Cloning and expression of nucleoprotein of influenza virus in *Escherichia coli*.

การโคลนนิ่งและการแสดงออกของ nucleoprotein ของไวรัสไข้หวัดใหญ่ใน *Escherichia coli*

Amornrat Sianglum (อมรรัตน์ เสียงล้ำ)* Dr. Wises Namwat (ดร.วิเศษ นามวาท)**
Dr. Sorujsiri Chareonsudjai (ดร.โสรัจสิริ เจริญสุดใจ)*** Dr.Viraphong Lulitanond (ดร.วีระพงศ์ ลุลิตานนท์)****

ABSTRACT

Influenza virus cause extensive damage to life, economic and populations throughout the world. Although there are available vaccines against influenza, the vaccine need to be changed each year according to the specific subtype of influenza virus circulated at that time as the protective immunity raised by influenza vaccine is subtype specific. Nucleoprotein (NP) of influenza virus is the alternative candidate antigen to be incorporate in the influenza vaccine as it is the conserved protein in all known influenza virus type A and can confer immunity cross-reacted with different subtypes of influenza A virus. In this study, NP gene was cloned into pQE-31 expression vector and expressed in *E. coli* XL1-blue. The NP protein was expressed in insoluble form with the molecular weight of 56 kDa with a hexahistidine tag at the N-terminal. In order to express the protein in soluble form to facilitate further study, optimization of growth condition by lowering the incubation temperature is now underway.

บทคัดย่อ

ไวรัสไข้หวัดใหญ่เป็นไวรัสที่ก่อให้เกิดความเสียหายมากมายทั้งชีวิตและเศรษฐกิจของประชากรทั่วโลก ถึงแม้จะมีวัคซีนป้องกันไข้หวัดใหญ่ใช้ แต่ต้องปรับการผลิตวัคซีนไข้หวัดใหญ่ตามสายพันธุ์ที่ระบาดอยู่ทุก ๆ ปี เนื่องจากภูมิคุ้มกันที่เกิดขึ้นจะมีความจำเพาะต่อ subtype ของไวรัส Nucleoprotein (NP) ของไวรัสจัดเป็นแอนติเจนตัวเลือกที่น่าสนใจในการนำมาใช้พัฒนาวัคซีน เนื่องจาก NP เป็นส่วนอนุรักษ์ (conserve) ที่มีความแปรผันน้อยมากใน influenza virus type A และสามารถที่จะกระตุ้นระบบภูมิคุ้มกันให้ต้านไวรัสไข้หวัดใหญ่ข้ามสายพันธุ์ได้ ดังนั้นในการศึกษานี้ ได้ทำการโคลน NP เข้าไปใน pQE-31 expression vector และแสดงออกของโปรตีน NP ใน *E. coli* XL1-blue พบว่าโปรตีน NP มีการแสดงออกในรูปโปรตีนที่ไม่ละลายน้ำ และมีน้ำหนักเท่ากับ 56 kDa โดยมี 6x His tag ที่ N-terminal ในขณะนี้ การศึกษากำลังดำเนินอยู่ เพื่อปรับให้โปรทีนมีการแสดงออกในรูปละลายน้ำเพื่อความสะดวกและเหมาะสมที่จะนำไปใช้ศึกษาต่อ โดยการปรับสภาวะการเจริญเติบโตด้วยการลดอุณหภูมิที่ใช้ในการบ่ม

Key Words : nucleoprotein, influenza, cloning, expression
คำสำคัญ : นิวคลีโอโปรตีน ไข้หวัดใหญ่ การโคลนนิ่ง การแสดงออก

**Student, Master of Science, program in Medical Microbiology, Department of Microbiology, Faculty of Medicine, Khon Kaen University*

***Lecturer, Department of Microbiology, Faculty of Medicine, Khon Kaen University*

****Assistant Professor, Department of Microbiology, Faculty of Medicine, Khon Kaen University*

*****Associated Professor, Department of Microbiology, Faculty of Medicine, Khon Kaen University*

Introduction

Influenza viruses, members of the *Orthomyxoviridae* family, are classified as types A, B, or C, based on antigenic differences in their nucleoprotein (NP) and matrix protein (M1) (Horimoto & Kawaoka, 2006). Type A influenza viruses are further divided into subtypes based on the antigenic relationships in the surface glycoproteins, haemagglutinin (HA) and neuraminidase (NA). At present 16 HA subtypes have been recognised (H1–H16) and 9 NA subtypes (N1–N9) (Alexander, 2007).

The emergence of epidemics of acute respiratory highly infectious illness has afflicted humans since remote times (Proenca-Modena, Macedo, & Arruda, 2007). Human influenza pandemics over the last 100 years have been caused by H1, H2, and H3 subtypes of influenza A viruses (Sandrock & Kelly, 2007). Since 1997, the increased circulation of highly pathogenic avian influenza virus subtypes (i.e., H5, H7) have been consistently detected in poultry and wild birds in many countries, mostly in Asia (Proenca-Modena et al., 2007; Sandrock et al., 2007). Current influenza virus vaccines try to induce strong antibody responses against the viral glycoproteins hemagglutinin (HA) and neuraminidase (NA). The presence of neutralizing antibodies specific for the HA at systemic or mucosal sites of infection provides immediate protection against infection with influenza viruses. Although antibodies specific for the NA do not neutralize infectivity, they prevent the release of new virus particles resulting in limitation of virus replication. (Subbarao & Joseph, 2007). Since influenza A virus undergoes genetic variation by two mechanisms genetic drift is point mutations in antigenically important positions caused by selective pressure from host immune response and genetic shift is a substitution of a whole genome fragment from one subtype to another resulting in changes in the antigenicity of the virus (Julkunen et al., 2001). This process made the pathogen escape from the host immunity (Chen & Deng, 2009). The seasonal influenza vaccine which are updated annually can not use for the new virus subtype. Therefore, the influenza vaccine must be updated according to the antigenic changes occurred in the pool of circulating virus strains (Gerhard et al., 2006). The idea to developing a vaccine based on an invariant influenza A protein that would induce long-lasting immunity providing cross-protection against different influenza variants is very tempting (Huang, Liu, Muller, Levandowski, & Ye, 2001; Jegerlehner, Schmitz, Storni, & Bachmann, 2004)

NP is the major structural protein that interacts with the RNA segments to form ribonucleoprotein. It is encoded by RNA segment 5 of influenza A virus and is 1,565 nucleotides in length containing 498 amino acids. Moreover, NP is a more conserved protein than the membrane glycoproteins. The NP gene is rarely mutated and is conserved by approximately 90% in many strains of influenza type A virus (Ohba et al., 2007). There are several reports indicating that NP is the viral target antigen recognized by cross-reactive cytotoxic T lymphocytes (CTL). Studies by Yewdell *et al.* demonstrated that cells infected with the recombinant vaccinia virus expressed the NP gene are lysed by cross-reactive CTL (Yewdell, Bennink, Smith, & Moss, 1985). Roy *et al.* has reported that murine and human cytotoxic T lymphocytes (CTLs) activated by NP epitopes are capable of cross-reacting with NP from different influenza A strains. They demonstrated the protective efficacy of a single dose of adenovirus vaccine

expressing the influenza A NP in protecting against two different avian H5N1 strains in chimpanzee (Roy et al., 2007).

In this study, NP gene was cloned into the pQE-31 expression vector, sequenced and expressed in *E. coli*. The expressed protein was purified and characterized by SDS-PAGE and immunoblotting.

Materials and methods

Cloning of NP

The full-length NP sequences were retrieved from Genbank (http://www.ncbi.nlm.nih.gov/). The BioEdit software was used for alignment of the multiple sequences retrieved. The primers were designed and analyzed by useing Oligonucleotide Properties Calculator Program (Oligo Calc) and Oligoanalyzer program. The genomic RNA of influenza virus type A strain H3N2 was isolated from seed viral culture by using HighPure Viral RNA Kit (Roche, Germany) and reverse transcribed the viral RNA using Transcriptor First Strand cDNA Synthesis Kit (Roche, Germany) according to the manufacturer's instructions. The NP gene was amplified using the following primers: 5'atatggtacctatggcgtcccaaggcaccaa ac 3' and 5' agtggtcgacttaattgtcgtactcttctgcattg 3' (*Kpn*I and *Sal*I restriction sites are underlined, respectively). PCR was performed under the condition of 95 ºC 15 min , 94 ºC 30 sec , 60 ºC 30 sec , 72 ºC 2 min, for 30 cycles and of 72 ºC 7 min for 1 cycle. The amplified products were firstly cloned into pGEM-T Easy vector and further subcloned to the expression vector, pQE-31. The recombinant plasmids were transformed into *E. coli* strain XL1-blue using electroporation method. The recombinant clones were selected by ampicillin selection, and further validated by restriction enzyme digestion and finally by DNA sequencing.

Expression of NP

A single colony of verified transformant from above was inoculated in 3 ml of LB broth containing both amplicillin (100 µg/ml) and tetracycline (25 µg/ml) and grown overnight at 37 ºC with shaking at 200 rpm. The overnight culture was added into LB prewarmed media (with both antibiotics) and grown at 37 ºC with vigorous shaking until the density reached OD_{600} of 0.4-0.6. NP expression was induced by adding 1mM isopropyl-µ-D-thiogalactoside (IPTG) to culture with a final concentration of 1 mM and the culture was further incubated for 4 hr. at 37 ºC with collection of 1 ml of culture sample each hour. After IPTG induction at specific time, the bacteria cells were harvested by centrifugation at 8000 rpm for 15 min at 4 ºC. The supernatant was discarded and the cell pellets were frozen at -70 ºC until use.

To determine protein solubility, the cell pellets were resuspened in 3 ml of lysis buffer with lysozyme (1 mg/ml) and incubated on ice for 30 min, then, disrupted by sonication for 6×20 sec with 20 sec pauses at 200-300W while keeping on ice all the time. After lysis, the lysate was centrifuged at 12,000 rpm for 30 min at 4 ºC. The supernatant containing soluble proteins were transferred to a new tube. The pellet containing insoluble protein was resuspended in 500 µl lysis buffer. The recombinant proteins either from soluble or insoluble fraction were stored at -20 ºC until use for analysis.

Analysis by SDS-PAGE and immunoblotting

The expressed proteins were subjected to 5% SDS-PAGE with stacking gel and 12% separating gel, stained with Coomasie Brilliant Blue R-250 blue and further analyzed by Western blot analysis using RGS-His Antibody (Qiagen GmbH, Germany). Following

electrophoresis, proteins in the polyacrylamide gel were transferred to a positively charged nitrocellulose membranes by semi-dry electroblotting. After transfer, the membrane was washed twice for 10 min each time with TBS buffer (10 mM Tris-Cl, pH 7.5, 150 mM NaCl), incubated for 1 hr. in blocking buffer (3% BSA (w/v) in TBS buffer) and washed twice for 10 min each time in TBS-Tween/Triton buffer (20 mM Tris-Cl, pH 7.5, 500 mM NaCl, 0.05% (v/v) Tween 20, 0.2% (v/v) Triton X-100) and then, in TBS buffer. The membrane was incubated in RGS-His antibody with the dilution of 1/1000 in blocking buffer for 1 hr. After incubation, the membrane was washed twice in TBS-Tween/Triton buffer and once in TBS buffer for 10 min each. The washed membrane was then incubated with goat anti-mouse IgG-HRP (Santa cruz) diluted to 1/1000 in blocking buffer (5% milk, 0.05% Tween-20 in TBS buffer) for 45 min. The membrane was washed 3 times for 5 min each with TBS-Tween buffer (20 mM Tris-Cl, pH 7.5, 500 mM NaCl, 0.05% (v/v) Tween 20) and once for 5 min with TBS buffer. The membrane was then added with the DAB (3, 3´-diaminobenzidine tetrahydrochloride) substrate solution and incubated until the desired band can be clearly visualized. All immunostaining processes were done at room temperature.

Results and discussion

Cloning of NP gene

In this study, the NP gene was amplified by PCR and has the PCR product length of 1,552 bp (Fig. 1). The PCR products were firstly cloned into pGEM-T Easy vector and transformed into *E. coli* XL-1 blue by electroporation. All selected recombinant clones were verified by *Eco*RI digestion. The digested recombinant plasmid DNA containing two DNA fragments with length of 3 kb and 1.6 kb (Fig. 2) was further confirmed by PCR using the same primer pairs. The recombinant plasmid can be amplified to give PCR product of 1,552 bp. (Fig. 3). The correctness of NP sequence was done by DNA sequencing

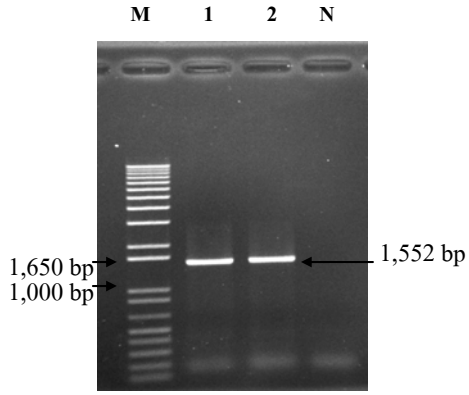

Fig. 1. Ethidium bromide staining gel of amplified products of NP gene. M = 1 kb plus DNA ladder, lane 1-2 = NP amplified product using influenza genome cDNA as template, lane N = negative control.

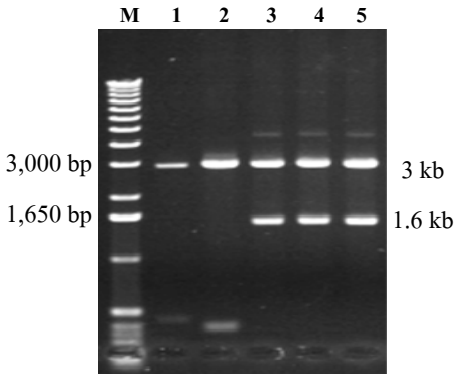

Fig. 2. Ethidium bromide staining gel of *Eco*RI-digested recombinant plasmid pGEM-T Easy containing NP gene. M = 1 kb plus DNA ladder, lane 1-2 = *Eco*RI-digested recombinant plasmids from negative clones, lane 3-5 = *Eco*RI-digested recombinant plasmids from positive clones.

The NP gene was further subcloned to expression vector, pQE-31. The pQE-31 vector was designed for the high expression of 6xHis tagged proteins in *E. coli* and

used T5 promoter transcription-translation system. The NP gene was cloned into pQE-31 vector at *Kpn*I and *Sal*I site. The derived recombinant clones were verified by PCR (Fig. 4) and *Kpn*I and *Sal*I -digestion. Figure 5 showed that the digested recombinant plasmid containing two DNA fragments with the length of 3.4 kb and 1,552 bp (lane1-2). Final verification by DNA sequencing showed that the corrected sequence of NP gene was inserted in the recombinant plasmid based on pQE-31 vector backbone.

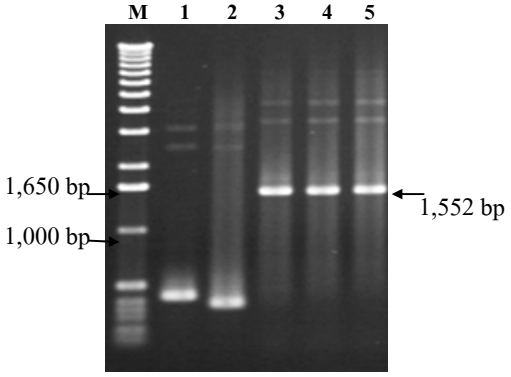

Fig. 3. Ethidium bromide staining gel of amplified products of NP gene using recombinant plasmid pGEM-T Easy containing NP gene as template. M = 1 kb plus DNA ladder, lane 1-2 = amplified products derived from negative clones, lane 3-5 = NP amplified products derived from positive clones.

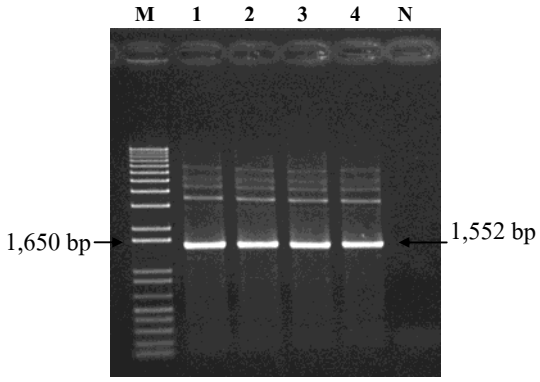

Fig. 4. Ethidium bromide staining gel of amplified products of NP gene using recombinant plasmid pQE-31 containing NP gene as template. M = 1 kb plus DNA ladder, lane 1-4 = NP amplified products derived from positive clones, lane N = negative control.

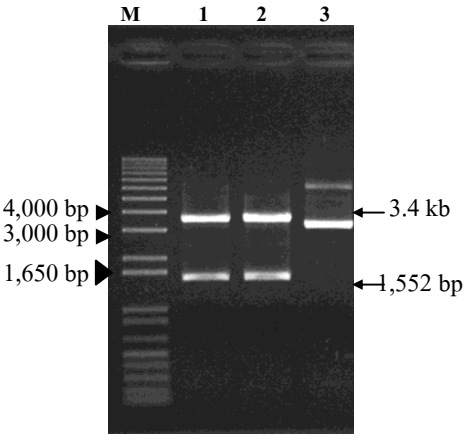

Fig. 5. Ethidium bromide staining gel of *Kpn*I and *Sal*I- digested recombinant plasmid pQE-31 containing NP gene. M= 1 kb plus DNA ladder, lane1-2 = *Kpn*I and *Sal*I- digested recombinant plasmids from positive clones, lane 3 = non-digested recombinant plasmid pQE-31 containing NP gene.

NP expression and analysis by SDS-PAGE and immunoblotting

For expression of the NP, the recombinant plasmid was transformed into *E. coli* XL1-blue. For small-scale expression, the incubation of transformed bacteria was done in 50 ml of prewarmed media (containing antibiotics with the overnight culture (1:50) and grew at 37 ºC with shaking until OD_{600} reached 0.4-0.6. After IPTG induction for 4 hr. with 1 ml sampling each hour, NP was expressed as 56 kDa-protein with 6xHis tagged and showed the highest expression at 4 hr. (Fig. 6).

The primary consideration for recombinant protein expression and purification is the experimental purpose for which the protein will be utilized. In our experiment, not only obtaining the recombinant protein but also requiring its functional activity for use in other future experiments. One critical point to

consider for recombinant protein production in *E. coli* is the protein solubility. From our studies, the expressed protein has been found as insoluble form in cell pellet fraction (Fig. 7, Lane 3). Previous studies by Hu, the expressed protein was found as insoluble form similar to our studies (Hu, 2005). As overproduction of heterologous proteins in cytoplasm of *E. coli*. is often accompanied by their misfolding and segregation into insoluble aggregates known as inclusion bodies (Baneyx, 1999), several experimental studies have tried to minimize the formation of inclusion bodies and improve protein folding (Makrides, 1996). An alternative, simpler and often equally successful approach, is to manipulate the growth condition, (Georgiou & Valax, 1996) such as reduction in growth temperature following induction. It was found that growth at lower temperature usually facilitate correct folding due to the decrease in the driving force for protein self-association, a slower rate of protein synthesis, and changes in the folding kinetic of polypeptide chain (Georgiou et al., 1996).

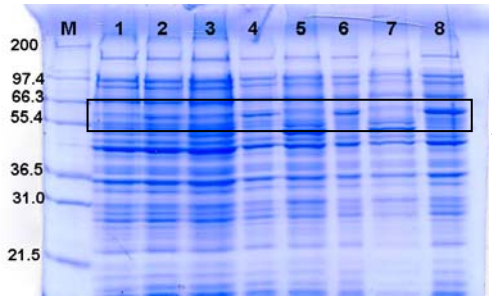

Fig. 6. SDS-PAGE analysis of expressed NP in *E. coli* XL1-blue. Samples from whole cells collected at 1,2,3,4 hr. after IPTG induction were electrophoresed on 12% SDS-PAGE. The protein bands were visualized by Coomasie Brilliant blue R-250 staining. Lane M = molecular weight marker Mark12 (kDa), lane 1, 3, 5, 7 = Non IPTG-induced cells at 1, 2, 3, 4 hr., lane 2, 4, 6, 8 = IPTG-induced cells at 1, 2, 3, 4 hr.

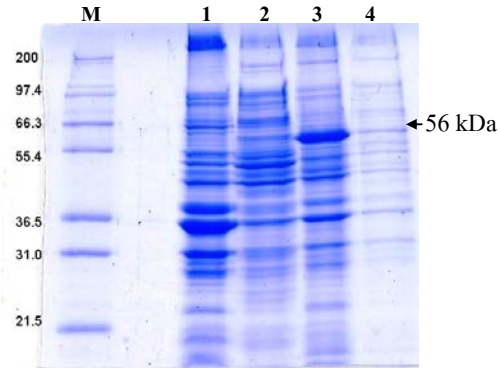

Fig. 7. SDS-PAGE analysis of NP expressed in *E. coli*. Samples from both pellet and supernatant at 4 hr. were electrophoresed on 12% SDS-PAGE. The protein bands were visualized by Coomasie blue R-250 staining. Lane M = molecular weight marker Mark12 (kDa), lane 1 = Non-IPTG induced cell lysate from pellet, lanes 2 = Non-IPTG induced cell lysate from supernatant, lanes 3 = IPTG-induced cell lysate from pellet, lanes 4 = IPTG induced cell lysate from supernatant.

In order to obtain NP as soluble protein, further optimization of the condition is now underway.

Conclusions

In this study, NP gene was cloned and expressed in *E. coli* XL1-blue. NP protein was expressed as the insoluble form. Further optimization is needed for obtaining the expressed protein in the soluble form.

Acknowledgements

This work was partially supported by grant from Research and Diagnostic Center for Emerging Infectious Diseases and from National Center for Genetic Engineering and Biotechnology (BIOTEC).

References

Alexander, D. J. 2007. An overview of the epidemiology of avian influenza. Vaccine, 25(30): 5637-5644.

Baneyx, F. 1999. Recombinant protein expression in *Escherichia coli.* Curr Opin Biotechnol, 10(5): 411-421.

Chen, J., & Deng, Y. M. 2009. Influenza virus antigenic variation, host antibody production and new approach to control epidemics. Virol J, 6: 30.

Georgiou, G., & Valax, P. 1996. Expression of correctly folded proteins in *Escherichia coli.* Curr Opin Biotechnol, 7(2): 190-197.

Gerhard, W., Mozdzanowska, K., & Zharikova, D. 2006. Prospects for universal influenza virus vaccine. Emerg Infect Dis, 12(4): 569-574.

Horimoto, T., & Kawaoka, Y. 2006. Strategies for developing vaccines against H5N1 influenza A viruses. Trends Mol Med, 12(11): 506-514.

Huang, X., Liu, T., Muller, J., Levandowski, R. A., & Ye, Z. 2001. Effect of influenza virus matrix protein and viral RNA on ribonucleoprotein formation and nuclear export. Virology, 287(2): 405-416.

Jegerlehner, A., Schmitz, N., Storni, T., & Bachmann, M. F. 2004. Influenza A vaccine based on the extracellular domain of M2: weak protection mediated via antibody-dependent NK cell activity. J Immunol, 172(9): 5598-5605.

Julkunen, I., Sareneva, T., Pirhonen, J., Ronni, T., Melen, K., & Matikainen, S. 2001. Molecular pathogenesis of influenza A virus infection and virus-induced regulation of cytokine gene expression. Cytokine Growth Factor Rev, 12(2-3): 171-180.

Makrides, S. C. 1996. Strategies for achieving high-level expression of genes in *Escherichia coli.* Microbiol Rev, 60(3): 512-538.

Ohba, K., Yoshida, S., Zahidunnabi Dewan, M., Shimura, H., Sakamaki, N., Takeshita, F., Yamamoto, N., & Okuda, K. 2007. Mutant influenza A virus nucleoprotein is preferentially localized in the cytoplasm and its immunization in mice shows higher immunogenicity and cross-reactivity. Vaccine, 25(21): 4291-4300.

Proenca-Modena, J. L., Macedo, I. S., & Arruda, E. 2007. H5N1 avian influenza virus: an overview. Braz J Infect Dis, 11(1): 125-133.

Roy, S., Kobinger, G. P., Lin, J., Figueredo, J., Calcedo, R., Kobasa, D., & Wilson, J. M. 2007. Partial protection against H5N1 influenza in mice with a single dose of a chimpanzee adenovirus vector expressing nucleoprotein. Vaccine, 25(39-40): 6845-6851.

Sandrock, C., & Kelly, T. 2007. Clinical review: update of avian influenza A infections in humans. Crit Care, 11(2): 209.

Subbarao, K., & Joseph, T. 2007. Scientific barriers to developing vaccines against avian influenza viruses. Nat Rev Immunol, 7(4): 267-278.

Yewdell, J. W., Bennink, J. R., Smith, G. L., & Moss, B. 1985. Influenza A virus nucleoprotein is a major target antigen for cross-reactive anti-influenza A virus cytotoxic T lymphocytes. Proc Natl Acad Sci U S A, 82(6): 1785-1789.