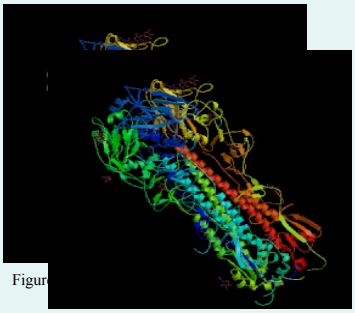
Cloning Of Hemagglutinin (HA) Protein Of Influenza A virus – Potential For Sialic Acid Linkage Discrimination.

Jakub Szczepaniak, O'Connell Michael and O'Connor Brendan

BAS Centre for BioAnalytical Science

Introduction:

The initial step in infection of a cell by influenza A virus is the attachment of a virus particle to the target cell. This is accomplished by interaction of a glycoprotein, hemagglutinin (HA), found on the surface of the viral lipid membrane with cell-surface oligosaccharides containing sialic acids. All influenza virus attachment requires terminal sialic acid residues and two major linkages between sialic acid (Neu5Ac) and the penultimate galactose (Gal) residues of carbohydrate side chains are found in nature, Neu5Ac(α2,3)-Gal and Neu5Ac(α2,6)-Gal. The HA's of different subtypes of influenza A virus exhibit different recognition specificities for these linkages and these linkage specificities have been correlated with host range specificity. The ability of the HA protein to differentiate sialic acid linkages makes it an interesting candidate for use in the characterization of glycoprotein's potentially facilitating the discrimination of alternate glycoforms of biopharmaceutical therapeutics and their subsequent purification.



Figur

Project Objectives:

- To amplify and clone the sialic acid binding domain (HA1 subunit) of influenza A virus A/Sichuan (H3N2) into an *E. coli* expression vector to incorporate C or N terminal affinity tags.
- To optimise the expression of the HA1 subunit of HA in *E. coli* and its subsequent purification via the incorporated affinity tags and to functionally and structurally characterise the affinity tagged protein.
- To alter the sialic acid binding specificity of the wild type human HA protein from α2-6 to α2-3 through the introduction of specific mutations.

Structure Of The HA Protein

The HA protein is synthesized as a precursor protein (550 aa's) that trimerizes prior to being transferred to the cell surface. Each HA0 monomer has an N-terminal HA1 domain, encompasing the sialic acid binding site, and a HA2 domain (see Figure 2). At the membrane the HAO protein is cleaved at a site between the HA1 and HA2 domains, which then only remain attached to each other by virtue of a disulfide bound. This cleavage generates the C-termius of the HA1 subunit (aa's 1-328) and N-terminus of the HA2 subunit (330-550: 221aa's).

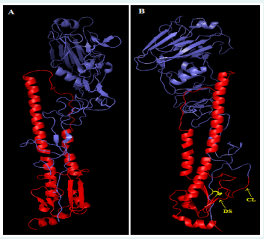


Figure 2: (A) Structure of HAO monomer with HA1 domain (blue) and HA2 domain (red) . **(B)** Cleaved loop (CL) and bridging disulfide bond (DS) that joins the subunits after cleavage are indicated.

Molecular Basis For Sialic Acid Linkage Specificity Of The HA1 Domain.

HAs of all 16 antigenic subtypes (H1–H16) of influenza found in avian influenza viruses bind preferentially to sialic acid in α2-3-linkage; equine viruses bind the α2-3 linkage; swine influenza viruses are reported to bind sialic acid in both α2-3- and α2-6-linkages and human viruses of the H1, H2, and H3 subtypes recognize sialic acid in a α2-6-linkage. Since an avian origin is proposed for the HA of human and swine viruses then these HA's must have effected a change in the sialic acid linkage specificity. The amino acids that make up the HA1 sialic acid binding site of avian HA's are highly conserved, even among the HAs of different subtypes. There is however variation when compared to human HA's. Residues at positions 226 and 228 in the HA1 subunit play a particularly significant role in determining sialic acid linkage specificity

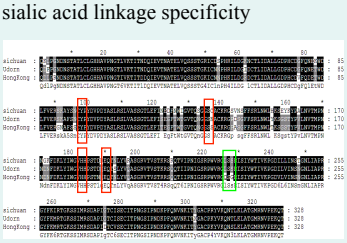


Figure 3: Alignment Of HA1 Sequence Of 3 H3N2 Subtype Influenza A Viruses. Udorn and Sichuan are both human while HongKong is a duck virus. The avian HA1 possess Glutamine (Gln) at 226 and Glycine (Gly) at 228 while human HA's of possess Leucine (Leu) at position 226 (L226) and Serine at position 228 . Other residues important for sialic acid binding are highly conserved across all subtypes and are also highlighted .

Sialic Acid Linkage Specificity Is Dependent On Oligosaccharide Conformations

The crystal structures of the HA protein bound to two pentasaccharides, LSTc and LSTa, of identical composition but differing in the linkage of the terminal sialic acid have shown these pentasaccharides bind with very different conformations. LSTc (Neu5Acα2-6Galβ1-4GlcNAcβ1-3Galβ1-4Glc) binds in a folded conformation while LSTa (Neu5Acα2-3Galβ1-3GlcNAcβ1-3Galβ1-4Glc) binds in an extended conformation.

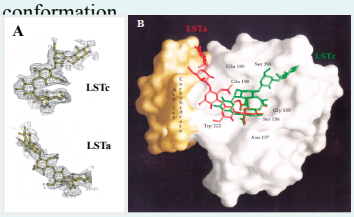


Figure 4: (A) Conformations of the LSTc and LSTa pentasaccharides. **(B)** Binding Of LSTc and LSTa by HA. In each case the terminal Neu5Ac is bound in a comparable orientation but the remaining portions of the oligosaccharide chains exit the binding site in different directions. Images from Eisen M.B. *et al* 1997 Virology **232**:19-31.

The capacity of HA binding sites to accommodate these distinctly different binding conformations of terminally sialylated oligosaccharides is likely to be responsible for different sialic linkage specificity and influenza host ranges.

Work Undertaken For Cloning And Expression Of The HA1 Domain.

The HA1 domain of the Human influenza virus A/Sichuan has been successfully amplified and cloned into an *E. coli* expression vector. This vector incorporates a C-terminal 6HIS tag and places the HA1 sequence under the control of a strong IPTG inducible *ptac* promoter.

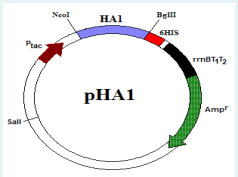


Figure 5: pHA1 expression vector

Work is currently underway to optimise the expression of the HA1 protein in *E. coli* and once this has been completed the purification of the protein will be optimised and its functionality evaluated.

Ultimately this human form of the HA1 gene will be mutated to alter residues Leu226 and Ser228 to Gln226 and GLy228 in order to obtain a HA1 domain with specificity for Neu5Acα2-3Gal .

DCU Bristol-Myers Squibb IDA IRELAND National University of Ireland, Galway Ollscoil na hÉireann, Gaillimh