Influenza Virus Hemagglutinin: A Model for Protein N-Glycosylation in Recombinant *Escherichia coli*

Behrokh Farahmand^{a, c} Mahvash Khodabandeh^c Fereidoun Mahboudi^b
Fateme Fotouhi^a Maryam Saleh^a Farzaneh Barkhordari^b
Mansoureh Tabatabaian^a Farnoush Parsaie Nasab^d Masoumeh Tavassoti Kheiri^a

^aDepartment of Virology (Influenza Unit) and ^bBiotechnology Research Center, Pasteur Institute of Iran,

^cDepartment of Industrial and Environmental Biotechnology, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran; ^dInstitute of Microbiology, Department of Biology, Eidgenössische Technische Hochschule Zurich, Zurich, Switzerland

Key Words

Hemagglutinin expression • Influenza • N-Glycosylation • Recombinant *Escherichia coli*

Abstract

Background: The hemagglutinin molecule of influenza virus is considered as an ideal model to study biological processes as well as the effect of glycosylation on the function of glycoproteins. **Objectives:** The large subunit of the influenza virus A/New Caledonia/20/99 (H1N1) hemagglutinin (HA1) was expressed in recombinant *Escherichia coli* containing the glycosylation system of *Campylobacter jejuni*. This viral glycoprotein contains glycosylation motifs recognized by prokaryotic and eukaryotic oligosaccharyltransferases. **Methods:** In order to express the hemagglutinin large subunit gene, the gene was amplified using reverse transcription polymerase chain reaction (RT-PCR), and it was cloned in pET22b for periplasmic expression. **Results:** Western blotting and lectin blotting bands confirmed glycosylation of the HA1 in recombinant *E. coli*. **Conclusion:** Such a successful accomplishment of hemagglutinin expression in recombinant

E. coli can be used to construct carbohydrates in hemagglutinin molecules of different strains in order to produce effective antigens for vaccine and rapid diagnostic kits against new emerging viruses.

Copyright © 2011 S. Karger AG, Basel

Introduction

N-Glycosylation is the most common co- or post-translational modification of proteins in eukaryotes [1]. Approximately 70% of human therapeutic proteins are N-linked glycoproteins [2].

Discovery of the N-glycosylation system in *Campylobacter jejuni*, and its transfer to *Escherichia coli* was a revolution in biotechnology and glycotechnology [1, 3]. This finding directed essential changes in glycosylation studies.

The *C. jejuni* glycosylation cluster genes (*pgl*) consist of putative glycosyltransferases and enzymes involved in sugar biosynthesis (fig. 1) [4].

Fig. 1. N-linked glycosylation pathway genes in *C. jejuni*.

pglB encodes the oligosaccharyltransferase, which is a homologue of the eukaryotic Stt3p, the most conserved protein in the oligosaccharyltransferase complex found exclusively in eukaryotes and archaea [1]. The *pglB* can detect prokaryotic N-glycosylation motif, D/ExNyS/T. The 'x' and 'y' could be any amino acid except P [5].

Function of the *pgl* gene locus of *C. jejuni* in a flexible organism like *E. coli* has provided the ability to express recombinant glycoproteins, and possibly to improve carbohydrate engineering [2, 3, 6].

One of the candidate glycoproteins to be studied in this recombinant *E. coli* is the hemagglutinin of influenza virus. The structure of this molecule was revealed by Watowich et al. [7]. Carbohydrate linkage of the molecule is through N-glycan, and there is no carbohydrate with the O-glycan linkage [8].

This valuable molecule is considered as an ideal model to study many biological processes such as membrane fusion, folding in secretion pathways as well as the effect of glycosylation on the function of glycoproteins [9–11]. Hemagglutinin is a membrane glycoprotein of the influenza virus that is expressed as a trimer. Each monomer consists of two subunits, of which HA1 (~36 kDa) and HA2 (~27 kDa) are glycosylated and linked together by a disulphide bond. HA1 is the amino-terminal part of the molecule which has most of the oligosaccharide chains and antigenic sites [12]. There is one prokaryotic N-glycosylation motif (D/ExNyS/T) among the glycosylation sites of most of the influenza A/H1N1 hemagglutinin amino acid sequences such as:

A/New Caledonia/20/99 (H1N1)
Met DTICIGYHANNSTDVDTVLEKNTVTTH

A/Puerto Rico/8/34 (H1N1)
Met DTICIGYHANNSTDVDTVLEKNTVTTH

A/California/07/2009 (H1N1)
Met DTLICIGYHANNSTDVDTVLEKNTVTTH

This study aims to demonstrate the expression of the influenza virus A/New Caledonia/20/99 (H1N1) hemagglutinin large subunit with an N-glycosylation prokaryotic motif in recombinant *E. coli* and analyze glycosylation using lectin blotting. Selection of hemagglutinin as a model to study the expression of the glycosylated form in recombinant *E. coli* and confirmation of its glycosylation can be a beneficial step in the study of eukaryotic glycoproteins in recombinant *E. coli*.

Materials and Methods

Virus and Bacteria

Human influenza virus A/New Caledonia/20/99 (H1N1) was used to inoculate in the confluent Madine-Darby canine kidney cell line grown in Dulbecco's modified Eagle's medium supplemented with 10% (v/v) FBS, 100 U/ml of penicillin G and 100 µg/ml of streptomycin [13]. *E. coli* Top10F' and *E. coli* BL21 (DE3) were used as host for cloning experiments and high level expression of proteins, respectively. A list of plasmids and primers used in this work is given in table 1.

Construction of pET-HA1

The HA1 gene lacking its 48-bp signal peptide was amplified by RT-PCR from influenza A/New Caledonia/20/99 (H1N1) genomic RNA as the template. The universal primer (Uni12) was used for cDNA synthesis and specific primers for cDNA amplification (table 1) [14].

The RT-PCR product was cloned into the pGEM-TEasy vector (Promega) for sequencing analysis. For periplasmic expression, HA1 gene was subcloned into pET22b downstream of the Sec-dependent signal peptide PelB (ssPelB).

Coexpression of pET-HA1 and pACYC (*pgl*)

The competent BL21 (DE3) cells were transformed by two constructs, pET-HA1 and pACYC (*pgl*). The colonies harboring both constructs were selected on LB agar containing ampicillin and chloramphenicol. The selected colonies were then grown in LB broth at 37°C with shaking at 185 rpm. To reach the optical density of 0.4–0.6 at a wavelength of 600 nm (OD₆₀₀), the induction was performed using different concentrations of IPTG. After 2 h of induction, the cells were harvested by centrifugation (9,000 g,

Table 1. Plasmids and primers used in this study

Plasmids and primers	Characteristics or sequences (5'–3')	Source
Plasmids		
pET22b (+)	ColE1, PelB signal peptide, Amp (~40 copy number)	Novagen, USA
pACYC (<i>pgl</i>)	P15A, Cm (10–12 copy number)	[1]
Primers		
F HA1	5'-GCAAGCTTATGGACACAATATGTATA-3': <i>HindIII</i>	This study
R HA1	5'-CACTCGAGACCTACTTTGGACACTC-3': <i>XhoI</i>	This study
Uni12	5'-AGCAAAAGCAGG-3'	[14]
F = Forward primer; R = reverse primer; Uni12 = universal primer with 12 mer.		

3 min at 4°C) and expression was determined by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) using the method of Laemmli [15].

Periplasmic Protein Extraction by Osmotic Shock Procedure

Periplasmic osmotic shock fluid was obtained by the method of Koshland and Botstein [16].

Periplasmic Expression and Glycosylation Analysis

The periplasmic proteins were separated under reducing conditions (12%) by SDS-PAGE. Proteins were then transferred onto the nitrocellulose membrane for immunoblotting and lectin blotting.

Immunoblotting

Western blotting was carried out using specific monoclonal antibody against the influenza A/New Caledonia/20/99 (H1N1) hemagglutinin (anti-HA; ab66189; Abcam) as first antibody (1:1,000 dilution), and horseradish peroxidase (HRP)-conjugated goat anti-mouse monoclonal antibody (A8924; Sigma-Aldrich) was used as the second antibody (1:2,000 dilution). The nitrocellulose membrane was blocked with 2.5% bovine serum albumin in PBS containing Tween 20. Blots were finally developed using diaminobenzidine (DAB) solution as an HRP substrate.

Lectin Blotting

The transferred membrane was blocked with a blocking solution containing 3% (w/v) bovine serum albumin in PBS-T overnight at 4°C. The membrane was incubated with HRP conjugate of soybean agglutinin (SBA; final concentration 10 µg/ml; L2650; Sigma-Aldrich) in blocking buffer for 2 h at room temperature. After washing 5 times with PBS-T, the membrane was incubated in DAB solution [17]. A recombinant HA1 [expressed in BL21 without pACYC (*pgl*)] was used as a negative control.

Lectin Blot Inhibition

The specificity of lectin binding was assessed by the overnight preincubation of lectin conjugate with a different concentration of N-acetylgalactosamine (A2795; Sigma-Aldrich) in PBS-T at 4°C followed by a lectin blotting procedure [17].

Competitive Hemagglutination Inhibition Assay

In this assay, periplasmic rHA1 was incubated (overnight at 4°C) with standard antiserum; antibodies remaining in the supernatant were measured by hemagglutination inhibition (HI; WHO protocol) [18]. HI titers were determined by the reciprocal value of the last dilution of standard antiserum which completely inhibited hemagglutination of chicken RBC.

Results

Expression of the HA1 from H1N1 Influenza Virus in E. coli

The HA1 containing 918 bp was amplified from the cDNA of influenza A/New Caledonia/20/99 (H1N1) using Taq high-fidelity polymerase (fig. 2a) and cloned in to the pGEM-TEasy vector without the signal peptide nucleotide sequence. Sequencing analysis showed that 48 nucleotides (16 amino acids) were removed from N-terminus. This sequence is compatible with reference sequence in NCBI (AJ344014), and there is no mutated nucleotide in extended N-glycosylation motif (D/ExNyS/T; data not shown).

The HA1 was subcloned in pET22b (fig. 2b). It was expressed and induced using IPTG (0.2 mM) in *E. coli* BL21(DE3) cells with and without a plasmid containing the wild-type *pgl* locus from *C. jejuni* (pACYC*pgl*; fig. 3). As a negative control for expression and glycosylation, we used cells carrying a backbone plasmid pET22b (fig. 3, lanes 2 and 3).

Glycosylation Analysis

For glycosylation detection of recombinant protein, periplasmic extracts were concentrated with PEG6000 and analyzed by SDS-PAGE, and subsequently confirmed

Fig. 2. RT-PCR of HA1 and pET-HA1 restriction analysis. **a** PCR product of HA1 gene using cDNA as template. Lane 1, amplified HA1 using Taq high fidelity; lane 2, negative control; lane 3, DNA marker (1 kb). **b** Restriction analysis of pET-HA1. Lane 1, uncut pET-HA1; lane 2, cutting of pET-HA1 using *NcoI*; lane 3, cutting of pET-HA1 using *HindIII/XhoI*; lane 4, DNA marker (1 kb).

Fig. 3. Analysis of periplasmic proteins of *E. coli* BL21 (DE3) over-expressing HA1 in the presence or absence of *C. jejuni* *pgl* locus. Periplasmic proteins were separated by SDS-PAGE and transferred to nitrocellulose, immunodetected with anti-influenza hemagglutinin antibody. **a** Lane 1, marker; lanes 2, 3, cells containing pET22b backbone only; lanes 4, 5, cells expressing influenza HA1. **b** Lane 1, marker; lanes 2, 3, cells transformed by pACYC (*pgl*) locus as first plasmid and pET22b backbone as second plasmid; lanes 4, 5, cells containing pACYC (*pgl*) locus and pET22b-HA1 as second construct.

using immunoblotting. The concentrated periplasmic extracts from cells expressing rHA1 showed a major protein with a mass of 36 kDa (fig. 3, lanes 4 and 5) and additional protein band with a mass of over 36 kDa (~37 kDa), when the pACYC (*pgl*) locus was present in the cells (fig. 3b, lane 4 and 5). Both of the proteins reacted with anti-HA antibody. The low migrated rHA1 band (~37 kDa) protein was not present in extracts from glycosylation-deficient cells (fig. 3a, lane 4 and 5); we hypothesized that this additional protein may be the single glycosylated form of HA1.

To identify carbohydrate structure covalently attached to the influenza rHA1, lectin blot was performed using peroxidase-labeled GalNAc-specific lectin SBA, and the glycosylated band was observed (fig. 4a, lane 5).

The GalNAc-specific lectin SBA recognizes the terminal GalNAc residues present on the *C. jejuni* N-glycans [19].

Expression of influenza HA1 in *E. coli* yielded non-glycosylated HA1 (fig. 3a, lane 4 and 5), whereas coexpression of HA1 with *C. jejuni* locus revealed glycosylated HA1, visualized by DAB substrate (fig. 4a, lane 5).

To confirm the glycosylated band, the lysate was blotted using inhibited HRP-conjugated SBA lectin (with 0.2 M GalNAc), which was determined by the band disappearance (fig. 4b, lane 5).

Competitive HI Assay of Glycosylated rHA1

Antigenicity of glycosylated recombinant HA1 (g-rHA1) and non-g-rHA1 was assessed by testing the protein's ability to remove HI of standard sheep (SH420) antiserum (NIBSC, 02/224) antibodies. Both rHA1 (g-rHA1 and non-g-rHA1) proteins decreased 2-fold in antibody titer (fig. 5).

Discussion

A variety of prokaryotic glycosylated proteins with N-terminally extended glycosylation motif (D/ExNyS/T) has been expressed in this recombinant *E. coli* [5].

According to Nita-Lazar et al. [20], prokaryotic N-glycosylation is accomplished in the periplasmic space where the oligosaccharyltransferase active site is present.

The large subunit of influenza virus is a good model to study N-glycosylation in this recombinant *E. coli*. In this study, the HA1 subunit from the influenza virus hemagglutinin was subcloned in pET22b vector containing ssPelB and successfully expressed in the recombinant *E. coli* (containing the *pgl* locus of *C. jejuni*).

Fig. 4. Analysis of periplasmic HA1 glycosylation in the presence or absence of the *C. jejuni* *pgl* locus. Periplasmic proteins separated by SDS-PAGE and transferred to nitrocellulose. **a** Probed with HRP-conjugated SBA lectin. **b** Probed with inhibited HRP-conjugated SBA lectin. Lane 1, marker; lane 2, cells containing plasmid backbone (pET22b) only; lane 3, cells containing pET22b-HA1; lane 4, cells containing pACYC (*pgl*) and (pET22b) as second plasmid; lane 5, cells containing pACYC (*pgl*) locus and pET22b-HA1 as second construct [D(g-rHA1) – disappeared g-rHA1].

Fig. 5. Competitive HI by glycosylated rHA1 and nonglycosylated rHA1. Serial dilution of inhibited standard antibodies (row 2, inhibition using g-rHA1; row 3, inhibition using non-g-rHA1). Antibodies were mixed with 4HAU of influenza virus A/New Caledonia/20/99(H1N1) and incubated for approximately 45 min at room temperature. Then, 0.5% suspension of chicken RBC was added and incubated for 60 min at room temperature. Row 1, control of standard antibody.

Nucleotide sequence analysis showed that the obtained sequence is compatible with the NCBI coding sequence (AJ344014), and 48 nucleotides (16 amino acids) were removed from the N-terminus. Based on the rHA1 and the ssPelB sequence, the molecular weight of rHA1 was calculated to be 36.04 kDa. Western blot analysis confirmed the obtained 36-kDa protein (fig. 3, lane 4 and 5). These results suggested that the HA1 protein was successfully expressed in the *E. coli* strains both in the presence and absence of the *pgl* locus.

In this experiment, the partially purified periplasmic hemagglutinin was concentrated with PEG6000. Following Western blotting, the two bands of rHA1 appeared in recombinant *E. coli*, showing a different molecular weight of rHA1. The lectin blot and lectin blot inhibition results confirmed the glycosylation of the rHA1. Disappearance of the glycosylated band after blotting with inhibited HRP-conjugated SBA lectin indicated that *C. jejuni*-specific oli-

gosaccharide was linked to rHA1 and successive glycosylation of rHA1. Therefore, the prokaryotic N-glycosylation motif of rHA1 consists of *C. jejuni*-specific oligosaccharide side chain, which could react with SBA lectin.

The results of competitive HI demonstrated that g-rHA1 and non-g-rHA1 have the same antigenicity. We confirmed that both rHA1s react with standard antiserum, and demonstrated that the prokaryotic N-glycosylation site does not interrupt its antigenicity.

Wang et al. [21] showed that hemagglutinin with short glycans at the glycosylation sites, even with a monosaccharide, possesses higher immunogenicity. Since the *C. jejuni* glycosylation system leads to a heptasaccharide, shorter than eukaryotic oligosaccharides, the rHA1 (with D/ExNyS/T motifs in all glycosylation sites) expressed in the recombinant *E. coli* using the *C. jejuni* glycosylation system may have higher antigenicity than the one expressed in eukaryotic cells.

On the other hand, the key enzyme of this system, pglB, is a flexible oligosaccharyltransferase, so different carbohydrate structures consisting of reduced monosaccharides such as FucNAc, GalNAc, GlcNAc in reductive ends can be transferred [3].

It has been shown that the glycosylation system of *C. jejuni* can only transfer heptasaccharide with a complete length or without any side branch to recipient protein via oligosaccharyltransferase (pglB) enzyme, whereas short oligosaccharide can be made in recombinant *E. coli* by destroying the genes of this pathway such as pglA, pglJ, pglH, and pglI which is able to transfer to the recipient protein [3]. Therefore, using this system for HA glycosylation with a shorter carbohydrate (monosaccha-

ride) and high immunogenicity could be produced. Thus, using these oligosaccharides and engineering the other eukaryotic N-glycosylation motifs of hemagglutinin to glycan recipient sequences in prokaryotic systems (D/ExNyS/T) can improve hemagglutinin glycosylation in *E. coli*.

Acknowledgements

We thank Prof. Markus Aebi (Institute of Microbiology, Department of Biology, Swiss Federal Institute of Technology Zurich) for the pACYC (*pgl*). This work was supported by the Influenza Unit of the Pasteur Institute of Iran, Tehran (grant No. 411).

References

- Wacker M, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR, Dell A, Wern BW, Aebi M: N-Linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* 2002;298:1790–1793.
- Pandhal J, Wright PC: N-Linked glycoengineering for human therapeutic proteins in bacteria. *Biotechnol Lett* 2010;32:1189–1198.
- Langdon RH, Cuccui J, Wern BW: N-Linked glycosylation in bacteria: an unexpected application. *Future Microbiol* 2009;4:401–412.
- Weerapana E, Imperiali B: Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology* 2006;16:91R–101R.
- Kowarik M, Young NM, Numao S, Schulz BL, Hug I, Callewaert N, Mills DC, Watson DC, Hernandez M, Kelly JF, Wacker M, Aebi M: Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J* 2006;25:1957–1966.
- Borman S: Glycosylation engineering, controlling personalities the wild sugars on proteins and natural products. *Chem Eng News* 2006;84:13–22.
- Watowich SJ, Skehel JJ, Wiley DC: Crystal structures of influenza virus hemagglutinin in complex with high-affinity receptor analogs. *Structure* 1994;2:719–731.
- Ward CW, Gleeson PA, Dopheide TA: Carbohydrate composition of the oligosaccharide units of the hemagglutinin from the Hong Kong influenza virus A/Memphis/10272. *Biochem J* 1980;189:649–652.
- Daniels R, Kurowski B, Johnson AE, Hebert DN: N-linked glycans direct the cotranslational folding pathway of influenza hemagglutinin. *Mol Cell* 2003;11:79–90.
- Imai M, Mizuno T, Kawasaki K: Membrane fusion by single influenza hemagglutinin trimers. *J Biol Chem* 2006;281:12729–12735.
- Schulze I: Effects of glycosylation on the properties and functions of influenza virus hemagglutinin. *J Infect Dis* 1997;176:S24–S28.
- Nicholson KG: Textbook of Influenza. London, Blackwell Science, 1998, pp 54–64.
- Shahidi M, Kheiri MT, Amini-Bavil-Olyae S, Hosseini M, Moattari A, Tabatabaeian M, Bashari R, Sarrafi-Foroshani R, Mahboudi F: Molecular and phylogenetic analysis of human influenza virus among Iranian patients in Shiraz, Iran. *J Med Virol* 2007;79:803–810.
- Chan CH, Lin KL, Chan Y, Wang Y, Chi Y, Tu H, Shieh H, Iu W: Amplification of the entire genome of influenza A virus H1N1 and H3N2 subtypes by reverse-transcription polymerase chain reaction. *J Virol Methods* 2006;136:38–43.
- Laemmli UK: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 1970;227:680–685.
- Koshland D, Botstein D: Secretion of beta-lactamase requires the carboxy end of the protein. *Cell* 1980;20:749–760.
- Tantrawatpan C, Maleewong W, Wongkham C, Wongkham S, Intapan PM, Nakashima K: Characterization of *Fasciola gigantica* adult 27-kDa excretory-secretion antigen in human fascioliasis. *Parasitol Res* 2003;91:325–327.
- World Health Organization: Manual on Animal Influenza Diagnosis and Surveillance. Geneva, World Health Organization, 2002, pp 28–39.
- Wacker M, Feldman MF, Callewaert N, Kowarik M, Clarke BR, Pohl NL, Hernandez M, Vines ED, Valvano MA, Whitfield C, Aebi M: Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. *Proc Natl Acad Sci* 2006;103:7088–7093.
- Nita-Lazar M, Wacker M, Schegg B, Amber S, Aebi M: The N-X-S/T consensus sequence is required but not sufficient for bacterial N-linked protein glycosylation. *Glycobiology* 2005;15:361–367.
- Wang C, Chen J, Tseng Y, Hsu C, Hung Y, Chen S, Chen C, Khoo K, Cheng T, Cheng Y, Jan J, Wu C, Ma C, Wong C: Glycans on influenza hemagglutinin affect receptor binding and immune response. *Proc Natl Acad Sci* 2009;106:18137–18142.