

Enhanced production of short-peptide tagged Ss3a recombinant protein as a novel potential biomarker for strongyloidiasis

Mohd-Hassan, N.H.¹, Noordin, R.¹ and Arifin, N.^{1*}

¹Institute for Research in Molecular Medicine (INFORMM), Universiti Sains Malaysia, 11800 Pulau Pinang, Malaysia

*Corresponding author email: syahida_arifin@usm.my

Received 16 January 2020; received in revised form 10 June 2020; accepted 15 June 2020

Abstract. Strongyloidiasis is a mysterious yet important parasitic disease that is hard to diagnose. While microscopic examination remains a “controversial” gold standard method, improved diagnosis is achieved through confirmatory assays with serological and/or molecular diagnostic approaches. In the current serodiagnosis of strongyloidiasis, recombinant proteins have been adopted in place of the use of native parasite antigens, although the availability of diagnostically potential proteins are still limited. Here, we introduce a novel *Strongyloides* recombinant protein that is uniquely attached to two different short peptide tags as a potential diagnostic biomarker for serodiagnosis of strongyloidiasis, namely lysine (7K) and aspartic acid (7D). The work presented focus on improving the yield and purity of the previously unexpressed recombinant protein. Preliminary diagnostic evaluation of the recombinant favors Ss3a7K protein owing to its higher antigenicity performance with 80% sensitivity and 100% specificity, respectively.

INTRODUCTION

Strongyloidiasis is an increasingly important parasitic disease which affects approximately 370 million people worldwide, mostly the poor’s living in tropical and subtropical countries. While good treatment exists for this infection, the diagnosis of the disease is very challenging, putting strongyloidiasis as “the most neglected of the neglected tropical disease” (Olsen *et al.*, 2009).

Definitive diagnosis of strongyloidiasis is achieved by stool examinations for larvae. However, the fact that most patients are asymptomatic or presented with uncharacterized symptoms mimic other parasitosis, low parasite burden, irregular larvae output and misidentification with other helminths, arguably, these limitations make fecal microscopy a rather controversial gold standard (Arifin *et al.*, 2018). The other

diagnostic methods of molecular detection and serological assays have also been performed; however, the latter had gained primacy as it can achieve high sensitivity, has the potential for point-of-care translation, up-scaling and can be used as a screening tool for post-treatment and epidemiological surveillance.

A serological test, particularly ELISA use the extract of *S. stercoralis* filariform larvae to measure serum antibodies of *Strongyloides* infections. Although somatic antigens are mostly used in making commercial ELISA kits for strongyloidiasis, the preparation of these antigens are often time-consuming and depends on the collection of faeces from heavily infected human or experimental animals. This has led to batch-to-batch variations and impractical for large scale production. In the current serodiagnosis of strongyloidiasis,

recombinant proteins are preferable in place of somatic antigens, although the availability of such diagnostically potential proteins are still limited. Hence, given the potential value of recombinant protein for commercialization, the production of *Strongyloides* recombinant protein would be immensely beneficial to reduce the gap in diagnostics and treatment of strongyloidiasis, two important prongs in control of this disease.

The use of peptide tags in protein recombinant technology has recently become attractive since its first inception 25 years ago (Kapust & Waugh, 1999; Waugh, 2005; Esposito & Chatterjee, 2006). This tag consists of repeated amino acids, polar, positively or negatively charged, and is usually in less than fifteen residues, thereby making it a very small protein that i) does not interfere with the protein structure – hence forward the protein folding, ii) causes less perturbation to the protein function, iii) avoids any additional metabolic burden on the bacterial hosts system and iv) requires no removal of protein tag (Kato *et al.*, 2006). Many studies have showed indicative enhanced solubility of a known low solubility protein, such as the mini body (Bianchi *et al.*, 1994), recombinant *Candida antarctica* lipase B (Jung *et al.*, 2011) and recombinant BabA (Hage *et al.*, 2015).

In our previous study, several cDNA clones were identified from screening of *S. stercoralis* cDNA library with anti-human IgG antibody and sera from patients and non-patients with strongyloidiasis, from which several immunogenic recombinant proteins of high diagnostic value were successively produced. One of these specific *Strongyloides* recombinant proteins is Ss3a, which was found to be highly similar to a hypothetical protein SRAE_2000383300 [*Strongyloides ratti*] with accession number CEF69183.1. Starting with no protein being expressed when the target protein is attached to a histidine tag in pET28 vector, the target gene was later fused to pET42a vector and expressed as GST-tagged protein. The initial stage of evaluation showed that this protein was 100% sensitive (n=14) and 90% specific (n=24) when tested by IgG-phage

immunoblot, a sensitivity and specificity of 100% (n=21) and 95.6% (n=91), respectively by IgG4-western blot and both 100% (n=10) by lateral flow dipstick test (Arifin, 2016). However, due to poor recovery and incomplete proteolytic cleavage inherent in the GST-fused protein, the gene was then inserted to a modified pET32a vector consisting of thioredoxin and seven residues of lysine/aspartic acid attached at the N-terminus, coupled with six residues of histidine at the C-terminus as solubility-enhancing tags. In this work, we produce two variants of diagnostically potential Ss3a recombinant protein which are attached to different short peptide tags namely lysine, otherwise recognised in this study as rSs3a7K and aspartic acid, which is subsequently recognised as rSs3a7D. We describe the steps toward producing good solubility and purity of previously unexpressed recombinant protein. Finally, the preliminary evaluation of the immunogenicity of these variants is also discussed.

MATERIALS AND METHODS

Cloning, expression and purification of recombinant protein

Firstly, the expression of rSs3a7K and rSs3a7D were compared in two different *E. coli* host cells namely BL21(DE3) and C41(DE3), cultured in 2-liter ampicillin-containing TB broth, at 37°C and 200 rpm. Once the optical density (OD₆₀₀) reaches 0.5, IPTG was added, and the temperature was lowered to 28°C. Scouting of the expression parameters was performed by varying the IPTG concentration at 0.2, 0.5, 0.8 and 1mM, and the period of post-induction incubation at 2, 4, 8, and 16 hours. Cells were then harvested by centrifugation at 10,000 x g for 10 minutes (4°C), resuspended in buffer A (50mM NaH₂PO₄, 500mM NaCl, and 10mM imidazole) containing a cocktail of protease inhibitors (Roche Diagnostics) and lysozymes, and then sonicated to lyse the cells (output of 4.5 Hz, 4 cycles, 30 seconds on/off). The final concentration of lysozymes was varied at 0.5mg/ml and 1.5mg/ml.

Clarified and filtered supernatant was loaded onto purification resin at ratio 1:10, whereby the efficacy of two nickel resins was compared namely Qiagen (Qiagen GmbH, Germany), Roche (Roche Diagnostics GmbH, Germany), in addition to one cobalt resin (Thermo Scientific, Waltham, MA, USA). At first, purification was performed by normal washing procedure using 10 column volumes (CV) of washing buffers (50mM NaH₂PO₄, 500mM NaCl) containing 20mM, 30mM and 40mM of imidazole (buffer B, C and D) and 5CV of elution buffer, EB (50mM NaH₂PO₄, 500mM NaCl, and 250mM imidazole). Once the optimal conditions of the above were achieved, the effects of repetitive washings were compared by adding buffer A and B each at 20CV, followed by 10CV of buffer C and D, and finally 5CV of EB.

Western blot and antigenicity evaluation

Eluted fractions containing the target protein were pooled, buffer exchanged with PBS and concentrated until the final volume reached 0.5ml. Protein analysis were performed by Bio Rad assay, 10% SDS-PAGE and anti-histidine western blot. To evaluate the immunoreactivity of the protein variants, western blot was performed in accordance to the general method described by Mohamad *et al.* (2009) using serum samples from patients (n=10) and non-patients (n=10) with strongyloidiasis. The conditions of western blot were determined beforehand by varying; i) antigen amount at 1µg, 2µg, 5µg and 10µg, ii) dilution of primary antibody at the ratio of 1:100 and 1:200, and iii) dilution of secondary antibody (IgG4-HRP) (Calbiochem, San Diego, CA, USA) at 1:2000, 1:4000 and 1:6000. Reactions were developed using SuperSignal® West Pico chemiluminescence substrate (Thermo Scientific, Waltham, MA, USA) in a dark room. The serum samples used in this study were archived and anonymous samples from serum bank at the Institute for Research in Molecular Medicine. Ethical clearance for the use of these samples was obtained from the Human Research Ethics Committee of Universiti Sains Malaysia (USM/JEPeM/17050273).

Cloning and expression of recombinant protein variants

In our experiments, rSs3a7K appeared to have better expression as a soluble protein in C41(DE3) cells than in BL21(DE3) host cells. Meanwhile for rSs3a7D, the BL21(DE3) host cells was shown to have better results than C41(DE3) cells; hence the former was selected. An average of pooled protein concentration was measured, and the two selected host cells for each respective variant showed inclinative support towards the opted parameter. While both variants are best induced with 0.5mM IPTG, no significant difference was observed when the post-induction incubation was prolonged to 16 hours; hence 4 hours was optimal for incubation (Figure 1, 2).

Purification of recombinant protein variants

Though the solubility of a protein can be theoretically calculated and predicted, protein purity is mostly dependent on the right combination of resin, the techniques and composition of buffers used. Since different resins have different binding capacities and strengths, it is therefore desirable to investigate different types of affinity resins at the appropriate resin and lysate ratio.

We found out that Qiagen works best to purify the target protein at the ratio of resin to lysate ratio 1:10, with high yield and modest purity for both variants (Figure 3a, 4a, Appendix B). Given that the target protein was hardly seen on the eluted fraction using two other resins, it is assumed that there was seemingly a weak binding between the target protein with the respective resins, wherein the possibility of target protein was removed during the washing step. Same situation happened when repetitive washing was applied on Qiagen resin (Figure 3c, 4c), as the effect of gradient imidazole concentration and higher column volumes have worsened the binding of low affinity protein to the resin, contributing to protein elution during washing step. In this case, only normal

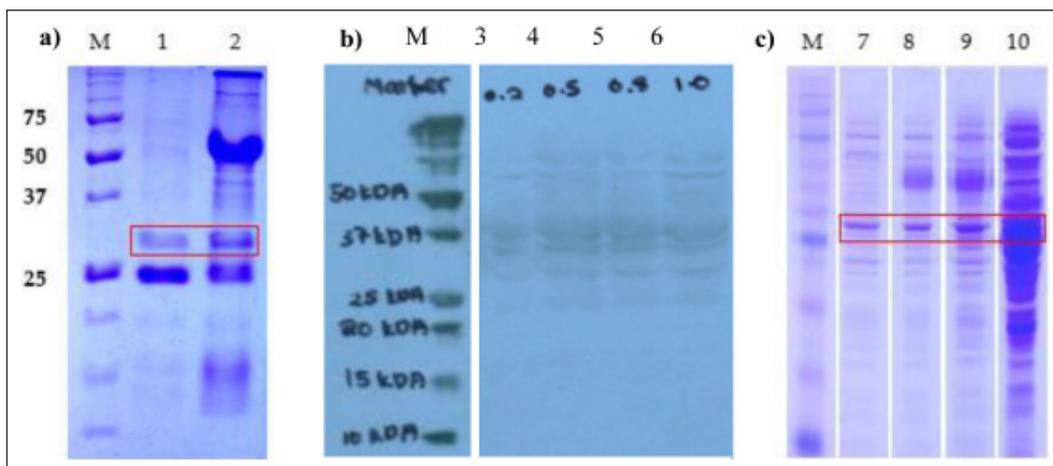


Figure 1. SDS-PAGE and western blot of crude rSs3a7D based on comparison of a) host cells (lane 1-2), b) IPTG concentrations (lane 3-6), and c) post-induction incubation (lane 7-10). Lane M: Marker, 1; rSs3a7D in BL21(DE3), 2; rSs3a7D in C41(DE3), 3; IPTG of 0.2mM, 4; 0.5mM, 5; 0.8mM, 6; 1.0mM, 7; 0 hrs post-induction (pi), 8; 2 hrs pi, 9; 4 hrs pi, and 10; overnight pi.

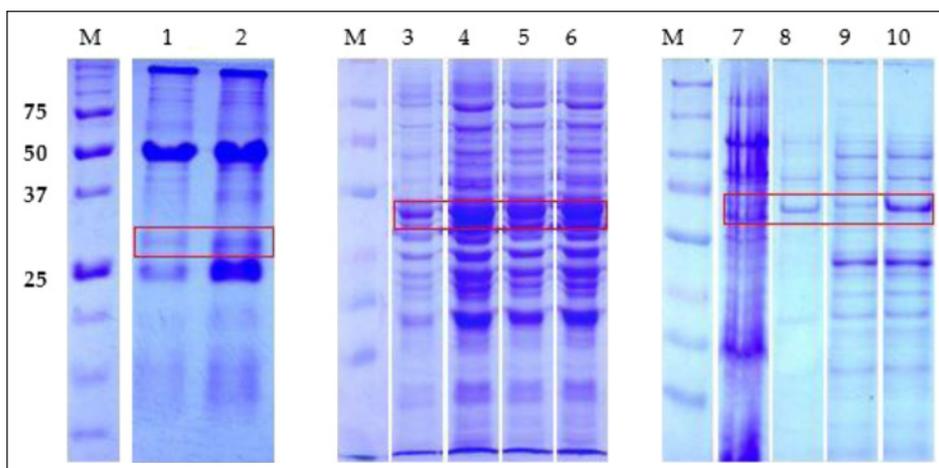


Figure 2. SDS-PAGE gel of crude rSs3a7K based on comparison of a) host cells (lane 1-2), b) IPTG concentrations (lane 3-6), and c) post-induction incubation (lane 7-10). Lane M: Marker, 1; rSs3a7K in BL21(DE3), 2; rSs3a7K in C41(DE3), 3; IPTG of 0.2mM, 4; 0.5mM, 5; 0.8mM, 6; 1.0mM, 7; 0 hrs post-induction (pi), 8; 2 hrs pi, 9; 4 hrs pi, and 10; overnight pi.

washing yielded higher yield at the satisfactory level of purity (Figure 3b, 4b). Therefore, the ideal purification condition was achieved using Qiagen resin at a ratio of 1:10, coupled with normal washing using buffers of 10-40mM imidazole each added at 10CV. The appearance of a specific band at size 34kDa showed a positive reaction when reacted with anti-His-HRP at dilution of 1:1000, confirming the presence of His-

fusion proteins at the expected band of rSs3a7D and rSs3a7K proteins at 34kDa (Figure 5).

In addition to the comparison between rSs3a7D and rSs3a7K, these two variants were also compared with the previous production in its earlier form in pET28 and pET42 vectors. The protein size of the former and the latter were 24.5 kDa and 42 kDa, respectively. When adding the

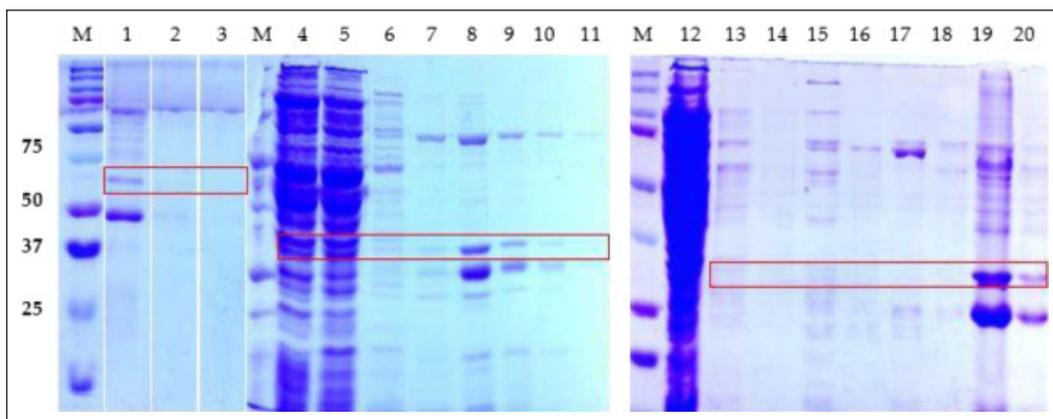


Figure 3. SDS-PAGE gel of rSs3a7D affinity chromatography purification using different a) resins (lane 1-3), b) normal washing strategy (lane 4-11), and c) repetitive washing strategy (lane 12-20). Lane M: Marker, 1; Purified rSs3a7D using Qiagen Superflow Ni-NTA resin, 2; ROCHE cCOMPLETE-His-tag resin, 3; Thermo-Scientific Cobalt resin, 4-11; Purification using normal washing, 12-20; Purification using repetitive washing. [4&12; Flowthrough, 13&14; 10mM; 5, 15&16; 20mM; 6&17; 30mM; 7&18; 40mM; 8-11&19-20; Elution.

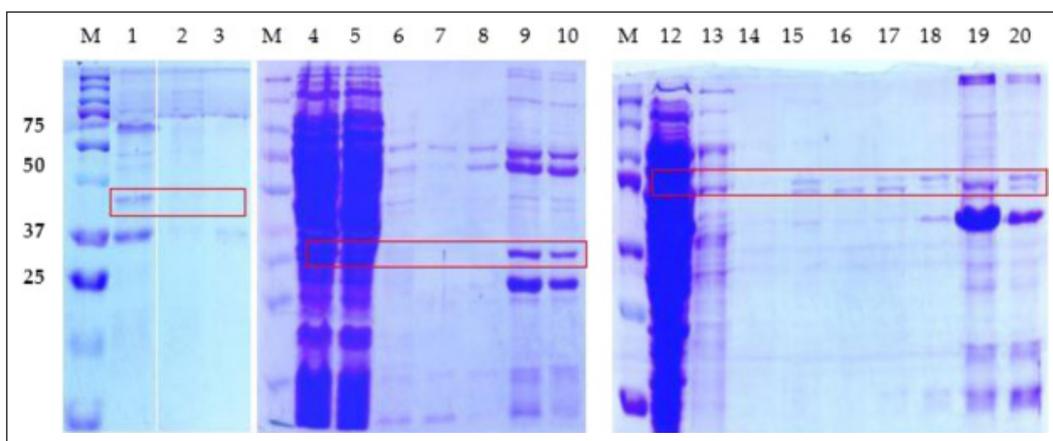


Figure 4. SDS-PAGE gel of rSs3a7K affinity chromatography purification using different a) resins (lane 1-3), b) normal washing strategy (lane 4-10), and c) repetitive washing strategy (lane 11-20). Lane M: Marker, 1; Purified rSs3a7K using Qiagen Superflow Ni-NTA resin, 2; ROCHE cCOMPLETE-His-tag resin, 3; Thermo-Scientific Cobalt resin, 4-10; Purification using normal washing, 12-20; Purification using repetitive washing. [4&12; Flowthrough, 5,13&14; 10mM; 6, 15&16; 20mM; 7&17; 30mM; 8, 18; 40mM; 9-10&19-20; Elution.

combinational tags in pET32 vector, the protein sizes of rSs3a7D and rSs3a7K changed to 34kDa, with *pI* of 4.8. Although rSs3a7D was slightly positively charged than rSs3a7K due to a higher protonation factor, the *pI* however remained the same for both variants because the added amino acid residues were sufficient to enhance the solubility of the original clone but not enough

to alter the *pI* of the recombinants, which was largely reliant on the pre-existing clone. In addition to the appearance of the distinct target bands on western blot film (Figure 5), the yield of the new protein variants increased significantly to 4-fold, giving an average of 1 to 2 mg per liter culture, thus affirming the advantage of peptide tag as solubilizing protein tag.

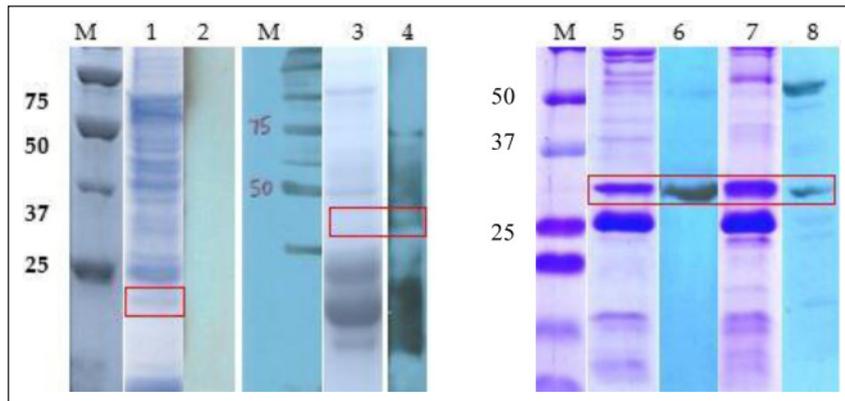


Figure 5. Comparison of the purity of rSs3a from different batches of purification and variants. Samples were loaded at 10ug/lane. Lane M: Marker; 1; Non-optimized (original) rSs3a-His, 2; anti-His western blot of rSs3a-His, 3; rSs3a-GST, 4; anti-GST western blot of rSs3a-GST, 5; rSs3a7D, 6; anti-His western blot of rSs3a7D, 7; rSs3a7K, 8; anti-His western blot of rSs3a7K.

Antigenicity evaluation

Further, preliminary evaluation of the immunogenicity of rSs3a7D and rSs3a7K were performed by western blot using patient and non-patient serum samples with strongyloidiasis, in combination with anti-human IgG4-HRP as the secondary antibody. This secondary conjugate was chosen as it is an important marker of chronic infection and is highly specific for detecting helminth infections, including strongyloidiasis. While other antibodies such as IgG and IgA are also important, the former is less specific and the latter is more abundant in saliva (Norsyahida *et al.*, 2013). Earlier, the parameters of IgG4-western blot, i.e. antigen amount, primary and secondary antibodies were optimized to obtain the best discrimination between positive and negative results.

At the protein amount of 2µg, serum dilution of 1:100 (overnight, 4°C), and anti-human IgG4 antibody at 1:6000 (2 hours, RT) the sensitivity and specificity of rSs3a7D resulted in 90% (n=10) and 80% (n=10) respectively. On the other hand, at the protein amount of 1µg, serum dilution at 1:100 (overnight, 4°C), and anti-human IgG4 antibody at 1:2000 (1-hour, RT), the sensitivity and specificity of rSs3a7K were found to be 80% (n=10) and 100% (n=10), respectively. Other parameters such as blocking incubation period, blocking agent,

concentration of Tween-20 in washing buffer (TBS) and temperature of incubation did not give any significant changes to the antigenicity of neither rSs3a7D nor rSs3a7K, but rather assisted in minimizing the high background and removing contaminants. The blocking period was fixed to 1 hour using Superblocking Solution (Thermo Scientific, Waltham, MA, USA) whereas the concentration of Tween 20 was fixed to 0.05% diluted in 1xTBS.

While recombinant rSs3a7D was shown to have a higher advantage in the perspective of purity as analyzed by SDS-PAGE, the antigenicity performance of the protein was marked to be lower than recombinant rSs3a7K. rSs3a7D has screened out nine out of ten (90%) of infected human serum samples, indicated by the absence of band at strip 1 (Figure 6a) whereas two out of ten individual negative serum samples contributed as false positives, indicated by the bands at strip 12 and 18 (Figure 6b). On the contrary, rSs3a7K has successfully screened out eight out of ten (80%) infected serum sample, as specified by the false negative bands from strip 1 and 9 (Figure 7a), this variant has also achieved ten out of ten (100%) specificity (Figure 7b). Meanwhile, the sensitivity and specificity of IgG4 crude protein-based immunoassay were previously reported to be 76.9% (n=26) and 92.7% (n=55),

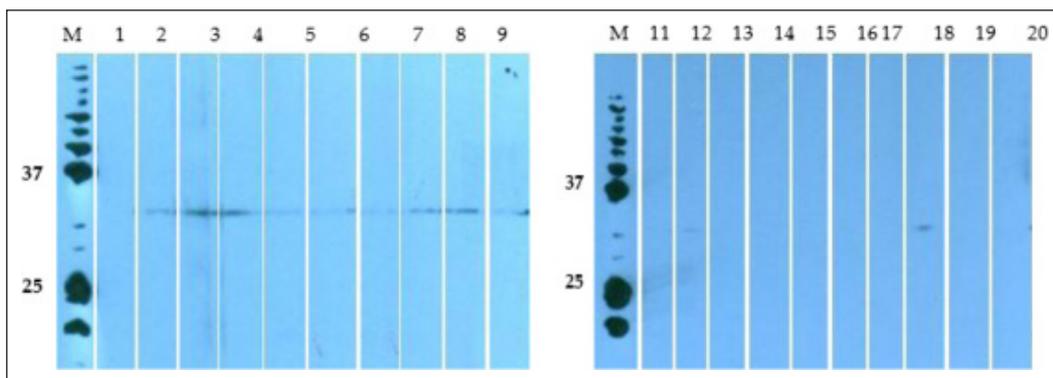


Figure 6. Western blot preliminary evaluation of rSs3a7D using ten human individual serum samples. a) 1-10 using individual positive serum samples, b) 11-20 using individual negative serum samples.

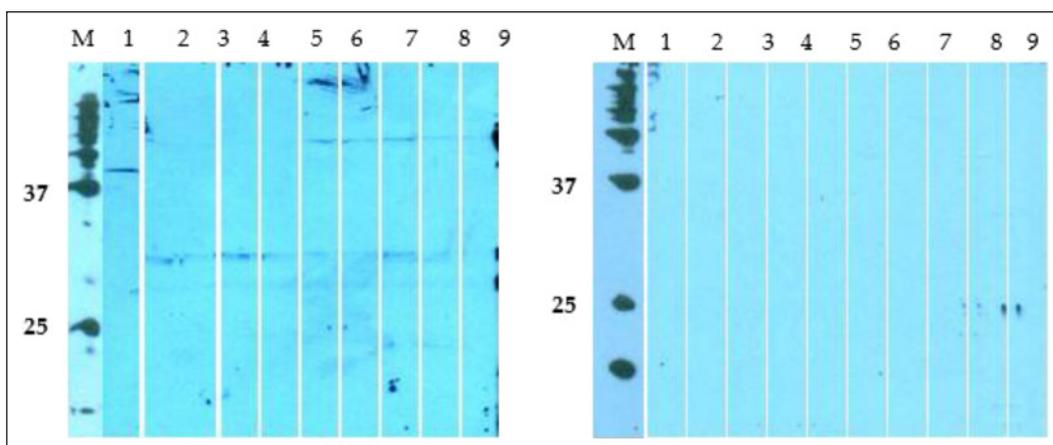


Figure 7. Western blot preliminary evaluation of rSs3a7K using ten human individual serum samples. a) 1-10 using individual positive serum samples b) 11-20 using individual negative serum samples.

respectively (Norsyahida *et al.*, 2013). Although the number of serum samples tested in this study was small, this study highlighted the good potential of *Strongyloides* recombinant protein as an antigen compared with the use of crude parasite antigen in immunoassay.

DISCUSSION

Previously, the Ss3a was hardly expressed as histidine-fused protein in pET28a vector and had low solubility when fused to GST tag in pET42a vector. In an attempt to enhance the production of previously unexpressed

protein while ensuring it is stably folded, proper selection of expression host and scouting on the expression parameters are essential in that process. While these physical factors usually have a great impact on protein solubility, the intrinsic factor such as the introduction of protein solubilizing tags also made a substantial difference. As the solubilizing effect of the protein tags seems to rely on the tags' correct folding, hence small protein tags that can exert solubility-enhancing effect without causing aggregation facilitates purification, and display active function are of interest worth investigation.

Peptide tags, i.e. lysine and arginine, are known to show preference as solubilizing tags due to their hydrophilic character, for which the higher charges of the longer peptides has bigger solubilizing effect. The enhancement of the solubility was attributed to the repulsive electrostatic interactions between similarly charged tags, which prevent aggregation and allow sufficient time for correct folding. Other than length, the type of amino acid (anionic or cationic tags) and the position where it is tagged (in the N- or C-terminus) are also reportedly seemed to affect protein solubility strongly. In our case, the gene coded for Ss3a protein was inserted in a modified pET32a vector consisted of 6x histidine, thioredoxin and seven residues of lysine (7K) and seven residues of aspartic acid (7D) at the N-terminus separately.

Originally, the molecular masses of rSs3a-His and rSs3a-GST were estimated to be 24.5 kDa, and 42 kDa, respectively. When adding the combinational tags, the protein size changed to 34kDa, with *pI* of 4.8 for both variants as analyzed. rSs3a7D comprised seven residues of aspartic acid hence contributing to an acidic property to the recombinant, whereas in contrast, rSs3a7K comprised seven residues of lysine, which express the essential basic properties. Taking this into account, while both variants possess hydrophilic properties, rSs3a7K is otherwise slightly more negatively charged as compared to rSs3a7D. This potentially led to affecting the chemical composition and the interaction of the recombinant protein, therefore constituting to different antigenic performances between the two variants. Not only did rSs3a7K showed better antigenicity, individual sera reacted specifically towards the target bands at 34kDa as expected which resulted in a clean profile. The most striking difference between these two variants is the protein stability, as we observed that rSs3a7K demonstrated better protein stability and consistent in data collection. Based on predictive patterns and the antigenicity performance by both variants, rSs3a7K, with its optimized condition in expression and purification was chosen as optimal and shall be proceeded with future work. As our current work is geared on proving the clone's potential

in serodiagnosis, a preliminary evaluation is sufficient to give us a lead.

In our previous study, we have produced rSs3a protein tagged with GST with substantial diagnostic value when tested by phage immunoblot, western blot and lateral flow immunochromatography dipstick test. The intrinsic complexity of GST tag has led to the use of histidine as protein tag of choice that befits diagnostic application. To further improve yield and purity of this protein, there are two important provisos to the methods and strategies implemented in this work. First, the incorporation of short peptide sequences to the original clone, which encapsulates the addition of seven lysine residues to the N-terminus of rSs3a7K protein, and seven aspartic acid residues to the N-terminal of rSs3a7D, in addition of two other solubility-enhancing tags (i.e. thioredoxin and histidine), and second, the optimization of the physical parameters of the protein expression and purification processes. Recombinant Ss3a7K was chosen as the optimal variant to be used in the development of lateral flow test as a proof-of-concept and other future potential works. When compared to the one without peptide tags, we have successively increased the yield of rSs3a-7K protein for up to 4-fold, with improved purity. However, due to limited number and volume of the available serum samples, we are not able to compare the antigenicity of this protein as many as we did previously. Nevertheless, evaluation of the antigenicity of rSs3a7K using limited serum samples shows promise as a good diagnostic biomarker for improved diagnosis of strongyloidiasis. Extensive efforts are needed to evaluate the efficacy of this protein for diagnostic application further.

Acknowledgements. We greatly appreciate the funding we received from Universiti Sains Malaysia (Short-term Grant num 304.CIPPM.6315071). Part of this work was supported by funding from the Malaysian Ministry of Education through the Higher Institution Centre of Excellence (HiCoE) Program (Grant No. 311/CIPPM/4401005). NHMH is funded by Graduate Assistant Scheme, USM.

REFERENCES

- Arifin, N. (2016). Serodiagnosis of strongyloidiasis: Identification of cDNA clones, production of recombinant antigens and immunoassay development. Thesis (Ph.D). Universiti Sains Malaysia.
- Arifin, N., Yunus, M.H., Nolan, T.J., Lok, J.B. & Noordin, R. (2018). Identification and preliminary evaluation of a novel recombinant protein for serodiagnosis of strongyloidiasis. *American Journal of Tropical Medicine and Hygiene* **98**: 1165-1170.
- Bianchi, E., Venturini, S., Pessi, A., Tramontano, A. & Sollazzo, M. (1994). High level expression and rational mutagenesis of a designed protein, the minibody: From an insoluble to a soluble molecule. *Journal of Molecular Biology* **236**: 649-659.
- Esposito, D. & Chatterjee, D.K. (2006). Enhancement of soluble protein expression through the use of fusion tags. *Current Opinion in Biotechnology* **17**: 353-358.
- Hage, N., Renshaw, J.G., Winkler, G.S., Gellert, P., Stolnik, S. & Falcone, F.H. (2015). Improved expression and purification of the *Helicobacter pylori* adhesin BabA through the incorporation of a hexa-lysine tag. *Protein Expression and Purification* **106**: 25-30.
- Jung, H.J., Kim, S.K., Min, W.K., Lee, S.S., Park, K., Park, Y.C. & Seo, J.H. (2011). Polycationic amino acid tags enhance soluble expression of *Candida antarctica* lipase B in recombinant *Escherichia coli*. *Bioprocess and Biosystems Engineering* **34**: 833-839.
- Kapust, R.B. & Waugh, D.S. (1999). *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science* **8**: 1668-1674.
- Kato, A., Maki, K., Ebina, T., Kuwajima, K., Soda, K. & Kuroda, Y. (2006). Mutational Analysis of Protein Solubility Enhancement Using Short Peptide Tags. *Biopolymers* **85**: 12-18.
- Mohamad, S., Azmi, N.C. & Noordin, R. (2009). Development and evaluation of a sensitive and specific assay for diagnosis of human toxocariasis by use of three recombinant antigens (TES-26, TES-30USM, and TES-120). *Journal of Clinical Microbiology* **47**: 61712-1717.
- Norsyahida, A., Mehdi, R., Sadjjadi, M.S., Yunus, M.H., Low, H.C., Zeehaida, M. & Noordin, R. (2013). Laboratory detection of strongyloidiasis: IgG, IgG4 and IgEELISAs and cross reactivity with lymphatic filariasis. *Parasite Immunology* **35**: 5-6 174-179.
- Olsen, A., Van Lieshout, L., Marti, H., Polderman, T., Polman, K., Steinmann, P., Stothard, R., Thybo, S., Verweij, J.J. & Magnussen, P. (2009). Strongyloidiasis- the most neglected of the neglected tropical diseases? *Transactions of the Royal Society of Tropical Medicine and Hygiene* **103**: 967-972.
- Waugh, D.S. (2005). Making the most of affinity tags. *Trends in Biotechnology* **23**: 316-320.