at 25°C with shaking at 270 rpm for 30 h. The cells were harvested by centrifugation at 10,000 g for 30 minutes at 4°C. and frozen at –80°C. The cell pellet was resuspended in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0; 10 mM imidazole, 1 mM DTT and PIC) and freeze thawed three cycles, resuspending thoroughly after thawing each time. Lysonase was added to the suspension and incubated at room temperature for 60 minutes with gentle agitation. The lysate was centrifuged at 14,500 g at 4°C for 50 minutes and the cleared supernatant incubated with Ni-IDA resin (Macherey Nagel) at 4°C for 120 minutes with gentle agitation. The supernatant was removed by centrifugation at 500 g for 3 minutes and the Ni-IDA resin washed with 50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0; 10 mM imidazole, 1 mM DTT and PIC. In subsequent washes, the imidazole concentration was increased. Bound hCNTF was eluted out of the resin (Additional file 1: Figure S3) by adding 250 mM imidazole to the buffer and the eluted fractions pooled, concentrated and buffer exchanged in 100 mM NaH_2PO_4, 300 mM NaCl, pH 8.0; 1 mM DTT using Amicon® Ultra-4 centrifugal filter units with a 3 kDa molecular weight cut off (Merck Millipore). The Ni-IDA eluted protein was further purified by size exclusion chromatography (SEC) by loading onto a HiLoad Superdex 200 prep grade filtration column (Additional file 1: Figure S4). The protein fractions were pooled, concentrated and buffer exchanged in 100 mM NaH_2PO_4, 50 mM NaCl, pH 8.0; 2 mM DTT using centrifugal filter units.

Binding assay of biotinylated hCNTF with hCNTFRα

hCNTF was labeled with biotin for binding studies with hCNTFRα. Briefly, hCNTF was reacted with a 20-fold molar excess of NHS-PEG4-biotin in a 100 mM carbonate buffer solution for 2 h at room temperature with gentle agitation. The reaction mixture was buffer exchanged with 100 mM NaH_2PO_4, 50 mM NaCl, pH 8.0 and concentrated using a spin filter having 3 kDa molecular weight cut off.

For the binding study, 1 μg of hCNTFRα (Sino Biologicals Inc.) in PBS was immobilized on the surface of a Nunc MaxiSorp flat-bottomed 96-well plate. The wells were washed with 50 mM Tris–HCl, 150 mM NaCl, 0.05% Tween 20, pH 7.0 (TBS-T) and blocked with a 2% BSA in TBS-T solution. After washings, biotinylated CNTF was added in varying amounts (0.2 – 3.0 μg) to the wells and incubated at room temperature for 2 h with gentle agitation. The wells were then washed with TBS-T followed by the addition of 50 ng of HRP (horseradish peroxidase)-conjugated streptavidin (Thermo Scientific) to each well and incubated at room temperature for 70 minutes with gentle agitation. The wells were washed and 200 μL of TMB (3,3', 5,5'-tetramethylbenzidine; Thermo Scientific) was added to each well. The reaction was stopped with 2 M H_2SO_4 and the optical density measured at 450 nm using a spectrophotometer (Varioskan Flash; Thermo Scientific).

Availability of supporting data

The data set supporting the results of this article are included within the article and in Additional file 1.

Additional file 1: Figures S1 - S4. Description: PCR amplified hCNTF, validation of hCNTF clones in expression vectors, affinity batch purification fractions, size exclusion chromatography profile.

Endnote

[a]During the course of this manuscript preparation, an alternate expression system expressing soluble hCNTF in *E. coli* was published. This has been discussed in the 'Results and Discussion' section.

Additional file

Additional file 1: Figure S1. PCR-amplified hCNTF insert. **Table S1** PCR primers for the amplification of the synthesized hCNTF. **Figure S2** PCR validation of cloned hCNTF in pOPIN vectors. **Figure S3** Representative eluted fractions from Ni-IDA batch purification of 2 × 450 ml cultures **(A and B)**. **Figure S4** Size exclusion chromatography (SEC).

Abbreviations

hCNTF: Human ciliary neurotrophic factor; E. coli: Escherichia coli; hCNTFRα: Human ciliary neurotrophic factor receptor alpha; LIFβ: LIF receptor; ALS: Amytropic lateral sclerosis; HD: Huntington's disease; AMD: Age related macular degeneration; tRNAs: Transfer RNAs; CAI: Codon adaptation index; TBONEX: Overnight expression terrific broth; PCR: Polymerase chain reaction; 6-His: 6-histidine; SUMO: Small ubiquitin related modifier; Trx: Thioredoxin; MsyB: *E. coli* acidic protein; GST: Glutathione S-transferase; Halo: Mutated dehalogenase; MBP: Maltose binding protein; TF: Trigger factor; NusA: N-utilization substance A; DTT: Dithiothreitol; Ni-NTA: Nickel nitriloacetic; PPB: Power prime broth; IPTG: Isopropyl β-D-1-thiogalactopyranoside; SDS-PAGE: Sodium dodecyl sulphate polyacrylamide gel electrophoresis; PIC: Protease inhibitor cocktail; Ni-IDA: Nickel iminodiacetic acid; SEC: Size exclusion chromatography; PBS: Phosphate buffered saline; NHS-PEG4: N-hydroxysuccinimide ester – polyethylene glycol; TBS-T: Tris buffered saline-tween; HRP: Horseradish peroxidase; TMB: 3,3', 5,5'-tetramethylbenzidine; SPP: Single protein production.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Centre for Drug Research, Division of Pharmaceutical Biosciences, Faculty of Pharmacy, University of Helsinki, P.O. Box 56, Viikinkaari 5E, Helsinki 00014, Finland. [2]School of Pharmacy, University of Eastern Finland, Kuopio, Finland. [3]Division of Structural Biology, Henry Wellcome Building for Genomic Medicine, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK. [4]OPPF-UK, Research Complex at Harwell, R92 Rutherford Appleton Laboratory, Harwell Oxford, Didcot, Oxford OX11 OFA, UK.

References

1. Campenot RB: **Local control of neurite development by nerve growth factor.** *PNAS* 1977, **74**:4516–4519.
2. Chao MV: **Neurotrophins and their receptors: A convergence point for many signalling pathways.** *Nat Rev Neurosci* 2003, **4**:299–309.
3. Vergara C, Ramirez B: **CNTF, a pleiotropic cytokine: emphasis on its myotrophic role.** *Brain Res Brain Res Rev* 2004, **47**:161–173.
4. Weissmiller AM, Wu C: **Current advances in using neurotrophic factors to treat neurodegenerative disorders.** *Transl Neurodegener* 2012, **1**:14.
5. Adler R, Landa KB, Manthorpe M, Varon S: **Cholinergic neuronotrophic factors: intraocular distribution of trophic activity for ciliary neurons.** *Science* 1979, **204**:1434–1436.
6. Barbin G, Manthorpe M, Varon S: **Purification of the chick eye ciliary neuronotrophic factor.** *J Neurochem* 1984, **43**:1468–1478.
7. Bazan JF: **Neuropoietic cytokines in the hematopoietic fold.** *Neuron* 1991, **7**:197–208.
8. McDonald NQ, Panayotatos N, Hendrickson WA: **Crystal structure of dimeric human ciliary neurotrophic factor determined by MAD phasing.** *EMBO J* 1995, **14**:2689–2699.
9. Stahl N, Yancopoulos GD: **The alphas, betas, and kinases of cytokine receptor complexes.** *Cell* 1993, **74**:587–590.
10. Dy M, Vasquez A, Bertoglio J, The'se J: **General aspects of cytokine properties and function.** In *The Cytokine Network and Immune Functions.* Hong Kong: Oxford Univ. Press; 1999:1–13.
11. Sleeman MW, Anderson KD, Lambert PD, Yancopoulos GD, Wiegand SJ: **The ciliary neurotrophic factor and its receptor, CNTFRα.** *Pharm Acta Helv* 2000, **74**:265–272.
12. Emerich DF, Winn SR, Hantraye PM, Peschanski M, Chen EY, Chu Y, McDermott P, Baetge EE, Kordower JH: **Protective effect of encapsulated cells producing neurotrophic factor CNTF in a monkey model of Huntington's disease.** *Nature* 1997, **386**:395–399.
13. Dutta R, McDonough J, Chang A, Swamy L, Siu A, Kidd GJ, Rudick R, Mirnics K, Trapp BD: **Activation of the ciliary neurotrophic factor (CNTF) signalling pathway in cortical neurons of multiple sclerosis patients.** *Brain* 2007, **130**:2566–2576.
14. Lambert PD, Anderson KD, Sleeman MW, Wong V, Tan J, Hijarunguru A, Corcoran TL, Murray JD, Thabet KE, Yancopoulos GD, Wiegand SJ: **Ciliary neurotrophic factor activates leptin-like pathways and reduces body fat,** without cachexia or rebound weight gain, even in leptin-resistant obesity. *Proc Natl Acad Sci U S A* 2001, **98**:4652–4657.
15. Ettinger MP, Littlejohn TW, Schwartz SL, Weiss SR, McIlwain HH, Heymsfield SB, Bray GA, Roberts WG, Heyman ER, Stambler N, Heshka S, Vicary C, Guler H-P: **Recombinant variant of ciliary neurotrophic factor for weight loss in obese adults: a randomized, dose-ranging study.** *JAMA* 2003, **289**:1826–1832.
16. Matthews VB, Febbraio MA: **CNTF: a target therapeutic for obesity-related metabolic disease?** *J Mol Med* 2008, **86**:353–361.
17. Wen R, Tao W, Li Y, Sieving PA: **CNTF and retina.** *Prog Retin Eye Res* 2012, **31**:136–151.
18. Baneyx F: **Recombinant protein expression in Escherichia coli.** *Curr Opin Biotechnol* 1999, **10**:411–421.
19. Balbás P: **Understanding the art of producing protein and nonprotein molecules in Escherichia coli.** *Mol Biotechnol* 2001, **19**:251–267.
20. Peti W, Page R: **Strategies to maximize heterologous protein expression in Escherichia coli with minimal cost.** *Protein Expr Purif* 2007, **51**:1–10.
21. Huang C-J, Lin H, Yang X: **Industrial production of recombinant therapeutics in Escherichia coli & and its recent advancements.** *J Ind Microbiol Biotechnol* 2012, **39**:383–399.
22. Chen R: **Bacterial expression systems for recombinant protein production: E. coli and beyond.** *Biotechnol Adv* 2012, **30**:1102–1107.
23. McDonald JR, Ko C, Mismer D, Smith DJ, Collins F: **Expression and characterization of recombinant human ciliary neurotrophic factor from Escherichia coli.** *Biochim Biophys Acta* 1991, **1090**:70–80.
24. Negro A, Corona G, Bigon E, Martini I, Grandi C, Skaper SD, Callegaro L: **Synthesis, purification, and characterization of human ciliary neuronotrophic factor from E. coli.** *J Neurosci Res* 2004, **29**:251–260.
25. Masiakowski P, Liu H, Radziejewski C, Lottspeich F, Oberthuer W, Wong V, Lindsay RM, Furth ME, Panayotatos N: **Recombinant Human and Rat Ciliary Neurotrophic Factors.** *J Neurochem* 1991, **57**:1003–1012.
26. Wagener E-M, Aurich M, Aparicio-Siegmund S, Floss DM, Garbers C, Breusing K, Rabe B, Schwanbeck R, Grötzinger J, Rose-John S, Scheller J: **The amino acid exchange R28E in ciliary neurotrophic factor (CNTF) abrogates IL-6R- but retains CNTFR-dependent signaling via gp130/LIFR.** *J Biol Chem* 2014, in press.
27. Baneyx F, Mujacic M: **Recombinant protein folding and misfolding in Escherichia coli.** *Nat Biotechnol* 2004, **22**:1399–1408.
28. Khow O, Suntrarachun S: **Strategies for production of active eukaryotic proteins in bacterial expression system.** *Asian Pac J Trop Biomed* 2012, **2**:159–162.
29. Kane JF: **Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli.** *Curr Opin Biotechnol* 1995, **6**:494–500.
30. Calderone TL, Stevens RD, Oas TG: **High-level Misincorporation of Lysine for Arginine at AGA Codons in a Fusion Protein Expressed in Escherichia coli.** *J Mol Biol* 1996, **262**:407–412.
31. Welch M, Villalobos A, Gustafsson C, Minshull J: **You're one in a googol: optimizing genes for protein expression.** *J R Soc Interface* 2009, **6**(Suppl 4):S467–S476.
32. Kudla G, Murray AW, Tollervey D, Plotkin JB: **Coding-Sequence Determinants of Gene Expression in Escherichia coli.** *Science* 2009, **324**:255–258.
33. Welch M, Govindarajan S, Ness JE, Villalobos A, Gurney A, Minshull J, Gustafsson C: **Design parameters to control synthetic gene expression in Escherichia coli.** *PLoS ONE* 2009, **4**:e7002.
34. Gvritishvili AG, Leung KW, Tombran-Tink J: **Codon preference optimization increases heterologous PEDF expression.** *PLoS ONE* 2010, **5**:e15056.
35. Burgess-Brown NA, Sharma S, Sobott F, Loenarz C, Oppermann U, Gileadi O: **Codon optimization can improve expression of human genes in Escherichia coli: A multi-gene study.** *Protein Expr Purif* 2008, **59**:94–102.
36. Vasina JA, Baneyx F: **Expression of aggregation-prone recombinant proteins at low temperatures: a comparative study of the Escherichia coli cspA and tac promoter systems.** *Protein Expr Purif* 1997, **9**:211–218.
37. Wall JG, Plückthun A: **Effects of overexpressing folding modulators on the in vivo folding of heterologous proteins in Escherichia coli.** *Curr Opin Biotechnol* 1995, **6**:507–516.
38. Galloway CA, Sowden MP, Smith HC: **Increasing the yield of soluble recombinant protein expressed in E. coli by induction during late log phase.** *BioTechniques* 2003, **34**:524–526. 528, 530.

39. Davis GD, Elisee C, Newham DM, Harrison RG: **New fusion protein systems designed to give soluble expression in Escherichia coli.** *Biotechnol Bioeng* 1999, **65**:382–388.
40. Kapust RB, Waugh DS: **Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused.** *Protein Sci* 1999, **8**:1668–1674.
41. Butt T, Edavettal S, Hall J, Mattern M: **SUMO fusion technology for difficult-to-express proteins.** *Protein Expr Purif* 2005, **43**:1–9.
42. Esposito D, Chatterjee DK: **Enhancement of soluble protein expression through the use of fusion tags.** *Curr Opin Biotechnol* 2006, **17**:353–358.
43. Waugh DS: **Making the most of affinity tags.** *Trends Biotechnol* 2005, **23**:316–320.
44. Kataeva I, Chang J, Xu H, Luan C-H, Zhou J, Uversky VN, Lin D, Horanyi P, Liu ZJ, Ljungdahl LG, Rose J, Luo M, Wang B-C: **Improving solubility of Shewanella oneidensis MR-1 and Clostridium thermocellum JW-20 proteins expressed into Esherichia coli.** *J Proteome Res* 2005, **4**:1942–1951.
45. Marblestone J, Edavettal S, Lim Y, Lim P, Zuo X, Butt T: **Comparison of SUMO fusion technology with traditional gene fusion systems: Enhanced expression and solubility with SUMO.** *Protein Sci* 2006, **15**:182–189.
46. Bogomolovas J, Simon B, Sattler M, Stier G: **Screening of fusion partners for high yield expression and purification of bioactive viscotoxins.** *Protein Expr Purif* 2009, **64**:16–23.
47. Knaust RKC, Nordlund P: **Screening for Soluble Expression of Recombinant Proteins in a 96-Well Format.** *Anal Biochem* 2001, **297**:79–85.
48. Berrow NS, Büssow K, Coutard B, Diprose J, Ekberg M, Folkers GE, Levy N, Lieu V, Owens RJ, Peleg Y, Pinaglia C, Quevillon-Cheruel S, Salim L, Scheich C, Vincentelli R, Busso D: **Recombinant protein expression and solubility screening in Escherichia coli: a comparative study.** *Acta Crystallogr Sect D: Biol Crystallogr* 2006, **62**:1218–1226.
49. Saez NJ, Vincentelli R: **High-Throughput Expression Screening and Purification of Recombinant Proteins in E. coli.** In *Structural Genomics. Volume 1091*. Edited by Chen YW. Totowa, NJ: Humana Press; 2014:33–53.
50. Ratanji KD, Derrick JP, Dearman RJ, Kimber I: **Immunogenicity of therapeutic proteins: Influence of aggregation.** *J Immunotoxicol* 2013, **11**:99–109.
51. Berrow NS, Alderton D, Sainsbury S, Nettleship J, Assenberg R, Rahman N, Stuart DI, Owens RJ: **A versatile ligation-independent cloning method suitable for high-throughput expression screening applications.** *Nucleic Acids Res* 2007, **35**:e45.
52. Berrow NS, Alderton D, Owens RJ: **The precise engineering of expression vectors using high-throughput In-Fusion PCR cloning.** *Methods Mol Biol* 2009, **498**:75–90.
53. Bird LE: **High throughput construction and small scale expression screening of multi-tag vectors in Escherichia coli.** *Methods* 2011, **55**:29–37.
54. Woestenenk EA, Hammarström M, van den Berg S, Härd T, Berglund H: **His tag effect on solubility of human proteins produced in Escherichia coli: a comparison between four expression vectors.** *J Struct Funct Genomics* 2004, **5**:217–229.
55. Saggio I, Paonessa G, Gloaguen I, Graziani R, De Serio A, Laufer R: **Nonradioactive receptor binding assay for ciliary neurotrophic factor.** *Anal Biochem* 1994, **221**:387–391.
56. Wang K, Zhou F, Zhu L, Zhu X, Zhang K, Zhu L: **High level soluble expression, purification, and characterization of human ciliary neuronotrophic factor in Escherichia coli by single protein production system.** *Protein Expr Purif* 2014, **96**:8–13.