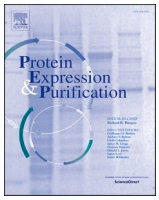
Protein Expression and Purification 222 (2024) 106542

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep



Production of stable and pure ZC3H11A – An extensively disordered RNA binding protein

Mostafa Fekry [a,b], Gun Stenberg [a], Doreen Dobritzsch [a], U. Helena Danielson [a,c,*]

[a] Department of Chemistry – BMC, Uppsala University, Uppsala, Sweden
[b] Biophysics Department, Faculty of Science, Cairo University, Giza, Egypt
[c] Science for Life Laboratory, Drug Discovery & Development Platform, Uppsala University, SE 751 23 Uppsala, Sweden

ARTICLE INFO

Keywords:
ZC3H11A
Zinc finger
RNA binding proteins
Insect cell expression
mRNA export

ABSTRACT

Human ZC3H11A is an RNA-binding zinc finger protein involved in mRNA export and required for the efficient growth of human nuclear replicating viruses. Its biochemical properties are largely unknown so our goal has been to produce the protein in a pure and stable form suitable for its characterization. This has been challenging since the protein is large (810 amino acids) and with only the N-terminal zinc finger domain (amino acids 1–86) being well structured, the remainder is intrinsically disordered. Our production strategies have encompassed recombinant expression of full-length, truncated and mutated ZC3H11A variants with varying purification tags and fusion proteins in several expression systems, with or without co-expression of chaperones and putative interaction partners. A range of purification schemes have been explored. Initially, only truncated ZC3H11A encompassing the zinc finger domain could successfully be produced in a stable form. It required recombinant expression in insect cells since expression in *E. coli* gave a protein that aggregated. To reduce problematic nucleic acid contaminations, Cys8, located in one of the zinc fingers, was substituted by Ala and Ser. Interestingly, this did not affect nucleic acid binding, but the full-length protein was stabilised while the truncated version was insoluble. Ultimately, we discovered that when using alkaline buffers (pH 9) for purification, full-length ZC3H11A expressed in Sf9 insect cells was obtained in a stable and >90 % pure form, and as a mixture of monomers, dimers, tetramers and hexamers. Many of the challenges experienced are consistent with its predicted structure and unusual charge distribution.

1. Introduction

Zinc finger (ZnF) proteins are among the most abundant proteins in eukaryotes, having diverse functions in various processes such as RNA packing, DNA recognition, transcriptional activation, regulation of apoptosis and protein folding. The zinc ions stabilizing their characteristic ZnF domains are coordinated by cysteine (C) and histidine (H) residues, with four of these residues typically being combined to C2H2, C3H or C4 coordination patterns giving rise to the correspondingly named ZnF subclasses. Classical ZnF domains are folded into one α-helix and two β-strands [1–3].

The protein in focus in this study is the Zinc finger CCCH domain–containing protein 11A (ZC3H11A). It consists of 810 amino acids (Mw = 89 kDa) and has three ZnF domains located at the N-terminus (amino acids 2 to 86). The rest of the protein appears to be intrinsically disordered. The protein is extensively post-translationally modified, primarily in the form of phosphorylations [4,5]. ZC3H11A belongs to the C3H-type of ZnF proteins. Many proteins in this subclass are involved in immune homeostasis, inflammation and immune cell maturation [6–8].

For a long time, the function of ZC3H11A remained elusive, but the protein has now been identified as a component of the multiprotein Transcription-Export (TREX) complex [9,10] which plays a key function in nuclear mRNA export [11,12]. More recently it was found that ZC3H11A is a stress-induced RNA binding protein localized to nuclear speckles and that it is indeed involved in mRNA export and required for effective reproduction of different human nuclear replicating viruses, such as HIV, herpes simplex virus, influenza virus, and adenovirus [13].

Abbreviations: IPTG, isopropylthiogalactopyranoside; MS, mass spectrometry; Mw, molecular weight; PCR, polymerase chain reaction; PEI, polyethylenimine; TEV, tobacco etch virus; ZnF, zinc finger; TREX, Transcription-Export.

* Corresponding author.
E-mail address: helena.danielson@kemi.uu.se (U.H. Danielson).

https://doi.org/10.1016/j.pep.2024.106542
Received 21 May 2024; Received in revised form 2 July 2024; Accepted 3 July 2024
Available online 4 July 2024
1046-5928/© 2024 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

These viruses appear to hijack ZC3H11A to promote their replication, as indicated by the drastic decrease in their proliferation observed upon infection of HeLa cells in which ZC3H11A was knocked-out. Furthermore, the proliferation of epidemic porcine viruses, such as pseudorabies virus (PRV) and porcine circovirus 2 (PCV2) has been found to be reduced by the inactivation of porcine ZC3H11A in PK-15 cells. These viruses are among the costliest pathogens affecting the swine industry [14]. A very recent study showed that human ZC3H11A expression is abnormal in glioma infected by human cytomegalovirus (HCMV). It also revealed that elevated ZC3H11A expression is primarily linked to a negative prognosis in most cancers but that there are variations depending on the cancer type [15]. These observations indicate that ZC3H11A may be a viable target for antiviral therapy [13].

Our goal in this study was to produce ZC3H11A in a pure form to enable biochemical studies and explore the possibility of establishing an antiviral discovery programme driven by biophysical and biochemical methods. Since little was known about the physico-chemical and structural characteristics of ZC3H11A at the start of this project, a standard approach for its recombinant production and purification was initially employed. However, it soon became clear that the production of stable and pure full-length ZC3H11A in useful amounts for biochemical studies was very challenging. It appeared to be a consequence of both its large size and its extensive disordered regions. Therefore, a series of recombinant expression trials aiming to produce full-length, truncated and mutated ZC3H11A variants were undertaken. It explored the effectiveness of various purification tags and fusion proteins, as well as co-expression with chaperones and putative interaction partners in several host systems. Here we provide a detailed description of the pursued strategies and their results, as well as basic biochemical characteristics of the ultimately obtained full-length and truncated protein variants.

2. Materials and methods

2.1. Molecular cloning and recombinant expression

The NCBI Reference Sequence NM_014827.5 was used as a template for generating expression constructs of ZC3H11A in mammalian and insect cells, but was codon optimised for expression in *E. coli*. The plasmids used for expression of ZC3H11A and accessory proteins, and their origin, are summarised in Tables S1, S2 and S3. Plasmids for the initial ZC3H11A constructs (pET15b/ZC3_N-his, BacPAK/ZC3_N-2x S-t, pET16b/ZC3D_N-2x S-t and pDSG-IBA105) were purchased from GenScript (Rijswijk, Netherlands) or IBA Lifesciences (Göttingen, Germany). The plasmids for the originally used constructs of accessory proteins (Hpr1-GST, hTex1-GST and SAP24_GST, BacPAK/Thoc5_C-his, BacPAK/UAP56_C-his, and BacPAK/ALYREF_C-his) were purchased from Addgene (Watertown, MA 02472, USA) or Genscript. All other clones were generated in house by subcloning. PCR primers (Eurofins Genomics, Ebersberg, Germany) are listed in Table S4. The bacterial clone pET15b/ZC3_N-his was used as template for PCR to generate a bacterial pKK expression construct as well as a construct of maltose binding protein (MBP) fused to the N-terminus of ZC3H11A.

All PCR, mutagenesis, restriction analysis, ligation, and transformation experiments were performed using standard procedures and reagents from Thermo Fisher Scientific (Waltham, MA, USA). The GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) was used for plasmid minipreps. DNA sequencing was performed by Eurofins Genomics (Ebersberg, Germany).

Expression in *E. coli*. Recombinant expression in *E. coli* was tested in bacterial strains BL21 (DE3) (Thermo Fisher Scientific), XL-1 Blue (Agilent, Santa Clara, CA, USA) or Top10 (Thermo Fisher Scientific) depending on the particular plasmid vector used. Heat-shock was used for transformation of chemically competent cells (produced in house). A small starter culture was prepared by picking individual transformant colonies for overnight growth at 37 °C, 220 rpm, in LB medium (10 g tryptone, 5 g yeast extract, 10 g NaCl per litre) supplemented with the antibiotics suiting the respective expression plasmids (Sigma-Aldrich, Darmstadt, Germany). The next day, the culture was added to larger medium volumes, diluting it 100-fold. The culture was grown at 37 °C, 220 rpm until an OD_{600} of approximately 0.6 was reached. Cultures were cooled and isopropyl-β-D-thiogalactopyranoside (IPTG) (Sigma-Aldrich) was added to 1 mM final concentration. Growth was then continued overnight at + 18 °C and cells were harvested by centrifugation. This protocol was used with smaller variations in expression times, induction start, concentration of IPTG, growth media and temperatures, but with little effect on the expression outcome.

Expression in *E. coli* was also attempted in strain XL-1 Blue using vector pKK223 (Pharmacia, Uppsala, Sweden) containing a tac promoter, or upon fusion truncated ZC3 to MBP in vector pMAL-c2X (New England Biolabs, Ipswich, MA, USA).

Expression in *Sf*9 insect cells. For protein production in *Sf*9 insect cells, the flashBAC ULTRA baculovirus system (Oxford Expression Technologies, Oxford, UK) was used. BacPAK8 constructs containing the gene of interest was used for transfection of *Sf*9 cells and generation of viral stocks. The instructions accompanying the kit was followed. A culture of 2×10^6 cells/mL in Insect-Xpress (Lonza, Basel, Switzerland) or Sf-900II SFM (Gibco, Thermo Fisher Scientific) was inoculated with 10 mL of virus stock/L of medium. Growth was at 28 °C, 90 rpm, for 30–72 h. For co-expressions, the virus stock additions were reduced to 8 mL of each. After expression, cells were harvested by centrifugation.

Mammalian cells. Full-length ZC311A with an N-terminal Twin-Strep-tag (GenScript) was produced in a mammalian cell system by transient expression of the ZC311A gene in a pDSG-IBA105 vector in Mexi-293E cells (IBA Lifesciences). Transfection was done using MEXi Culture Medium (MEXi-CM) and MEXi Transfection Medium (MEXi-TM). The cells were subsequently grown in MEXi Culture Medium (MEXi-CM) by shaking at 125 rpm at 37 °C and 5 % CO_2. Cells were harvested to be seeded in MEXi-TM (pre-warmed at 37 °C) at 1.5×10^6 cells/mL. Equilibration of pH in the medium was then done by shaking at 125 rpm, 37 °C and 5 % CO_2 for 1 h. The ZC3H11A DNA plasmid was added to the transfection medium at 1.5 mg/L followed by addition of linear polyethylenimine (PEI) at 5 mg/mL. The flask was then shaked for 4 h at 125 rpm, 37 °C and 5 % CO_2. After 4 h, the transfected cell culture was diluted to 1:2 with MEXi-CM (resulting in a cell density of 0.75×10^6 cells/mL) and culturing was continued. Cells were harvested by centrifugation after 120 h when the cell viability had come down to *ca.* 75 %.

For all above production methods, the isolated cell pellets were either used directly for protein purification or flash frozen in N_2 (l) for storage at −80 °C, awaiting purification.

2.2. Protein purification and handling

Strep-Tactin®XT affinity column chromatography was used to purify proteins produced with a Twin-Strep-tag (IBA Lifesciences). Initially, the protocol provided by IBA was followed. Cells were resuspended by shaking in ca 30 mL of Lysis buffer A (100 mM Tris-HCl pH 7.4 or pH 8, 150 mM NaCl) supplemented with RNAse, DNAse and cOmplete protease inhibitor)/10 g of cells and subsequently lysed by sonication or by a cell disruptor. Lysed cells were centrifuged for 45 min @20 000 rpm in a Beckman JA25.5 rotor.

To minimise the unspecific binding of biotinylated protein to Strep-Tactin®XT beads, the lysate was supplemented with Bio-Lock (IBA Lifesciences) and placed on ice for a minimum of 20 min. Strep-Tactin®XT resin (IBA Lifesciences) was equilibrated in Lysis buffer A and poured into a disposable column. The lysate was passed through the column at least twice at gravity flow. The column was subsequently washed with Lysis buffer A until the flow-through had an $Abs_{280} \approx 0$. Bound protein was eluted with Lysis buffer A containing 50 mM biotin. Fractions were collected and the absorbance at 280 nm determined. Fractions were further analysed by SDS-PAGE.

Purification of ZC3H11A$_{1\text{-}100}$ under denaturing conditions was done by thawing the *E. coli* pellet, and resuspending it in 5 mL denaturing lysis buffer (100 mM Tris-HCl pH 7.4, 150 mM NaCl, 6 M urea) per gram of pellet. The slurry was stirred for 60 min at room temperature, followed by centrifugation at 10 000×*g* for 30 min at room temperature. The supernatant was then purified by chromatography using a Strep-Tactin®XT resin. Elution was done using the denaturing lysis buffer supplemented with 50 mM biotin.

The buffer conditions for purification of ZC3H11A$_{1\text{-}100}$ and full-length ZC3H11A expressed in *Sf9* cells were optimised in order to improve protein purity and stability. Ultimately, the selected lysis buffer had a higher pH and different concentrations of other components (20 mM Tris-HCl pH 9, 200 mM NaCl). Nucleic acid was removed by adding 100 μL of 5 % PEI solution to 1 g of cells (Merck, Darmstadt, Germany). To remove the Twin-Strep-tag, the protein was bound to the Strep-Tactin®XT resin and Tobacco Etch Virus (TEV) protease was added (1 mg/mL). Cleavage was done overnight at + 4 °C and the flow-through containing tag-free ZC3H11A was collected by gravity flow the next day.

Immobilised metal affinity chromatography (IMAC) was usually done in a batch mode. *Ca.* 30 mL of Lysis buffer B (20 mM Tris-HCl, 150 mM NaCl, 10 mM imidazole) supplemented with RNAse, DNAse and cOmplete protease inhibitor)/10 g of cells was added to the cells and shaken until cells were suspended. Lysis and centrifugation were as above. Equilibrated IMAC resin (Cytiva, Uppsala, Sweden) was added to the lysate and allowed to bind for 1 h at +4 °C with slow rotation. The gel slurry was poured into a disposable column and washed with lysis buffer until the flow-through had an $Abs_{280} \approx 0$. The protein was eluted with lysis buffer supplemented with 250 mM imidazole. Fractions were analysed for absorbance at 280 nm followed by SDS-PAGE.

IMAC was also used to remove the TEV protease from the full-length ZC3H11A, co-eluted in the Strep-Tactin®XT affinity purification. The IMAC resin was equilibrated in 20 mM Tris-HCl pH 9, 200 mM NaCl buffer and the sample was incubated with the resin for 30 min. Eluate containing full-length ZC3H11A free of TEV protease was then collected.

Amylose affinity chromatography. MBP-ZC3H11A$_{1\text{-}100}$ was purified using amylose resin (New England Biolabs, Ipswich, Massachusetts, US) and protocols from the provider. TEV protease was added to the pooled fractions at a concentration of 10 % relative to the eluted protein concentration, and incubated overnight. The eluate was subsequently concentrated for a second size exclusion chromatography step.

Size exclusion chromatography (SEC). SEC was done using an ÄKTA explorer and a HiLoad 16/600 Superdex 200 column (Cytiva, Uppsala, Sweden) equilibrated with the same buffer as used for lysing the cell pellet containing the protein. The concentrated protein sample was loaded at a flow rate of 0.5 mL/min. The flow rate was subsequently increased to 1 mL/min. Fractions were collected after approximately 40 mL flow time and analysed as above.

Protein concentration and storage. Pooled protein fractions from the chromatographic steps were concentrated using Amicon Ultra centrifugal filters (Burlington, Massachusetts, US). The concentrated protein was subsequently flash frozen in N_2 (l) and stored at −80 °C.

2.3. Analytical methods

Electrophoresis. Selected fractions were analysed by SDS-PAGE using precast gels (Invitrogen, Waltham, Massachusetts, US). Native electrophoresis was used to determine the molecular weight of native ZC3H11A. Electrophoresis was performed following instructions by the supplier of gels.

Mass spectrometry (MS). The identity of proteins separated by SDS-PAGE was determined using mass spectrometry of bands cut out from the gel (MS-based proteomics facility, Dept. Chemistry-BMC, or Mårten Larsson, Dept. of Med. Biochem. and Microbiology, Uppsala university, Uppsala, Sweden).

Structural and bioinformatic analysis. The predicted structure of ZC3H11A AF-O75152-F1-model_v4.pdb was downloaded from AlphaFold [16,17]. The isoelectric point of ZC3H11A was calculated using ProtParam in Expasy (www.expasy.org) [18]. PropKa was used to estimate the amino acid side chain pKa values of ZC3H11A [19]. The four TREX proteins used for our co-expression studies were originally identified as ZC3H11A binding partners by the STRING software tool [20].

3. Results

In order to obtain pure and stable ZC3H11A for subsequent experiments, a series of expression and purification trials using different expression constructs, conditions and hosts were performed with varying success (schematically outlined in Fig. 1).

3.1. Production of full-length ZC3H11A in E. coli

First, we attempted heterologous production of full-length ZC3H11A in *E. coli*, since expression in bacterial host systems is fast and relatively cost effective. The expression system was based on a pET vector and the T7 RNA polymerase promoter. This is generally a very efficient system that can yield large amounts of protein. The initial construct contained an N-terminal His-tag to ease protein purification. The tested constructs and outcomes of all performed expression and purification trials are summarised in Table 1. The relative success of each experiment was evaluated based on achieved solubility and purity, followed by stability and yield. Generally, the protein stability, in terms of resistance to proteolytic degradation, was evaluated by SDS-PAGE 2–3 days after purification. Yields obtained using different expression/purification strategies are estimated relative to each other, depending on different expression/purification strategies. Other biophysical methods, such as dynamic light scattering (DLS) were also used from time to time to analyse protein stability and aggregation.

Although SDS-PAGE analysis of cell lysates and samples taken at different steps during the purification confirmed that ZC3H11A was produced, the yields were very low and there were consistently many smaller contaminant bands visible in the gel (Table 1). Mass spectrometry (data not shown) verified that many of these bands were N-terminal fragments of ZC3H11A, which is consistent with the position of the purification tag. It indicates that the protein was degraded, either already in the cells, or rapidly upon cell lysis. Inclusion of protease inhibitors during purification did not prevent degradation. The purified full-length protein disappeared entirely upon storage at −80 °C, showing that degradation was difficult to avoid.

A range of alternative strategies were consequently explored to improve the production of the protein. Since the purity achieved after IMAC was very poor, the His-tag was replaced with an N-terminal Twin-Strep-tag to enable the use of an alternative affinity chromatographic procedure. Although this increased the purity of the protein, degradation and thus contamination with N-terminal fragments remained a problem. To increase the chance of isolating only full-length protein, a construct with a C-terminal His-tag was generated. Nevertheless, obtained ZC3H11A yields were poor and the problem of fast protein degradation persisted. To ease the burden on the expression host, we also tested a less efficient expression system based on the pKK223 vector and the tac promoter. Though generally successful, this approach did not solve the degradation problem (Table 1).

Next, we explored co-expression of ZC3H11A with the molecular chaperones GroEL/GroES, which significantly increased expression levels of the target protein. The purified ZC3H11A remained soluble and stable, however, only for as long as the chaperones were not removed. Thus, all attempts at bacterial expression of ZC3H11A did not give promising results, prompting us to explore alternative expression systems (Table 1).

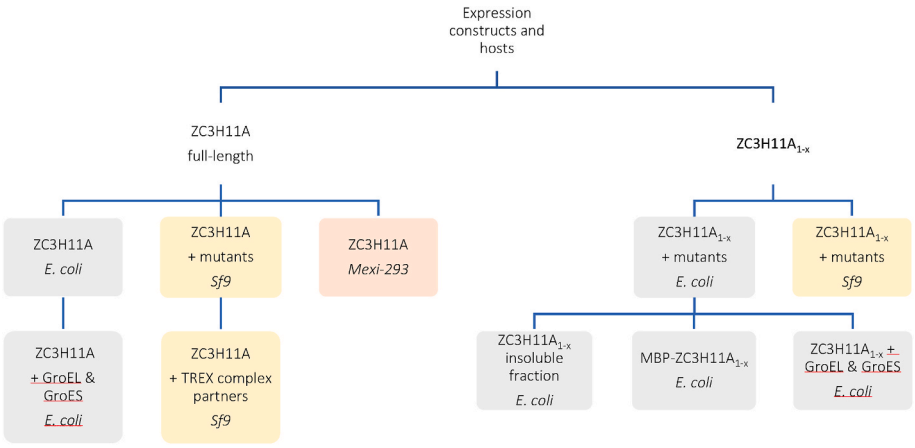


Fig. 1. Overview of strategies explored for recombinant production of ZC3H11A. *E. coli* (grey boxes), *Sf9* insect cells (yellow boxes), *Mexi-293* cells (rose). In the *E. coli* system, $ZC3H11A_{1-x}$ indicates a truncated version, where x is typically 100, but it was also 150, 200, 300 and 329.

Table 1
Summary of production outcomes for full-length ZC3H11A with different tags and co-expressions with chaperones and TREX proteins in mammalian cells, *E. coli* and *Sf9* cells. Successful production is highlighted by bold letters. The most successful strategy is highlighted by bold letters and framed. 2xS-t indicates that the construct had a Twin-Strep-Tag, while N-His and C-His represent N- or C-terminal His-tags, respectively.

Protein(s)	Host	Solubility[a]	Purity	Stability	Yield	Stable complex with co-expressed proteins
ZC3H11A 2xS-t	HEK293	Good	Low	Poor	Low	–
ZC3H11A N-His	*E. coli*	Poor	Low	Poor	Low	–
ZC3H11A C-His	*E. coli*	Poor	Low	Poor	Low	–
ZC3H11A 2xS-t	*E. coli*	Poor	Low	Poor	Low	–
ZC3H11A C-His and GroEL/ES	*E. coli*	Good	Low[b]	Transient	Intermediate	No
ZC3H11A 2xS-t and GroEL/ES	*E. coli*	Good	Low[b]	Transient	High	No
ZC3H11A 2xS-t and Thoc5 C-His	*E. coli*	Poor	Low	Transient	Low	No
ZC3H11A 2xS-t and Thoc7 C-His	*E. coli*	Poor	Low	Transient	Low	No
ZC3H11A 2xS-t	***Sf*9**	**Poor => good[c]**	**Low => high[c]**	**Poor => good[c]**	**Low => intermediate[c]**	–
ZC3H11A N190Q 2xS-t	*Sf*9	Poor	Low	Poor	Low	–
ZC3H11A C8S 2xS-t	***Sf*9**	**Good**	**High**	**Good**	**Intermediate**	**-**
ZC3H11A C8A 2xS-t	***Sf*9**	**Good**	**High**	**Good**	**Intermediate**	**-**
ZC3H11A 2xS-t and Thoc5	*Sf*9	Intermediate	Low	Transient	Intermediate	No
ZC3H11A 2xS-t and ALYREF	*Sf*9	Intermediate	Low	Transient	Intermediate	No
ZC3H11A 2xS-t and Thoc7	*Sf*9	Intermediate	Low	Transient	Low	No[d]
ZC3H11A 2xS-t and UAP56	*Sf*9	Poor	Low	Poor	Low	No
ZC3H11A 2xS-t, Thoc5 and ALYREF	*Sf*9	Intermediate	Low	Transient	Intermediate	No
ZC3H11A 2xS-t, Thoc5 and Thoc7	*Sf*9	Intermediate	Low	Transient	Intermediate	No
ZC3H11A 2xS-t, Thoc7, ALYREF and UAP56	*Sf*9	Intermediate	Low	Transient	Intermediate	No
ZC3H11A 2xS-t, Thoc5, Thoc7 and ALYREF	*Sf*9	Intermediate	Low	Transient	Intermediate	No
ZC3H11A 2xS-t, Thoc5, Thoc7, ALYREF and UAP56	*Sf*9	Intermediate	Low	Transient	Intermediate	No

[a] Protein in eluted fraction can be detected in a Coomassie stained SDS-PAGE gel, in some cases only for a limited time (a few days), and when co-expressed with chaperones, before removal of chaperones.
[b] Primarily contaminated by GroEL/ES.
[c] Initial procedure => optimised procedure.
[d] Thoc7 aggregates.

3.2. Production of full-length ZC3H11A in mammalian cells

Full-length ZC3H11A with an N-terminal Twin-Strep-tag was transiently expressed in Mexi-293E cells. The transfection was done at a cell viability of ~100 % for 120 h, after which the viability went down to ~75 %, the level recommended for cell harvest. According to the manufacturer information, this was more than double the expected time (~48 h). The protein was purified via Strep-Tactin®XT affinity column chromatography, and eluted fractions were analysed by SDS-PAGE (Fig. 2).

Since there were several bands present (Fig. 2) and none corresponded to the expected molecular weight of full-length ZC3H11A (~89 kDa), mass spectrometry was used to establish the identity of the proteins causing the major bands. It revealed that ZC3H11A was indeed successfully expressed in HEK-293 cells. However, it had a mass of ~110 kDa, which is significantly higher than the theoretical mass of the protein with the Twin-Strep-tag included. It thus appears that post-translational modifications (PTMs) are introduced during expression.

Moreover, two major contaminants were identified after purification, the 70 kDa heat shock chaperone protein (HSP70) and the multi-tasking SET protein (smaller amounts). It was not possible to remove these proteins from ZC3H11A despite exploration of many different expression and purification conditions. For example, $ZnCl_2$ was added to the transfection medium and salmon sperm DNA was added to the

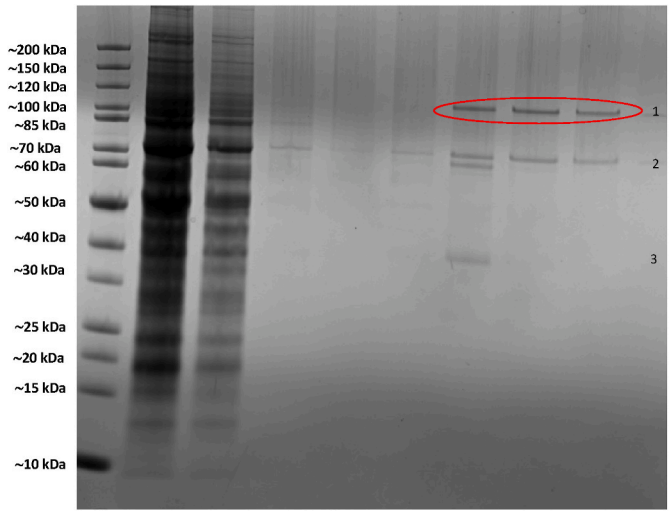


Fig. 2. SDS-PAGE gel of full-length ZC311A with N-terminal Twin-Strep-tag produced in mammalian cells and purified via Strep-Tactin®XT affinity chromatography. MS analysis identified 3 proteins in the purified sample (3 lanes on right hand side): 1) ZC3H11A, ~110 kDa, 2) heat shock 70 kDa protein 1A, 3) SET protein, 33.5 kDa.

medium after transfection to decrease stress on the cells, Mg-ATP was included in washing buffers and SEC was added as an extra purification step. We even added a denatured cell lysate at different concentrations to the washing buffer to capture free HSP70 in the lysate, with the aim to reduce its potential to interact with ZC3H11A [21]. None of these approaches were successful, indicating that HSP70 is clearly very tightly bound to the produced full-length ZC3H11A. Furthermore, yields were consistently relatively low, stability of the purified protein was poor, degrading quickly even at −80 °C, and the overall procedure was very time consuming, with 8 days passing between transfection and harvesting of cells. Taken together, this was not a useful system for production of ZC3H11A.

3.3. Production of full-length ZC3H11A in Sf9 cells

Initial experiments of heterologous production of ZC3H11A in insect cells were performed for the full-length protein alone, testing expression times ranging from 26 to 96 h in order to optimize protein yields. It was soon realized that although the protein was expressed, it was not stable in solution. The presence of protease inhibitors in all purification buffers did not prevent degradation. In fact, it was suspected that the process started already during expression. Nevertheless, efforts were made to speed up the purification process and keep temperatures low. Although the purification was done at +4 °C and the protein fractions were kept on ice whenever possible, it was not sufficient to prevent degradation. Already after the first purification step (affinity chromatography via the Twin-Strep-tag), considerable degradation was observed. The protein also degraded within a few weeks even when the protein was snap-frozen in liquid nitrogen and directly transferred to −80 °C. Moreover, the purity was poor, with both nucleic acid and proteins found as contaminants. Standard methods for removal of nucleic acid, such as addition of nucleases or high salt in buffers, did not work.

3.4. Production of ZC3H11A$_{1\text{-}100}$ in E. coli

The difficulties and problems experienced when expressing and purifying full-length ZC3H11A were assumed to be due to the predicted intrinsic disorder of the C-terminal region. A truncated form of ZC3H11A was therefore designed, including the first 100 amino acids at the N-terminus (ZC3H11A$_{1\text{-}100}$) which comprise the three zinc fingers, proceeded by a Twin-Strep-tag (Table 2). All expression and purification

Table 2
Summary of production outcomes for ZC3H11A$_{1\text{-}100}$ with different tags, fusion protein, mutants and co-expressions in *E. coli* and *Sf9* cells. Successful production is highlighted by a frame and bold letters. 2xS-t indicates that the construct had a Twin-Strep-Tag, while N-His represents an N-terminal His-tag.

Protein	Host	Result
ZC3H11A1-100 2xS-t	*E. coli*	Expressed but aggregates
ZC3H11A1-100 N-His	*E. coli*	Expressed but aggregates
ZC3H11A1-100 MBP-fusion	*E. coli*	Good expression, ZC3H11A1-100 and MBP not separable
ZC3H11A1-100 2xS-t and GroEL/ES	*E. coli*	ZC3H11A1-100 soluble, strongly bound to GroEL/ES
ZC3H11A1-100 C8A 2xS-t	*E. coli*	Expressed but aggregates
ZC3H11A1-100 C8S 2xS-t	*E. coli*	Expressed but aggregates
ZC3H11A1-100 2xS-t	***Sf9***	**Soluble**
ZC3H11A1-100 C8A 2xS-t	*Sf9*	Expressed but aggregates
ZC3H11A1-100 C8S 2xS-t	*Sf9*	Expressed but aggregates

experiments using this construct resulted in ZC3H11A$_{1\text{-}100}$ being exclusively found in the insoluble fraction (Fig. 3a).

All subsequent expression optimization experiments that explored *e.g.* the effect of using different constructs, co-expression of chaperones, and varying, induction temperatures and times, were equally unsuccessful - the protein consistently appeared in the insoluble fraction or tightly bound to chaperones (Fig. 3a). In contrast, the eluted ZC3H11A$_{1\text{-}100}$ fractions obtained using denaturing conditions were essentially pure (Fig. 3b). Different strategies, methods and commercial kits, were therefore used to refold the protein upon its extraction and purification from the insoluble fraction. However, the protein typically precipitated and all attempts failed.

To overcome the poor solubility of ZC3H11A$_{1\text{-}100}$, a construct fusing the truncated protein with Maltose Binding Protein (MBP) was generated. It includes a Tobacco Etch Virus (TEV) protease cleavage site between the two proteins, allowing MBP to be removed at the end of the purification process. The expression and purification were both successful but the TEV cleavage efficiency was low. Attempts to optimize cleavage conditions by variation of TEV protease concentration and buffer conditions were not successful. Nevertheless, ZC3H11A$_{1\text{-}100}$ represented around 50 % of the total protein purified. SEC was therefore employed to isolate the ZC3H11A$_{1\text{-}100}$ from MBP and un-cleaved MBP-ZC3H11A$_{1\text{-}100}$. Despite the large difference in size between ZC3H11A$_{1\text{-}100}$ and the fusion protein, ZC3H11A$_{1\text{-}100}$ was eluted together with the un-cleaved MBP-ZC3H11A$_{1\text{-}100}$ and cleaved MBP as a single peak. Using amylose affinity chromatography in addition to IEX chromatography in an attempt to separate them was also not successful (data not shown), indicating that there is a strong interaction between MBP- ZC3H11A$_{1\text{-}100}$ and ZC3H11A$_{1\text{-}100}$. It was clearly not a very promising strategy to produce ZC3H11A$_{1\text{-}100}$ in *E. coli.*

To establish if the length of the protein was critical for solubility, we also generated a set of truncated variants comprising the N-terminal 150, 200, 300 and 329 residues (Table S2). The exact length of the truncated constructs was arbitrarily chosen. Only the expression of the two shorter constructs was successful, however, the added 50 or 100 amino acids as compared to the original truncated construct ZC3H11A$_{1\text{-}100}$ did not improve solubility.

3.5. Production of ZC3H11A$_{1\text{-}100}$ in Sf9 insect cells

The production of ZC3H11A$_{1\text{-}100}$ was also explored in *Sf9* insect cells. A construct with a Twin-Strep-tag was used (Table 2). The few remaining impurities were removed in a second purification step using SEC. SDS-PAGE showed that that the protein elutes as one band and that no degradation products are present, *i.e.* is pure and stable (Fig. 4).

This production method was for a long period the only reliable source of any version of ZC3H11A, explaining why ZC3H11A$_{1\text{-}100}$ was initially used for our biophysical studies (Fekry et al., manuscript in

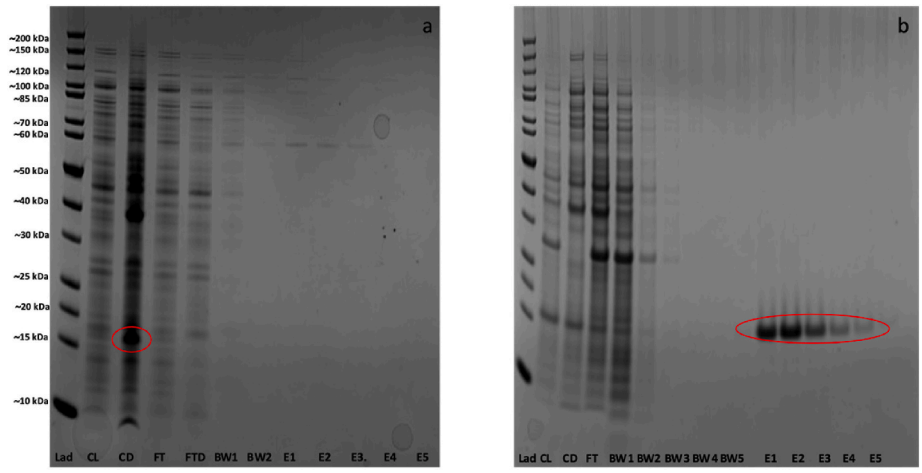


Fig. 3. SDS-PAGE gels of ZC3H11A$_{1\text{-}100}$ after expression in *E. coli.* and purification under (**a**) native conditions and (**b**) denaturing condition. ZC3H11A$_{1\text{-}100}$ is highlighted by red circles. In **a**) it is in the insoluble pellet fraction, while in **b**) it is in eluted fractions. The lanes in both gels are labelled as follows; Lad: SDS PAGE marker, CL: cell lysate, CD: cell debris, FT: flow through, FTD: flow through diluted, BW: buffer wash, and E: elution.

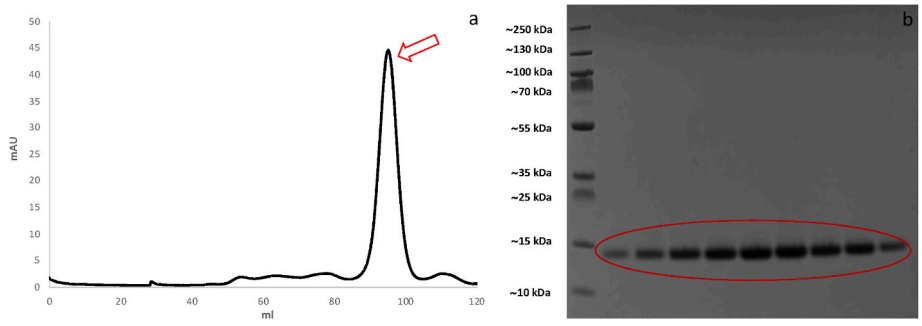


Fig. 4. Purification of ZC3H11A$_{1\text{-}100}$ expressed in insect cells. **a**) SEC chromatogram with red arrow pointing to the ZC3H11A$_{1\text{-}100}$ peak, and **b**) SDS-PAGE gel for purified ZC3H11A$_{1\text{-}100}$. The red oval surrounds the protein from the SEC chromatogram peak.

preparation). The yield varied from batch to batch since it depends on many factors, such as the viral stock status and the passage number. It was typically 3 mg ZC3H11A$_{1\text{-}100}$/L.

3.6. *Optimised purification of full-length ZC3H11A and ZC3H11A$_{1\text{-}100}$ (in Sf9 cells)*

While optimizing buffers for biophysical characterization of ZC3H11A$_{1\text{-}100}$ expressed in *Sf9* cells, it was found that it was more stable using an alkaline pH of 9 than the neutral buffers that had previously been used for production of the protein. Consequently, purification of both full-length ZC3H11A and ZC3H11A$_{1\text{-}100}$ at pH 9 was tried, using the otherwise identical production procedures established previously, comprising a Strep-Tactin®XT affinity chromatography step followed by SEC. Almost pure full-length ZC3H11A was obtained which, interestingly, eluted from the SEC column as an oligomer with an apparent molecular weight corresponding to ~400 kDa or higher (Fig. 5). The stability of the purified protein is very high, also upon storage at −80 °C (Table 1), and there were no signs of degradation occurring even after buffer exchange to pH 7.4 buffers. Obtained yields were typically around 0.8 mg ZC3H11A/L culture.

The oligomeric state of ZC3H11A was investigated using native PAGE (Fig. 5c). The gel shows a distribution of oligomers ranging in molecular weight between approx. 130 and 700 kDa. These bands are

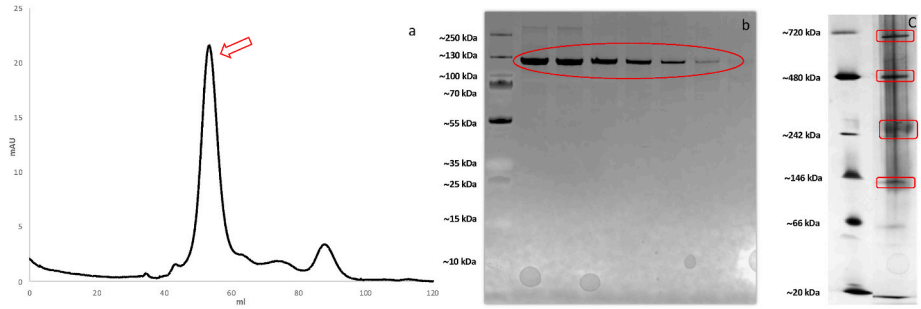


Fig. 5. Purification of ZC3H11A expressed in insect cells and purified at pH 9. **a**) SEC chromatogram with red arrow pointing to the ZC3H11A oligomer, **b**) SDS-PAGE gel for sample from the SEC peak in **a**). ZC3H11A is surrounded by the red oval, and **c**) native PAGE gel of the purified protein sample shown in b. The loaded sample contained 10 μg of protein and 100 mM NaCl. The red frame surrounds the ZC3H11A bands.

interpreted to represent species comprised of one, two, four and six subunits. The relative amounts of the different oligomers were affected by salt concentration but all bands were visible at the tested conditions of 50–200 mM NaCl.

3.7. Production of C8A and C8S variants of ZC3H11A in Sf9 cells and E. coli

In parallel with the ongoing production of ZC3H11A *in vitro*, *in vivo* studies with cysteine mutants of ZC3H11A revealed that Cys8, one of the putative zinc binding residues in the ZnF was critical for nucleic acid binding (personal communication, Tanel Punga, Uppsala university). This indicated that it would be possible to engineer ZC3H11A variants that did not bind tightly to nucleic acids. Therefore, Cys8 was targeted by point mutations to yield the alanine and serine substituted variants of truncated ZC3H11A$_{1\text{-}100}$ and full-length ZC3H11A for production attempts in both *E. coli* and insect cells (Tables S1 and S2).

The Cys8Ala and Cys8Ser variants of both full-length and ZC3H11A$_{1\text{-}100}$ were successfully expressed in insect cells. The serine variants gave higher yields than the alanine variants. Interestingly, the full-length protein gained stability by the substitutions and no degradation was seen after days of storage at 4 °C. Remarkably, the opposite was seen for ZC3H11A$_{1\text{-}100}$, with both cysteine variants ending up in the insoluble fraction, contrasting the successful purification of non-mutated ZC3H11A$_{1\text{-}100}$ expressed in insect cells. In *E. coli,* the expression of both cysteine variants of ZC3H11A$_{1\text{-}100}$ was the same as for the protein without substitutions, *i.e.* the protein is expressed but aggregates due to insolubility. In addition, it was observed that the substitutions did not reduce the binding of nucleic acid to either full-length or truncated ZC3H11A *in vitro*, which contradicts the *in vivo* experiments.

3.8. Co-expression of full-length ZC3H11A with TREX complex proteins in Sf9 cells

Since ZC3H11A interacts with other proteins in the TREX complex, a selection of TREX complex proteins were co-expressed with ZC3H11A to explore if this could have a stabilizing influence (Table 1). The TREX proteins were also expressed on their own (Table 3). The first small-scale co-expression trials were performed in *E. coli* with GST-fusion constructs of Thoc1, Thoc3 and Thoc7. No effect on the expression of ZC3H11A was seen for the first two. Subsequently, His-tagged UAP56, Thoc5, ALYREF (Thoc4) and Thoc7 were selected for co-expression. Different combinations were tested involving one, two, three or four of the TREX proteins, in addition to ZC3H11A.

Three out of four tested TREX proteins (ALYREF, Thoc5 and Thoc7) were found to have a positive impact on the stability of ZC3H11A. However, the mechanism by which this is achieved is not understood since no TREX protein-associated bands were visible in the gel upon SDS-PAGE analysis of ZC3H11A-containing fractions eluted from the affinity chromatography resin (with the possible exception of ALYREF). An example of elution fractions obtained by Strep-Tactin®XT affinity chromatography analysed by SDS-PAGE is shown in Fig. 6a. Possibly, the TREX proteins involve in transient interactions too weak to promote co-elution with ZC3H11A.

That the TREX proteins were indeed expressed could be confirmed by capture of the His-tagged TREX proteins from the pooled flow-through fractions of the Strep-Tactin®XT chromatographic step via IMAC. The SDS-PAGE gels of IMAC fractions from co-expression of ZC3H11A and four TREX proteins is shown in Fig. 6b. The analysis revealed that the lack of a stabilizing-effect of UAP56 on ZC3H11A is not explained by a failure to co-express, and that Thoc7 can positively affect ZC3H11A expression despite being prone to aggregate (it was identified in the non-soluble fraction).

ZC3H11A co-expressed with TREX proteins and purified by Strep-Tactin®XT affinity chromatography appears as a distinct band in the SDS gel corresponding to the expected molecular weight of approximately 110 kDa (Fig. 6a), *i.e.* the same molecular weight estimated from SDS-PAGE analysis of ZC3H11A expressed in mammalian cells. A single affinity chromatography purification making use of the strep-tag was not sufficient to separate ZC3H11A from all contaminants, and neither was an IMAC purification step sufficient to obtain reasonably pure TREX proteins (Fig. 6b). IEX and SEC were thus added as further purification steps, resulting once again in degraded ZC3H11A.

Table 3
Summary of TREX protein production in E. coli and Sf9 cells. Successful production is highlighted in bold.

Protein	Host	Result
Thoc5	***E. coli***	**Expressed and soluble**
Thoc7	*E. coli*	Expressed but aggregates
Thoc5	***Sf*9**	**Expressed and soluble**
Thoc7	*Sf*9	Expressed, but aggregates
ALYREF	***Sf*9**	**Expressed and soluble**
UAP56	***Sf*9**	**Expressed and soluble**

3.9. Structural and bioinformatic analysis

During the course of the project, an AlphaFold-predicted structure of ZC3H11A became available. It shows that only the N-terminal zinc finger domain is likely to be stably structured (Fig. 7). The remainder of the polypeptide chain is predicted to be largely disordered (Fig. 7). A few stretches are likely to form α-helical secondary structure, which together with extensive loop regions wrap themselves in a loose and solvent-exposed layer around the zinc-finger core.

Since the pH was critical for the production of the protein, we focused on analysing the distribution of charges in the structure (Fig. 8). As to be expected from a nucleic acid-binding domain, the predicted structure of the protein core exhibits a distinctly positively charged groove on its surface. And due to the high content (30 %) of amino acid side chains that are charged at neutral pH (Fig. S2), a general characteristic of intrinsically disordered proteins, basically all other parts of the ZC3H11A are highly hydrophilic as well.

ZC3H11A contains 119 negatively charged side chains (38 Asp and 81 Glu), and 128 positively charged side chains (85 Lys and 43 Arg), which are frequently clustered together, and occasionally form highly uniformly charged stretches (Fig. 8). For example, the 16-aa stretch from aa 161 to 176 contains eight Asp and three Glu residues, but no positively charged side chains. ProtParam estimates an isoelectric point of 8.6 for ZC3H11A, which is the pH where, according to the PropKa server, the protein is predicted to have the most negative free energy of folding, *i.e.* to be most stable.

It is generally assumed that chaperones have a high affinity for exposed hydrophobic regions of not properly folded target proteins. The preferred binding site of HSP70/DnaK has been described as a stretch of 4–5 hydrophobic acids (Leu, Ile, Val, Phe or Tyr) flanked by basic regions, and a strong bias against acidic residues [22]. A search for DnaK signature binding motifs yielded only two potential binding sites for this chaperone, stretch 441–453 (KKTVVLPPIVASR) and stretch 595–606 (KPVLTAVPGITR). Two additional exposed hydrophobic patches (677–683 and 700–708) that are however flanked by negatively charged residues on one site may represent binding sites for other chaperones, *e.g.* GroEL/ES.

4. Discussion

Despite the many challenges encountered during the production of ZC3H11A, we ultimately succeeded in obtaining full-length ZC3H11A in a pure and stable form. This required an extensive exploration of various expression and purification strategies. Although most were unsuccessful, their setups and outcomes are described here as they provide important insights into the characteristics of ZC3H11A, such as its inherently poor stability and high affinity for several other proteins and

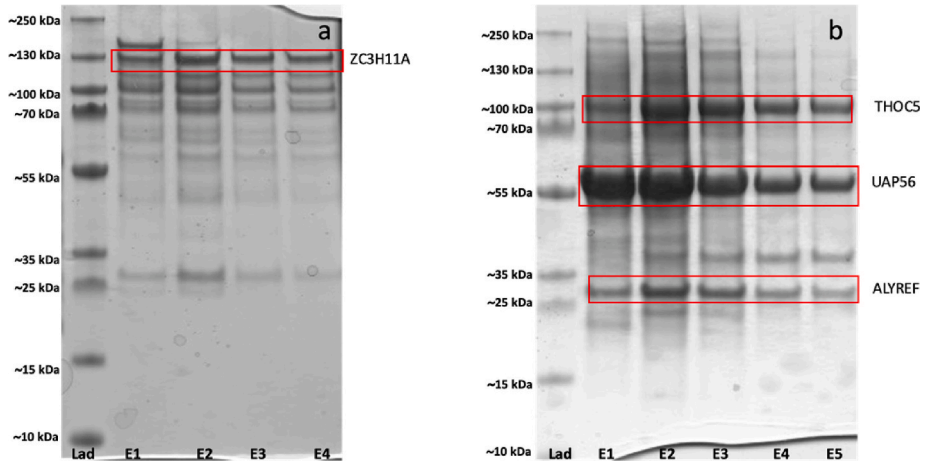


Fig. 6. SDS-PAGE gels showing an example of elution fractions from affinity chromatography after co-expression of ZC3H11A with the four TREX proteins Thoc5, UAP56, ALYREF and Thoc7. **a**) Eluted fractions from Strep-Tactin®XT affinity chromatography containing Strep-tagged ZC3H11A. **b**) IMAC elution fractions containing three of the four his-tagged TREX proteins, except Thoc7 did not enter the gel due to aggregation.

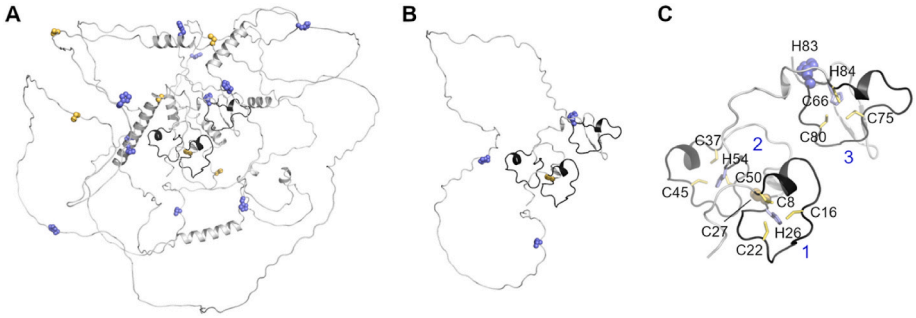


Fig. 7. Three-dimensional structure of ZC3H11A as predicted by AlphaFold (AF-O75152-F1-model_v4.pdb) [16,17]. (The predicted aligned error (PAE) for the structure is shown in Fig. S1). ZC3H11A is shown in cartoon presentation in light grey, with the three zinc fingers coloured dark grey. The cysteine (yellow) and histidine (blue) residues involved in zinc ion coordination are shown as sticks, all others as spheres. Full-length ZC3H11A (**A**) and the truncated versions ZC3H11A$_{1\text{-}200}$ (**B**) and ZC3H11A$_{1\text{-}100}$ (**C**) are all shown in the same orientation. In (**C**) the cysteine and histidine residues are labelled, and the three zinc fingers indicated by blue numbers (in order 1–3 counting from the N-terminus).

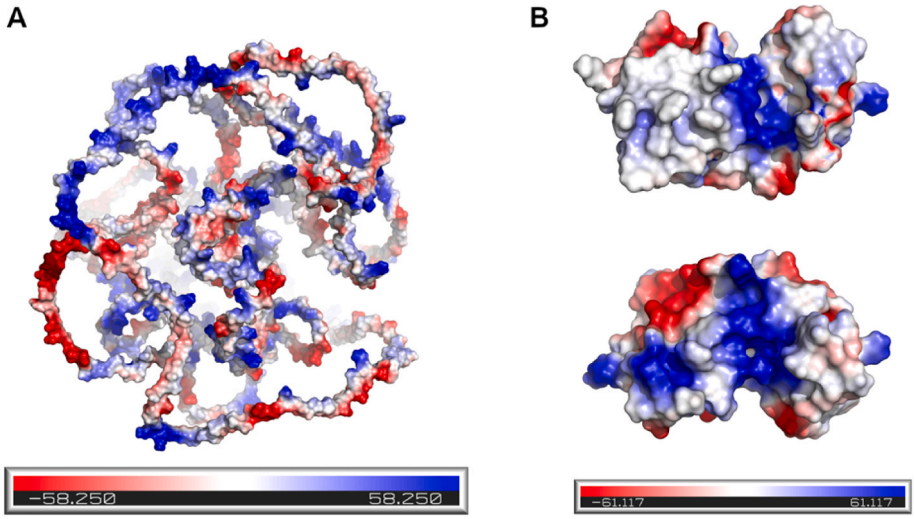


Fig. 8. Representation of the electrostatic surface potential of full-length ZC3H11A (**A**) and variant ZC3H11A$_{1\text{-}100}$ (**B**) as generated by the PyMOL Molecular Graphics System (Ver. 3.0 Schrödinger, LLC). In (**B**) the two views are related by a 180° rotation around an axis lying horizontally in the paper plane.

nucleic acids.

We initially assumed that the ZC3H11A would require to be fully folded to be functional, and thus focused on the full-length protein and conditions that would allow its production in a stable form. Nevertheless, we soon realized that this was very challenging and therefore explored production of various truncated constructs as well. Most of

these yielded poorly soluble or aggregated protein. Initially, only the ZC3H11A$_{1\text{-}100}$ variant encompassing the zinc-fingers was found to be soluble. Although other C-terminally truncated forms with up to 100 additional residues could be expressed in *E. coli*, there was no beneficial effect of these additional residues on solubility. For studies of the isolated ZnF domains, the short ZC3H11A$_{1\text{-}100}$ variant was thus preferred.

The fact that HSP70/DnaK was found as a major contaminant in several of the expression systems used was attributed to the tendency of this chaperone to be highly expressed and that its affinity for disordered regions makes it prone to bind to ZC3H11A [23]. Chaperone association was to a certain extent beneficial as it protected the protein against proteolytic degradation, but was ultimately problematic as they remained tightly bound and were difficult to remove in many of the experimental set-ups tested. Nevertheless, it inspired us to test whether chaperones could increase the stability of ZC3H11A. Co-expression with GroEL/GroES in *E. coli* resulted in the production of full-length ZC3H11A that was unstable once the chaperones were removed. Chaperone co-expression with ZC3H11A$_{1\text{-}100}$, *i.e.* the protein variant devoid of the disordered C-terminal region, resulted in a high yield of soluble protein, but it could not be separated from the chaperones. We have currently no explanation for the unusually strong association of the chaperones with ZC3H11A.

Similarly, being an mRNA-binding protein, ZC3H11A also has an innate affinity for nucleic acids. This resulted in difficulties in removing unspecified nucleic acid contaminants (likely mRNA, the natural binding partner for ZC3H11A) originating from the expression host. Standard enzymatic methods such as treatment with DNase, RNase, benzonase or high-salt buffers were unsuccessful, but precipitation with PEI, a cationic polyelectrolyte, worked. The predicted structure of ZC3H11A suggests that nucleic acids bound by the core may not be easily accessible to enzymes, may explain why only the precipitation approach was useful.

Our approach to reduce nucleic acid contaminations by targeting a zinc-coordinating cysteine, Cys8, for site directed mutagenesis was partially successful. Although the original goal of decreasing ZC3H11A's affinity to nucleic acids was not achieved, the full-length Cys8Ala and Cys8Ser mutant variants produced in insect cells were almost free of protein contaminants after purification, and considerably more stable in solution than the native protein. Surprisingly, the problem of contamination by nucleic acids persisted. Moreover, the positive effect of the amino acid exchanged on full-length protein stability was not transferable to the ZC3H11A$_{1\text{-}100}$ variant, which instead became completely insoluble. It thus appears that Cys8 is structurally important, and not only for binding of nucleic acid.

A novel approach to overcome all (or at least some) of the problems described above, was to co-express ZC3H11A with its natural binding partners within TREX complex. This was originally promising but none of the co-expressions resulted in pure or stable ZC3H11A. Production of individual TREX proteins was also attempted. All four were easily expressed, and all except Thoc7 were stable in solution. However, they were difficult to purify and appeared to bind to many cellular proteins that could not be removed by affinity chromatography. This might not be surprising considering that the TREX complex is a very large multimeric structure, as has recently been illustrated by cryo-electron microscopy of mRNA and some of the proteins in the complex (not ZC3H11A) [24].

Ultimately, it was a shift in the pH of lysis and purification buffers towards the alkaline (from pH 7.4 or 8.0 to 9.0) that yielded the most profound breakthrough regarding production of full-length ZC3H11A. This pH shift towards the pI should render the protein less soluble, but, according to the PropKa server prediction, also more stable.

The stability of ZC3H11A is predicted to be highest at 8.6, (*i.e.* at its predicted pKa), and the protein is expected to retain 80 % of its maximum stability in the pH range 6.5–9.8. This covers the pH values of the cell lysis and affinity chromatography buffers (pH 7.4 or 8) initially used. Interestingly, once the protein was obtained in a pure and stable form using high pH buffers, it tolerated being transferred to a buffer with lower pH.

To rationalise how a net charge shift closer to 0 and/or a switch in charge state of specific residues somehow can induce favourable interactions between previously mobile parts of ZC3H11A and thus an increase in distinct substructures, we looked at the structure and sequence of the protein. The only residue types changing charge state within this pH range are Cys and possibly Tyr and Lys residues, for which the pK_a may shift towards the acidic depening on their local environment. However, with almost the entire polypeptide chain being intrinsically disordered and thus solvent-exposed, such pKa shifts are unlikely to occur. Fittingly, the PropKa server does not predict a shift of sufficient magnitude and in the correct direction for any of the lysine and tyrosine residues. A possibility is that Cys residues are critical for general protein stability and not only zinc-finger formation. The currently available data are not sufficient to answer this question satisfactorily.

5. Conclusions

The characteristics and functions of intrinsically disordered proteins are now considerably better understood than when this project started [25], making the loose arrangement of the largest part of ZC3H11A a relevant property of the protein required for function, with instability as a natural consequence rather than caused by suboptimal recombinant production procedures. In hindsight, our experience with the production of ZC3H11A for biochemical studies reflect typical challenges of intrinsically disordered proteins. The establishment of protocols for recombinant production of full-length, truncated and mutant forms of ZC3H11A has been a major effort but gave important insights into how to tackle production of structurally disordered proteins as well as the features of this protein. The availability of a predicted 3D structure and other prediction tools have clearly been useful once we had experimental data, but much work remains before we understand the biophysical and functional characteristics of ZC3H11A. The reliable production of ZC3H11A variants suited for such experiments, is critical, and has been the aim of this work.

However, many questions remain to be resolved to further our understanding of the function ZC3H11A in a biological context. For example, the role of the disordered C-terminal region, and how it differs from the ordered N-terminal domain with respect to its interactions with mRNA and the other components of the TREX complex involved in mRNA export. Specifically, it would be interesting to obtain a deeper understanding of the role of Cys8, since it does not appear to be critical for nucleic acid binding. Further studies can inspire novel therapeutic strategies for diseases involving ZC3H11A function, such as those caused by nuclearly replicating viruses.

Declaration of interests

The authors have no competing interests to declare.

CRediT authorship contribution statement

Mostafa Fekry: Writing – review & editing, Writing – original draft, Methodology, Investigation. **Gun Stenberg:** Writing – review & editing, Supervision, Methodology, Investigation. **Doreen Dobritzsch:** Writing – review & editing, Visualization. **U. Helena Danielson:** Writing – review & editing, Supervision, Resources, Funding acquisition, Conceptualization.

Data availability

Data will be made available on request.

Acknowledgements

This work was funded by the Knut and Alice Wallenberg Foundation grant no. 2017.0071. The patience and continued support from Prof. Leif Andersson (Department of Medical Biochemistry and Microbiology, Uppsala University), the leader of the overall project is much appreciated. Tanel Punga (Department of Medical Biochemistry and Microbiology, Uppsala University) has been very helpful by sharing information about his cellular studies. Mårten Larsson (Department of Medical Biochemistry and Microbiology, Uppsala University) performed some of the MS analysis in the project. Lidija Savic, Julia Schwekendiek and Daria Kovryzhenko (MSc students in biochemistry at the Dept. of Chemistry – BMC, Uppsala University) helped to reproduce and confirm some of the results.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pep.2024.106542.

References

[1] M. Cassandri, et al., Zinc-finger proteins in health and disease, Cell Death Discov 3 (2017) 17071.
[2] J.H. Laity, B.M. Lee, P.E. Wright, Zinc finger proteins: new insights into structural and functional diversity, Curr. Opin. Struct. Biol. 11 (1) (2001) 39–46.
[3] J. Miller, A.D. McLachlan, A. Klug, Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes, EMBO J. 4 (6) (1985) 1609–1614.
[4] S.A. Beausoleil, et al., A probability-based approach for high-throughput protein phosphorylation analysis and site localization, Nat. Biotechnol. 24 (10) (2006) 1285–1292.
[5] N. Dephoure, et al., A quantitative atlas of mitotic phosphorylation, Proc Natl Acad Sci U S A 105 (31) (2008) 10762–10767.
[6] K. Maeda, S. Akira, Regulation of mRNA stability by CCCH-type zinc-finger proteins in immune cells, Int. Immunol. 29 (4) (2017) 149–155.
[7] M. Fu, P.J. Blackshear, RNA-binding proteins in immune regulation: a focus on CCCH zinc finger proteins, Nat. Rev. Immunol. 17 (2) (2017) 130–143.
[8] J. Liang, et al., Genome-wide survey and expression profiling of CCCH-zinc finger family reveals a functional module in macrophage activation, PLoS One 3 (8) (2008) e2880.
[9] M.Y. Hein, et al., A human interactome in three quantitative dimensions organized by stoichiometries and abundances, Cell 163 (3) (2015) 712–723.
[10] E.G. Folco, et al., The proteins PDIP3 and ZC11A associate with the human TREX complex in an ATP-dependent manner and function in mRNA export, PLoS One 7 (8) (2012) e43804.
[11] V.O. Wickramasinghe, R.A. Laskey, Control of mammalian gene expression by selective mRNA export, Nat. Rev. Mol. Cell Biol. 16 (7) (2015) 431–442.
[12] Z. Hajikhezri, et al., Role of CCCH-type zinc finger proteins in human adenovirus infections, Viruses 12 (11) (2020).
[13] S. Younis, et al., Multiple nuclear-replicating viruses require the stress-induced protein ZC3H11A for efficient growth, Proc Natl Acad Sci U S A 115 (16) (2018) E3808–E3816.
[14] L. Yang, et al., Porcine ZC3H11A is essential for the proliferation of pseudorabies virus and porcine circovirus 2, ACS Infect. Dis. 8 (6) (2022) 1179–1190.
[15] J. Li, et al., Analysis of the mRNA export protein ZC3H11A in HCMV infection and pan-cancer, Front. Microbiol. 14 (2023) 1296725.
[16] J. Jumper, et al., Highly accurate protein structure prediction with AlphaFold, Nature 596 (7873) (2021) 583–589.
[17] M. Varadi, et al., AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models, Nucleic Acids Res. 50 (D1) (2022) D439–D444.
[18] E. Gasteiger, et al., Protein identification and analysis tools on the Expasy server, in: J.M. Walker (Ed.), The Proteomics Protocols Handbook, Humana Press, 2005.
[19] M. Rostkowski, et al., Graphical analysis of pH-dependent properties of proteins predicted using PROPKA, BMC Struct. Biol. 11 (2011).
[20] D. Szklarczyk, et al., STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets, Nucleic Acids Res. 47 (D1) (2019) D607–D613.
[21] D.V. Rial, E.A. Ceccarelli, Removal of DnaK contamination during fusion protein purifications, Protein Expr. Purif. 25 (3) (2002) 503–507.
[22] S. Rudiger, et al., Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries, EMBO J. 16 (7) (1997) 1501–1507.
[23] M.R. Fernandez-Fernandez, et al., Hsp70 - a master regulator in protein degradation, FEBS Lett. 591 (17) (2017) 2648–2660.
[24] B. Pacheco-Fiallos, et al., mRNA recognition and packaging by the human transcription-export complex, Nature 616 (7958) (2023) 828–835.
[25] A.S. Holehouse, B.B. Kragelund, The molecular basis for cellular function of intrinsically disordered protein regions, Nat. Rev. Mol. Cell Biol. 25 (3) (2024) 187–211.