

## Crystallographic conditions of the heterologally expressed recombinant metal-binding protein Ts-PCHTP

H. I. Sbirikova<sup>1\*</sup>, G. A. Radoslavov<sup>2</sup>, P. I. Hristov<sup>2</sup>, B. L. Shivachev<sup>1</sup>

<sup>1</sup> Institute of Mineralogy and Crystallography, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 107, 1113 Sofia, Bulgaria

<sup>2</sup> Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 25, 1113 Sofia, Bulgaria

Received February, 2013; Revised May, 2013

We have optimized the expression of Polycysteine and histidine-tailed protein (Ts-PCHTP). The protein contains a natural His-tag for metal binding, used in the protein purification process. The “tag” amino acid sequence binds very strongly to conventional Ni affinity chromatography columns and permits the protein elution with EDTA. The protein purification should be monitored by Native-PAGE as the high content of cysteine in the Ts-PCHTP protein sequence produces false positives, when SDS-PAGE technique is used. According to the DLS data a dimmer form of Ts-PCHTP is observed. Sparse matrix crystallization screens were used for the determination of the most favorable conditions for growth of crystals suitable for X-ray single crystal.

**Key words:** recombinant protein; His Tag affinity chromatography; Metal binding protein; Dynamic Light Scattering; Protein crystallization.

### INTRODUCTION

The ultimate goal of structural biology is to understand the structural basis of proteins in cellular processes. In structural biology, the most critical issue is the availability of high-quality samples. Thus “structural biology grade” proteins must be produced in a suitable quantity and quality allowing crystallization (screening for crystallization conditions) and structure determination. The purification procedure must be reproducible and yield homogeneous protein in milligram quantities. Therefore the choice of protein purification and quality control procedures play a key role in obtaining protein samples in sufficient amounts and with high purity.

PolyCysteine and Histidine-Tailed protein (Ts-PCHTP) is a metalloprotein expressed in muscle larvae of nematode *Trichinella spiralis*. Metalloprotein is a term used for proteins (e.g. ferritin) which binds, transports or stores, metal ions or enzymes (metalloenzymes), in which the metal atom is involved in important catalytic processes. Ts-PCHTP is the first described member of a new nematode-specific pro-

tein family PCHTP – Poly-cysteine proteins, which are unique to Order Trichocephalida [1]. Ts-PCHTP binds different bivalent metal ions such as Fe, Ni and Zn. Its most probable function is transport or storage of the metal ions in the parasite. Heavy metal ions exhibit positive and negative effects on biological processes. On one hand they act as a cofactor for a number of enzymes, which catalyze numerous metabolic reactions. On the other hand, they can destroy the native structure of proteins or nucleic acids, which requires their detoxification. Proteins, which bind bivalent metal cations, contain specific amino acids (aa) sequences forming metal binding sites. These sequences contain numerous aa such as cysteine, histidine, methionine, serine, threonine, tyrosine, asparagine, glutamine, aspartate and glutamate [2, 3].

The *Ts-pchtp* gene length is 1896 bp [1]. It contains six exons and five introns. The nucleotide sequence of the mRNA comprises 1576 bp. The transcript has a single open reading frame consisting of 1272 bp. The primary structure of Ts-PCHTP includes 424 aa with a molecular weight of 47744 Da. Protein sequence shows a high content of the following amino acids: 36 cysteine (8.5%), histidine 26 (6.4%), 10 tyrosine and 8 tryptophan. The protein consists of a signal peptide for extracellular localization in

\* To whom all correspondence should be sent:  
E-mail: sbirikova@mail.bg

the N-term, two homologous polycysteine domains (pcd) – pcd-1 and pcd-2, poly-histidine domain (histidine tag) at the C-term. Another characteristic of the cysteine rich and metal binding proteins is predominantly  $\beta$ -structure [4, 5]. Ts-PCHTP is mainly characterized by  $\beta$ -structural conformation. It is typical for many metalloproteins that they can also bind specifically or nonspecifically hydrophobic ligands, such as fatty acids or hem [4, 6].

Poly-histidine regions are often associated with the binding of transition metal ions such as  $\text{Ni}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Co}^{2+}$ ,  $\text{Mn}^{2+}$ ,  $\text{Fe}^{2+}$ ,  $\text{Fe}^{3+}$ . These metal binding sites can be constructed either from consecutive histidine residues or repeats of histidine in combination with one to three aa between them [7–9]. The poly-histidine motifs common observed in the middle part of protein chain rather than C- and N-term (Histidine rich glycoproteins etc.).

The poly-histidine motif is one of the most frequently used in the purification of recombinant proteins (*his-tag* protein) [9]. Typically the construct (*his-tag*) is composed by six or more histidine residues, which are located in C- or N-term of a recombinant protein. The method is one of the most common and is based on the affinity of the histidine to bivalent metal ions. Purification was performed by metal affinity chromatography –  $\text{Ni}^{2+}$  or  $\text{Co}^{2+}$  chromatography.

In this study we present successful conditions for production of the recombinant Ts-PCHTP (rTs-PCHTP) in *Escherichia coli* strains. Recombinant protein with natural His-tag at the C-end was obtained by metal affinity chromatography similarity to the common recombinant His-tag purification techniques. Purified rTs-PCHTP was used for a crystallographic assay and Dynamic Light Scattering (DLS) analyses.

## EXPERIMENTAL METHODS

### *cDNA synthesis and PCR amplification*

The cDNA synthesis and amplification were performed as described by Radoslavov *et al.* [1]. Briefly, total RNA from *Trichinella spiralis* (T1 ISS03) muscle larvae was extracted after homogenization with Trizol reagent according to the manufacturer's instructions (Invitrogen).

Reverse cDNA transcription was performed with 4  $\mu\text{g}$  of RNA, gene specific primer AS3 and M-MuLV Reverse Transcriptase (Fermentas). The nucleotide sequence from the second exon to the stop codon ORF of the Ts-TCHTP (GeneBank™ accession JF899252 and GQ497342) was amplified with the following primers: S22 5'-ggaattccatgatgacaaaatttcgtcgccga-3' and AS3 5'-cgcgatccttatca

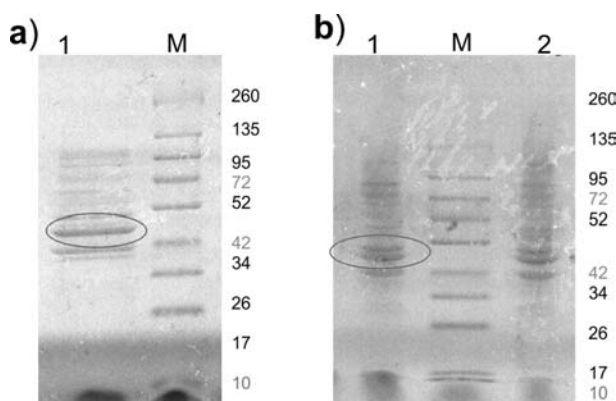
atgatgatgatgatgatg-3', containing *Bam*HI and *Nde*I restriction sites for direct cloning into pJC20 expression vector [10]. PCR product and vector were ligated with T4-DNA Ligase (Fermentas) by standard protocol [11] and transformed into *Escherichia coli* DH5 $\alpha$  cells. The positive colonies were checked by colony PCR and DNA sequencing (Macrogen).

### *Production and purification of recombinant protein*

The *E. coli* BL21(DE3), C43(DE3) and Rosetta-gami(DE3) cells were used for the expression of the recombinant protein. The induction was performed with 1 mM IPTG (final concentration), for 4 hours at 37 °C. Cells were lysed through sonication on ice. The soluble fraction contains Ts-PCHTP. In consequence Ni-affinity chromatography, HisTrap Kit (GE Healthcare BioSciences), was employed for rTs-PCHTP purification.

The protein concentration was determined initially by the method of Bradford [12] and spectrophotometrically using a molar extinction coefficient of 1.332  $\text{M}^{-1}$  at 280 nm as calculated on the basis of the aromatic amino acid of 10 Tyrosine (Tyr) and 8 Tryptophan (Trp) residues with ProtParam tool [13] (www.expasy.org). Protein fractions were dialyzed with PD-10 desalting column, Sephadex G25 (Amersham Bioscience) against 0.02M Tris buffer, pH 7.2. The protein was concentrated up to 4 mg/ml with Centricon columns, 10kDa (Millipore).

The molecular size and purity of the protein was determined by 12% SDS-PAGE and 12% Native-PAGE [11]. Proteins were visualized by Colloidal Coomassie G-250 Blue Safe staining (National Diagnostics) (Fig. 1).



**Fig. 1.** a) 12% Native-PAGE gel of purified rTs-PCHTP lysate (elution with 0.05M EDTA) in to C43DE3 cells and b) 12% SDS-PAGE gel of purified rTs-PCHTP lysate (elution with 0.05M EDTA) in to C43DE3 cells. Proteins were visualized by Colloidal Coomassie Blue staining

### Dynamic Light Scattering (DLS)

DLS measurements on the protein solutions were obtained using a 90Plus (Brookhaven Instruments), equipped with a 657 nm, 35mW laser. Time dependent fluctuations in the scattered intensity were measured using an avalanche photo detector (APD) and digital correlator. Protein solutions were analyzed in triplicate at 25.0 °C and scattering angle of 90°. The buffer solutions were filtered through 0.44 µm and 0.022 µm filters and were also analyzed by DLS. Standard NIST traceable polystyrene 22 nm ± 1.8 nm, 92 ± 2 nm latex standards and a blank, 0.02 µm filtered ultrapure water (Nanopure, Thermo Scientific), were also run as standards. Data was collected as co-added runs of 2 min collected for a total of 10 min. The autocorrelation functions were deconvoluted to obtain size distributions using both the non-negatively constrained least squares fit (multiple pass NNLS) algorithm. The size distributions obtained from the NNLS algorithm were presented since the distributions are bimodal. The intensity of scattered light is proportional to the particle size to the sixth power, which results in a higher scattered intensity for larger particles. The intensity weight distributions, measured by DLS, were converted to number weighted distributions using analysis software provided by Brookhaven (BIC).

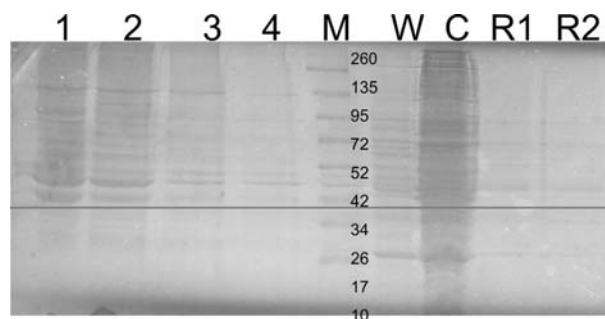
### Crystallographic conditions

Crystallization attempts were performed using the vapor-diffusion method, in which 1 µL of Ts-PCHTP (3.6 mg ml<sup>-1</sup> in 0.02 M Tris 7.3 pH) solution was added to 1 µL of reservoir solution in a hanging drop suspended over 0.6 ml of reservoir solution. The reservoir (crystallization) conditions were those of the commercial Hampton Research Crystal screen HT, PEG/Ion and PEG/Ion 2.

## RESULTS AND DISCUSSION

DNA sequencing of the pJC20 construct showed that the inserted fragment length is 1221 bp. Detailed analysis of the sequencing results showed that was the amplification and cloning, that the insert has an additional ATG initial codon at 5' end of the sequence. Thus the construct will translate 408 aa from the Native Ts-PCHTP (424 aa), with predicted molecule weight of 46 kDa. The 408 aa sequence includes the two full poly-cysteine domains and the poly His-Tag end, while the starting signaling peptide sequence is omitted. DNA sequencing showed also that the construct reading frame starts and ends correctly, so thus will allow subsequent expression.

The three *E. coli* bacterial strains were transformed with the expression vector: BL21(DE3), C43(DE3) and Rosetta-gami(DE3). The recombinant protein is isolated from soluble protein fraction under normal aerobic conditions, even that the cysteine constitutes 8.5% of all aa in Ts-PCHTP. 12% SDS-PAGE of bacterial lysates showed that only the C43(DE3) cells produced sufficient quantity of the recombinant protein (Fig. 2). The C43(DE3) cells are designed for expression of difficult or even toxic proteins. This suggests that Ts-PCHTP may be toxic due to its metal binding properties. Ts-PCHTP contains 26 histidine residues or 6.1% of the amino acid content in the protein. Fourteen histidine residues are localized in the C-terminus of the protein, which gives the name of the domain – poly-histidine tag. Seven consecutive histidine residues form a motif at the C-end of the protein. This histidine repeat at the C-term allows rTs-PCHTP to be purified with Ni-affinity chromatography (His Trap kit / GE Healthcare BioSciences). Sonication of C43(DE3) was performed in 10 mM imidazole pH 7.9, followed by centrifugation. The supernatant was used for loading on a HisTrap column. Attempting to elude the protein with high concentration of Imidazole (up to 1M) in some of the cases failed. The protein was successfully eluted with 0.05 M EDTA, pH 8.0. The concentration of the eluted protein was measured by Bradford [12] and spectrophotometrically at 280 nm. The molecular weight and purity of the protein was determined by 12% SDS-PAGE and 12% Native-PAGE. The protein is visualized on the gel at the expected molecular weight of 46 kDa (Fig. 1). The Native-PAGE showed better separation and single bands, in contrast to the SDS-PAGE, where multi bands



**Fig. 2.** 12% Native-PAGE of rTs-PCHTP lysate (elution with 0.05M EDTA): bands 1–4 correspond to C43DE3, M – protein marker, W – Wash cytosol after IPTG, C – Cytosol after IPTG, R1-R2 – rTs-PCHTP in Rosetta cells. Proteins were visualized by Colloidal Coomassie Blue staining

were visualized. This is an indication that the protein is more stable into non reductive conditions.

rTs-PCHTP was used for DLS analysis to evaluate protein purity. The estimation of the molecular weight of rTs-PCHTP was carried out according to the Mark-Houwink-Sakurada equation  $D = KM^a$  [15, 16] where  $D$  is the diffusion coefficient (assessed by DLS,  $D = kT/(6\pi\eta R_h)$ ),  $M$  is the molecular weight and  $K$  and  $a$  are two empirically determined parameters that are solvent, temperature and protein specific. The molecular weight for the first size (peak of 39.4 nm) of the bimodal distribution corresponds to 110 kDa (Fig. 3b and 3c). This value is more than two times bigger than the theoretical one of 46

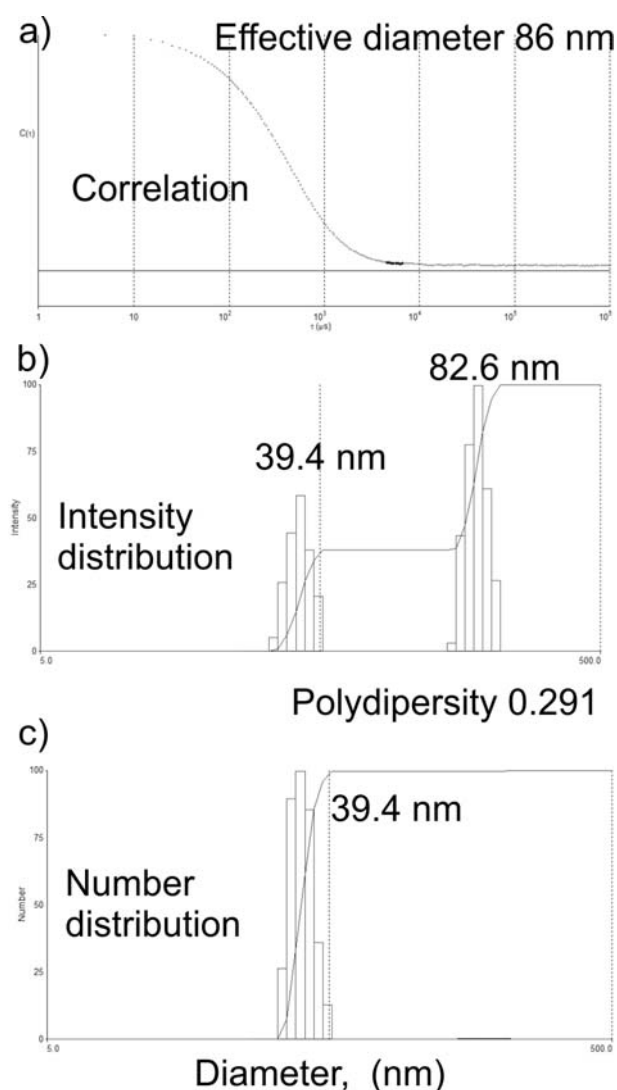
kDa. Thus the first peak in the size distribution can be related to a dimer Ts-PCHTP. Although proteins generally fold into compact globular domains that are tightly packed (no voids or water molecules are accommodated in the interior of the protein) the surface of the protein is rather “bumpy” and is covered by an outer shell of water/solvent molecules. The DLS results for  $R_h$  will be affected by that shell and will overestimate the protein molecular weight. The second peak (82.6 nm,  $R_h$  of 41.3 nm) can be attributed to additional aggregation of Ts-PCHTP (octamer or dodecamer). As one can see from the multimodal distribution “number vs. diameter” the contribution of Ts-PCHTP dimer is close to 100%, while the higher 8- or 12-mer is almost indiscernible (Fig. 3c).

#### Preliminary crystallization results

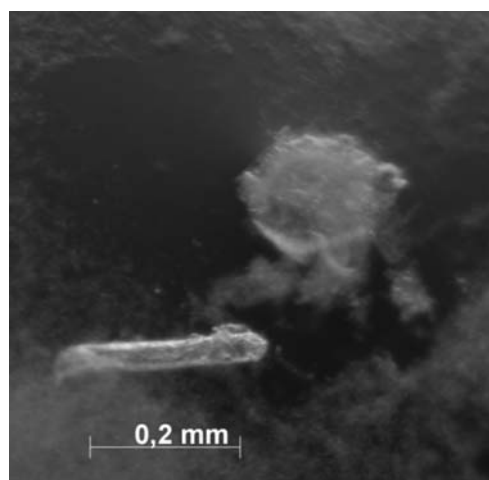
The initial sparse matrix screening for Ts-PCHTP crystallization showed that the crystallization conditions listed in Table 1 are suitable for the growth of crystals for X-ray data collection (Fig. 4).

#### CONCLUSIONS

The present investigation provides insights into the heterologous expression, purification and crystallization of rTs-PCHTP. The protein was expressed and purified from soluble protein fraction at native condition. The purification of the rTs-PCHTP with Ni affinity chromatography is based on its nature poly his tag (7 histidine residues). Ts-PCHTP



**Fig. 3.** Representative DLS for rTs-PCHTP 3.6 mg ml<sup>-1</sup> in Tris pH 7.3 a) correlation function b) intensity particle size distribution and c) number particle size distribution



**Fig. 4.** Observed crystal of rTs-PCHTP, No 41 of Hampton Crystal screen 2

**Table 1.** Formulation of the crystallization conditions for Ts-PCHTP crystal growth

Hampton Screen	No	salt	buffer	precipitant
Crystal screen 1	46	0.2 M Ca acetate hydrate	0.1 M Na cacodylate trihydrate, pH 6.5	18% w/v PEG 800
Crystal screen 2	33	2M Ammonium formate	0.1M HEPES pH 7.5	
Crystal screen 2	41	0.01M Ni(II) chloride hexahydrate	1M Tris pH 8.5	1M Lithium sulfate monohydrate 0.
PEGI	2	0.2 M Potassium fluoride		20% w/v PEG 3350
PEGI	41	0.2 Potassium phosphate monobasic		20% w/v PEG 3350

describes for the first time a natural protein with this amino acid motif. DLS analysis showed that protein forms thoroughly dimmer formations which are not visualized at native gel conditions. There are identified several different suitable crystallization conditions for Ts-PCHTP. The data can be used for X-ray analysis.

**Acknowledgments:** The authors thank the financial support of the Bulgarian National Science Fund through grant DRNF 02/1.

#### REFERENCES

- G. Radoslavov, R. Jordanova, D. Teofanova, K. Georgieva, P. Hristov, M. Salomone-Stagni, E. Liebau, I. Bankov, *PLoS One*, **5** (10), e13343 (2010).
- D. Gregory, A. Martin, J. Cheetham, A. Rees, *Protein Engineering, Design and Selection*, **6**, 29 (1993).
- M. Yamashita, L. Wesson, G. Eisenman, D. Eisenberg, *Proc. Natl. Acad. Sci., USA*, **87** (15), 5648 (1990).
- D. Matthews, *Curr. Opin. Biotechnol.*, **6** (4), 419 (1995).
- N. Shu, T. Zhou, S. Hovmöller, *Bioinformatics*, **24** (6), 775 (2008).
- L. Yi, *Metalloprotein Design & Engineering*. Encyclopedia of Inorganic Chemistry, Second Edition, ed. R. Bruce King, **5**, 3159 (2005).
- A. Jones, M. Hulett, *Immunol. Cell Biol.*, **83** (2), 106 (2005).
- R. Garrett, C. Grisham, *Biochemistry of Metal Ions*. (Book Reviews: The Inorganic Chemistry of Biological Processes). *Science*, **4104**, 181, 1995, p. 1037.
- P. Hengen, *Trends Biochem. Sci.*, **20** (7), 285 (2005).
- J. Clos, S. Brandau, *Protein Expr. Purif.*, **5** (2), 133 (1994).
- J. Sambrook, E. Fritsch, T. Maniatis, *Molecular Cloning, A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, New York, USA, 1989.
- M. Bradford, *Anal. Biochem.*, **72**, 248 (1976).
- E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. Wilkins, *Protein Identification and Analysis Tools on the ExPASy Server*, in: John M. Walker (ed), *The Proteomics Protocols Handbook*, Humana Press, USA, 2005, p. 571.
- P. Hiemenz, T. Lodge, *Polymer Chemistry*, Second ed. Boca Raton: CRC P, 336, (2007).
- G. Borgstahl, *Methods Mol. Biol.*, **363**, 109 (2007).
- A. Ferre-D'Amare, S. Burley, *Structure*, **2**, 357 (1994).

## КРИСТАЛИЗАЦИОННИ УСЛОВИЯ НА ХЕТЕРОЛОЖНО ЕКСПРЕСИРАН РЕКОМБИНАНТЕН МЕТАЛОСВЪРЗАЩ БЕЛТЪК TS-PCNTP

Х. И. Сбиркова<sup>1\*</sup>, Г. А. Радославов<sup>2</sup>, П. И. Христов<sup>2</sup>, Б. Л. Шивачев<sup>1</sup>

<sup>1</sup> *Институт по Минералогия и Кристалография, Българска Академия на Науките, ул. „Акад. Г. Бончев“, бл. 107, 1113 София, България*

<sup>2</sup> *Институт по Биоразнообразие и Екосистемни Изследвания, Българска Академия на Науките, ул. „Акад. Г. Бончев“, бл. 25, 1113 София, България*

Постъпила февруари, 2013 г.; приета май, 2013 г.

(Резюме)

В настоящата работа е оптимизирана експресията на протеин с полицистеинови домейни (Ts-PCNTP). Протеинът съдържа естествен фрагмент (His-Tag) за метална координация, използван в процеса на пречистване, чрез Ni афинитетна хроматография. Чистотата на белтъка се следи чрез нативна-полиакриламидна електрофореза, тъй като високото съдържание на цистеини в Ts-PCNTP протеиновата последователност води до подвеждащи резултати, когато се използва техника с  $\text{CH}_3(\text{CH}_2)_{11}\text{OSO}_3\text{Na}$  – полиакриламидна електрофореза. Според данните, получени чрез динамично светлоразсейване, се наблюдава димерната форма на Ts-PCNTP. За определяне на условията за кристализация на белтъка е използвана матрица от произволни условия за израстване на монокристали, подходящи за рентгеноструктурен анализ.