Makara Journal of Science

Volume 14 | Issue 2 Article 16

11-25-2010

TROUBLESHOOTING IN EXPRESSION AND PURIFICATION OF RECOMBINANT SEVERE ACUTE RESPIRATORY SYNDROME-ASSOCIATED CORONAVIRUS NUCLEOCAPSID PROTEIN IN Escherichia coli BL21

Andi Yasmon
Department of Microbiology, Faculty of Medicine, University of Indonesia, Jakarta 10320, Indonesia, andiyasmon@yahoo.com

Fera Ibrahim
Department of Microbiology, Faculty of Medicine, University of Indonesia, Jakarta 10320, Indonesia

Budiman Bela
Department of Microbiology, Faculty of Medicine, University of Indonesia, Jakarta 10320, Indonesia

MAKARA, SAINS, VOL. 14, NO. 2, NOVEMBER 2010: 140-144

TROUBLESHOOTING IN EXPRESSION AND PURIFICATION OF RECOMBINANT SEVERE ACUTE RESPIRATORY SYNDROME-ASSOCIATED CORONAVIRUS NUCLEOCAPSID PROTEIN IN *Escherichia coli* BL21

Andi Yasmon[*)], Fera Ibrahim, and Budiman Bela

Department of Microbiology, Faculty of Medicine, University of Indonesia, Jakarta 10320, Indonesia

[*)]*E-mail: andiyasmon@yahoo.com*

Abstract

Considering importance of N protein for study of viral pathogenesis or development of immunodiagnostic assay, we reported effects of several conditions on purity and homogeneity of recombinant SARS-CoV N protein expressed in *E. coli* BL21. The SARS-CoV N gene was reverse transcribed and amplified by the reverse transcription-polymerase chain reaction (RT-PCR) technique. The amplicons were cloned into pGEX-6P1 and followed by subcloning of the target gene into pQE-80L. After inserting the recombinant plasmid (pQE80-N) into *E. coli*, the recombinant protein (6 x His tag-N protein fusion) was expressed by inducing the bacterial cells with 0.1-0.5 mM isopropyl-1-thio-D-galactopyranoside (IPTG) for 1-5 h. The protein recombinant were extracted from the bacterial cells by NTT buffer containing 0-20 mM imidazol, and followed by Ni-NTA affinity resin purification. The results showed that induction of *E. coli* BL21 with 0.2 mM IPTG for 4 h and followed with lysis of bacterial cells in NTT buffer containing 10 mM imidazol were optimal conditions to obtain the pure recombinant SARS-CoV N protein.

Keywords: imidazol, IPTG, N-lauroylsarcosine, Triton X, SARS-CoV N protein

1. Introduction

Genome of severe acute respiratory syndrome coronavirus (SARS-CoV) is approximately 30 kb with a structural gene arrangement: 5'- Replicase (orf 1a)-Protease (orf 1b)-Spike (S)-Envelope (E)-membrane (M)-Nucleocapsid (N)-3' [1]. Among the SARS-CoV proteins, N protein constitutes an unique protein because of having a putative nuclear localization signal (KKDKKKK, a.a. 370–376), playing a role in replication and transcription of viral RNA [2]. Moreover, N protein is highly immunogenic and earliest expressed, and N protein-anti Abs are most abundant circulated during infection [3-6]; therefore, several groups developed the SARS-CoV N protein-based serological assays for SARS diagnosis [7-11].

In this study, we report several parameters technically affecting on the purity of the purification of the recombinant SARS-CoV N protein expressed in *E. coli* BL21. The information is useful in the use of the recombinant SARS-CoV N protein to develop immunodiagnostic assay or to study viral pathogenesis.

2. Methods

Viral RNA standard and SARS positive sera. SARS-CoV genome were obtained from Matthias Niedrig, Robert Koch-Institut Berlin and provided by the European Network for Diagnostics of "imported' Viral Diseases (ENIVD).

Cloning, Expression and purification of recombinant SARS-CoV N protein. Reverse transcription and PCR were performed with primers: forward (5'-GC*G GATCC*ATGTCTGATAATGGACCCCAA-3') and reverse (5'-AC*GTCGAC*TTATGCCTGAGTTGAATC-3'). N gene was ligated into *Bam*HI and *Sal*I sites of prokaryotic expression vector pGEX-6P1 (Amersham Pharmacia) and followed by subcloning the gene into vector pQE-80L (Qiagen). The fidelity of the insert sequence and in frame between the insert and vector DNA sequences was confirmed by the sequence of recombinant DNA. The recombinant plasmid (pQE80-N) was transformed into *E. coli* strain BL21. The recombinant protein (6 x His tag-N protein fusion) was expressed by inducing the bacterial cells with 0.1-0.5 mM

isopropyl-1-thio-D-galactopyranoside (IPTG) at 25 °C or 37 °C for 1-5 h. Bacterial pellets were sonicated and solubilized in lysis buffer from kit (50 mM NaH2PO4; 300 mM NaCl; 10 mM imidazole, pH 8.0 [Qiagen]) or in NTT buffer (1.5% N-lauroylsarcosine, 1% Triton X-100, 150mM NaCl, and 10mM Tris, pH 8.0) containing 0-20 mM imidazol. Finally, purification of the recombinant protein was performed with following the manufacturer's instructions (Qiagen).

Sodium dodecyl sulfate–polyacrylamide gel (SDS-PAGE). A 12% SDS-PAGE with gel thickness of 0.75 mm was consisted of separating and stacking gels. Compositions of the gels were made in accordance with Sambrook [12]. Electrophoresis of proteins on SDS-PAGE was performed at 150 mV for 40 min. Gels were stained with coommasie blue (50% methanol (Merck], 0,05% coommasie brilliant blue G-250 [Sigma], 10% asam asetat [Merck], and 40% H_2O) for 10 min followed by destaining with destaining solution (5% methanol [Merck], 7% asam asetat [Merck], 88% H_2O) for 16-18 h.

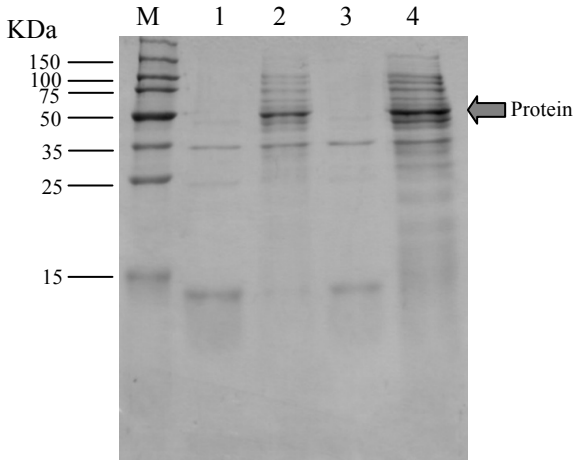

Figure 1. Analysis of Inclusion Body from *E. coli* Cells Induced at 25 °C and 30 °C. M: Protein Ladder. Lane 1 and 3: Supernatant Phases of the Cells Induced at 37 °C and 25 °C. Lane 2 and 4: Pellet Phase of the Cells Induced at 37 °C and 25 °C.

3. Results and Discussion

Cloning of SARS-CoV N gene into pQE80L. Sequencing of the gene nucleotide was performed to confirm the exactness of SARS-CoV N nucleotide sequence and its in-frame with pQE80L. Based on the analysis, the nucleotide sequence showed 100% homology with the N gene of SARS-CoV strain Urbani (Genebank accession number: AAP13445) and was truly in-framed with pQE80L (data not shown).

Optimal time for expression of recombinant SARS-CoV N protein. To express protein, the pQE80-N was transformed into *E. coli* BL21 following protocol reported by Sambrook [12]. In order to know optimal time for expression of the recombinant protein, the cells containing the pQE80-N was induced with 1 mM at 37 °C for 1-5 h. Since there was no significant difference between protein bands induced for 4 and 5 h (data not shown); thus, the optimal induction time in expressing the protein is for 4 h.

Analysis of inclusion body. A formed possible inclusion body was analyzed by means of detecting the expected protein in both supernatant and pellet phases. The result showed that the expected protein was detected in pellet phase (Figure 1, lane 2 and 4) but not detected in supernatant phase (Figure 1, lane 1 and 3); indicating that the recombinant protein formed an inclusion body when it was expressed in *E. coli* BL21. Induction temperatures at 25 °C and 37 °C were also compared to solve the inclusion body (Figure 1). However, there were no effects of induction temperatures at 25 °C and 37 °C (Figure 1), but intensity of protein band from *E. coli* BL21 induced at 25 °C showed stronger intensity (Figure 1, lane 4); thus, the optimal induction temperature was 25 °C.

Since difference of visually protein bands in supernatant phases induced at 25 °C and 37 °C was insignificance (Figure 1, lane 1 and 3), we analyzed effect of IPTG concentrations on solubility of the recombinant protein. The result showed that the recombinant protein was more dominant in supernatant (Figure 2, lane 2s) than in pellet phases (Figure 2, lane 2p) when *E. coli* BL21 containing pQE80-N was induced with 0.2 mM IPTG. In contrast, the bacterial cells induced with 0.1, 0.3-0.5 mM IPTG showed dominant protein in pellet phases (Figure 2); therefore, we decide 0.2 mM IPTG as optimal concentration to make the recombinant protein become more soluble.

The supernatant and pellet phases showed in Figure 1-2 were obtained from bacterial cells that were lysed in lysis buffer from kit (Qiagen). To get a better buffer, we compared between the lysis buffer from kit and NTT buffer. This comparison study showed that the recombinant protein was dominantly detected in supernatant phase when the cells were lysed in NTT buffer (Figure 3, lane Ns). While cells lysed in the lysis buffer from kit showed the recombinant protein that was dominant in pellet phase (Figure 3, lane Qp). Based on the results, NTT buffer was better solution to solve inclusion bodies.

Purification of recombinant SARS-CoV N protein. To obtain pure recombinant proteins, the proteins were purified by Ni-NTA affinity resin. In purification procedure, washing and elution steps were performed five and four times, respectively. Result of each elution

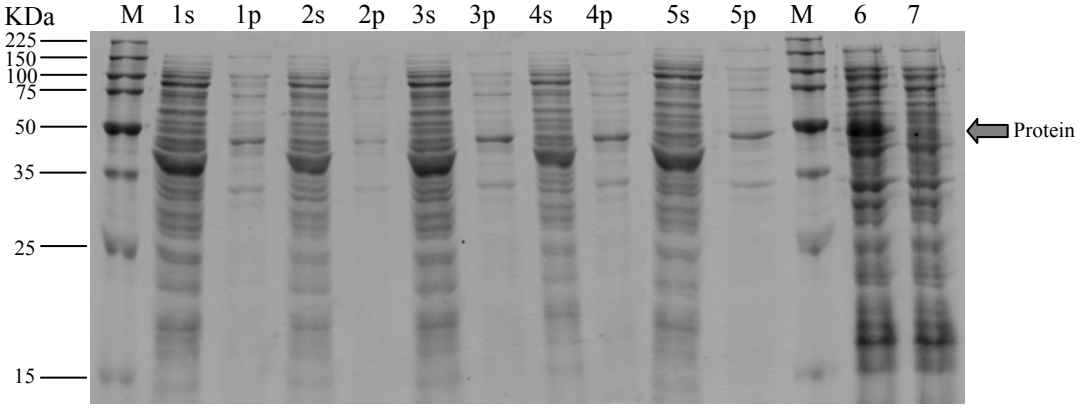

Figure 2. Analysis of Effect of Variable IPTG Concentrations on Solubility of the Recombinant Protein. M: Protein Ladder. Lane 1s-5s: Supernatant Phases of *E. coli* Cells Induced with 0.1-0.5 mM IPTG. Lane 1p-5p: Pellet Phases of *E. coli* Cells Induced with 0.1-0.5 mM IPTG. Lane 6 and 7: Cell Extract Controls Expressing Recombinant Protein and no Recombinant Protein, Respectively. KDa. Kilo Dalton

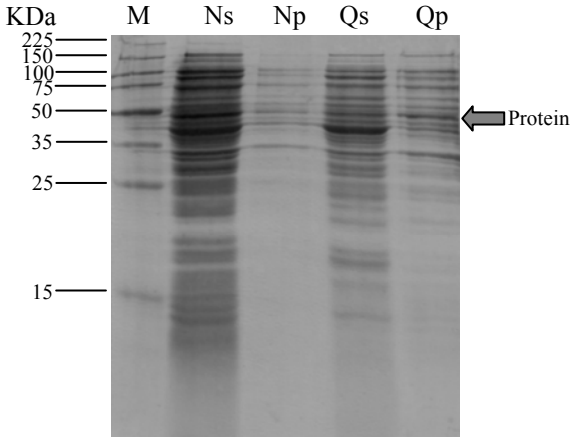

Figure 3. Result of Comparison between the Lysis Buffer from Kit and NTT Buffer to Solve Inclusion Bodies. M: Protein Ladder. Lane Ns and Qs: Supernatant Phases of Bacterial Cells Lysed in the NTT Buffer and Lysis Buffer from Kit, Respectively. Lane Np and Qp: Pellet Phases of Bacterial Cells Lysed the NTT Buffer and Lysis Buffer from Kit, Respectively. KDa: Kilo Dalton

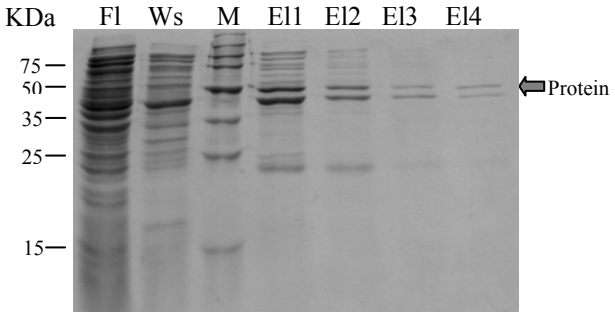

Figure 4. Result of Purification of the Recombinant Protein by Ni-NTA Affinity Resin. Lane F1: Elute of Flowthrough Step. Lane Ws: Elute of Washing Step. M: Protein Ladder. Lane El1-4: Elutes of Elution Step 1-4. KDa: Kilo Dalton

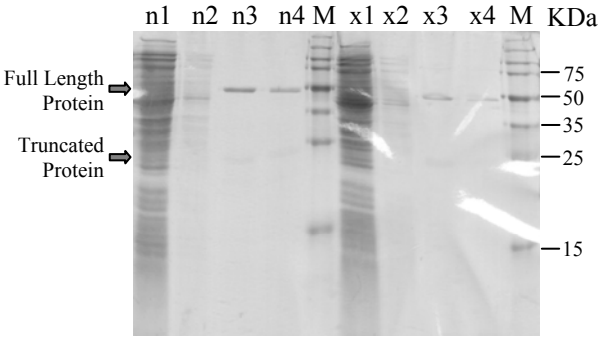

Figure 5. Effect of Adding 10 mM or 20 mM Imidazol into NTT Buffer on Purity of Purification Result of the Recombinant Protein and on Efficiency of the Recombinant Protein Purification. n: NTT Buffer Containing 10 mM Imidazol. x: NTT Buffer Containing 20 mM Imidazol. 1-4: Elutes of Flowthrough, Washing, First Elution, and Second Elution Steps, Respectively. M: Protein Ladder. KDa: Kilo Dalton

was analyzed on SDS-PAGE for knowing which elute of four elutes contains the purer protein. The result of the analysis showed that all elutes contained expected recombinant proteins as well as bacterial proteins (Figure 4, lane EL1-4).

Because of the present of the contaminated protein from *E. coli* (Figure 4), we analyzed effects of adding 10 mM or 20 mM imidazol into NTT buffer on purity of purification result of expected recombinant protein. The results showed that the recombinant protein was purer when bacterial cells were lysed in NTT buffer containing 10 or 20 mM imidazol (Figure 5, lane n3 and n4), but NTT buffer containing 20 mM resulted weak density of protein bands (Figure 5, lane x3 and x4). Therefore, we decided that NTT buffer containing 10 mM imidazol was better solution for extracting of the

recombinant N SARS-CoV protein expressed in *E. coli* BL21.

In this study, gene encoding N protein of SARS-CoV has been cloned into pQE-80L. The plasmid can be used as expression vector in all strains of *E. coli* cells including strain BL21 [13]. Thus, recombinant SARS-CoV N protein has been expressed in *E. coli* BL21 and further purified by using Ni-NTA affinity resin. In expression and purification, we found one main problem i.e. the recombinant proteins formed inclusion body (Figure 1). The inclusion body is a protein mass caused by hydrophobic interactions among the proteins during stages in forming protein secondary structures [13]. Several works reported that soluble level of particular recombinant protein expressed in *E. coli* cells can be increased by means of decreasing growth temperature (from 37 ^{o}C to 20 ^{o}C) when the bacterial cells were induced with IPTG [13-15]. However, we found no significant difference of the soluble level of the recombinant protein expressed in bacteria that was induced at 37 ^{o}C and 25 ^{o}C (Figure 1). It is thought that the soluble level of the protein is unable to be detected visually because the quality of the solubility is involved in correct folding of certain protein while forming secondary structure.

In addition to temperature treatment, IPTG induction at particular concentration also influence the protein solubility [13]. In the present study, bacteria induced at 0.2 mM IPTG showed a protein band with less strong intensity in supernatant than in pellet phases. It means that induction at 0.2 mM IPTG lead the recombinant protein to more soluble but not optimal solubility (Figure 2). To achieve the optimal protein solubility, we used NTT buffer to lyse the bacterial cells instead of the lysis buffer from kit. The NTT buffer was firstly reported by Chang and colleagues to lyse *E. coli* strain m15 expressing recombinant SARS-CoV N protein [2]. They found that the recombinant SARS-CoV N protein-anti polyclonal antibody recognized native SARS-CoV N protein on Vero E6 cell line that was infected by SARS-CoV [2]. According to their findings, it can be concluded that the NTT buffer can be used for lysis of *E. coli* cells expressing the recombinant SARS-CoV N protein without eliminating the native protein features.

Moreover, the lysis buffer containing N-lauroylsarcosine and Triton X-100 with low concentrations (≤2%) can be used for solving the inclusion bodies without eliminating native protein conformation [12,16,17]. Sambrook reported that 1.5% N-lauroylsarcosine and 1% Triton X-100 can be used to solve glutathine-S-transferase (GST)-fused proteins without interfering interactions between GST domains and glutathione-agarose resins in which the interactions are a native protein conformation-dependent interaction [12]. The evidences have also been proven previously in another study [18].

Even though the solubility of the recombinant protein showed a better level when NTT buffer used as the lysis buffer, there were still protein contaminations from *E. coli* BL2 (Figure 4). To eliminate protein contaminations, we added 10 mM or 20 mM imidazol into NTT buffer. Principle of the addition of the imidazol is that the lysis buffer containing 10-20 mM imidazol can reduce unspecific interactions between Ni-NTA resins and unexpected proteins; however, high imidazol concentration lead to dissociate specific interactions [13].

A comparison study showed that NTT buffer containing 10 mM imidazol (NTT + 10 mM imidazol) yielded the recombinant proteins that were purer than NTT buffer containing 20 mM imidazol (Figure 5). In addition, the NTT + 10 mM imidazol also led the recombinant protein bands to higher intensity (Figure 5). Thus, 10 mM imidazol added into NTT buffer can increase the purity of the recombinant protein and is more efficient for purification of the recombinant protein compared with 20 mM imidazol or without imidazol. As shown in Figure 5 that there were truncated proteins detected on SDS-PAGE gel. The proteins were not contaminant proteins from *E. coli* BL21 but truncated SARS-CoV N proteins. We have proved the evidences by Western blot and enzyme-linked immusorbent assay (data not shown).

In conclusion, optimal expression and purification of recombinant SARS-CoV N protein in *E. coli* BL21 are conducted with the following two important conditions: bacterial cells were induced at 0.2 mM IPTG for 4 h and lysed in NTT buffer (1.5% N-lauroylsarcosine, 1% Triton X-100, 150mM NaCl, and 10mM Tris, pH 8.0) containing 10 mM imidazol.

Acknowledgement

This research was funded by Ministry of Research and Technology, Indonesia. We are grateful to Matthias Niedrig, Robert Koch-Institut Berlin for providing the SARS-CoV genome.

References

[1] D.A. Groneberg, R. Hilgenfeld, P. Zabel, Respir. Res. 6 (2005) 8.

[2] M.S. Chang, Y.T. Lu, S.T. Ho, C.C. Wu, T.Y. Wei, C.J. Chen, Y.T. Hsu, P.C. Chu, C.H. Chen, J.M. Chu, Y.L. Jan, C.C. Hung, C.C. Fan, Y.C. Yang, Biochem. Biophys. Res. Commun. 314 (2004) 931.

[3] X.Y. Che, W. Hao, Y. Wang, B. Di, K. Yin, Y.C. Xu, C.S. Feng, Z.Y. Wan, V.C. Cheng, K.Y. Yuen, Emerg. Infect. Dis. 10 (2004) 1947.
[4] Y. He, Y. Zhou, H. Wu, Z. Kou, S. Liu, S. Jiang. J. Clin. Microbiol. 42 (2004) 5309.
[5] Y.H. Li, J. Li, X.E. Liu, L. Wang, T. Li, Y.H. Zhou, H. Zhuang, J. Virol. Methods 130 (2005) 45.
[6] B. Shang, X.Y. Wang, J.W. Yuan, A. Vabret, X.D. Wu, R.F. Yang, L. Tian, Y.Y. Ji, V. Deubel, B. Sun, Biochem. Biophys. Res. Commun. 336 (2005) 110.
[7] Q. He, I. Manopo, L. Lu, B.P. Leung, H.H. Chng, A.E. Ling, L.L. Chee, S.W. Chan, E.E. Ooi, Y.L. Sin, B. Ang, J. Kwang, Clin. Diagn. Lab. Immunol. 12 (2005) 321.
[8] M. Saijo, T. Ogino, F. Taguchi, S. Fukushi, T. Mizutani, T. Notomi, H. Kanda, H. Minekawa, S. Matsuyama, H.T. Long, N.T. Hanh, I. Kurane, M. Tashiro, S. Morikawa, J. Virol. Methods 125 (2005) 181.
[9] Y. Shi, Y. Yi, P. Li, T. Kuang, L. Li, M. Dong, Q. Ma, C. Cao. J. Clin. Microbiol. 41 (2003) 5781.
[10] P.C. Woo, S.K. Lau, B.H. Wong, H.W. Tsoi, A.M. Fung, K.H. Chan, V.K. Tam, J.S. Peiris, K.Y. Yuen, J. Clin. Microbiol. 42 (2004) 2306.
[11] P.C. Woo, S.K. Lau, B.H. Wong, H.W. Tsoi, A.M. Fung, R.Y. Kao, K.H. Chan, J.S. Peiris, K.Y. Yuen, J. Clin. Microbiol. 43 (2005) 3054.
[12] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning, A Laboratory Manual, Vol 3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, p.15.113.
[13] Qiagen, The QIA*expressionist*. A Handbook for High-level Expression and Purification of 6xHis-tagged Proteins, 2003, p.126
[14] M. Martinez-Alonso, N. Gonzalez-Montalban, E. Garcia-Fruitos, A. Villaverde, Microb. Cell Fact. 8 (2009) 4.
[15] R. Rudolph, H. Lilie, In Vitro Folding of Inclusion Body Proteins, FASEB J 10 (1996) 49.
[16] F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl (Eds), Current Protocols in Molecular Biology, Vol. 2, Greene Publishing Associates and Wiley-Interscience, New York, 1990, p.16.13.7.
[17] T. Hackstadt, J. Bacteriol. 173 (1991) 7046.
[18] A. Yasmon, Tesis Magister, Ilmu Biomedik, Fakultas Kedokteran Universitas Indonesia, Indonesia, 2005.