Structural Insights on the Potential Significance of the Twin Asn-Residue Found at the Base of the Hemagglutinin 2 Stalk in All Influenza A H1N1 Strains: A Computational Study with Clinical Implications

Marni E. Cueno, Kenichi Imai, and Kuniyasu Ochiai

Abstract

Influenza hemagglutinin (HA) is a homotrimeric glycoprotein responsible for binding to sialic acids found in the host cell surface. HA has a prominent 75 Å-long α -helix (HA2 stalk) that contributes to overall HA structural stability. Among the H1N1 strains, a high level of predicted disorder is found at the base of the HA2 stalk, predominantly containing Asn residues. Surprisingly, the significance of Asn residues at the base of the HA2 stalk has not been elucidated. In this study, we analyzed the HA2 stalk base of 2830 amino acid sequences of the influenza H1N1 subtype obtained from human, swine, and avian strains throughout 1918–2012. We detected a structurally conserved twin Asn-residue (N145₂–N146₂) present at the HA2 stalk base in all H1N1 strains. In addition, we found that the twin Asn-residue maintains both a 2.13 Å salt bridge and 11.74 Å 110-helix:B-loop distance measurement in all H1N1 strains studied. Both observations were consistent with known HA crystal structures. Interestingly, amino acid substitutions in either or both residues 145₂ and 146₂ were found to alter these measurements and, likewise, shift the empirical distribution of HA amino acid residues. Thus, we hypothesize that both the N145₂ and N146₂ residues found at the HA2 stalk base in all H1N1 strains provide the necessary structural requirements to stabilize the HA protein. More importantly, this would imply that the twin-Asn-residue is an ideal target for anti-influenza therapies.

Introduction

INFLUENZA A VIRUS is an RNA virus that initiates infection to the host cell by binding to sialic acids found on the host cell surface (Kongchanagul et al., 2008; Skehel and Wiley, 2000). Influenza virus binding to the host cell is primarily attributed to the hemagglutinin (HA) glycoprotein. HA is a homotrimeric glycoprotein that constitutes most of the virus surface (Cherry et al., 2009; Das et al., 2010) and consists of a fibrous globular stem and a globular head containing three sialic acid binding sites (Daniels et al., 2003; Das et al., 2010; Wang et al., 2009).

HA consists of two disulfide-linked polypeptide chains, HA1 and HA2, created by a proteolytic cleavage of the precursor protein (Wiley and Skehel, 1987; Wilson et al., 1981). The HA glycoprotein has multiple roles in the virus life cycle (Daniels et al., 2003) and, during interaction with host cells, causes a dramatic structural reorganization of the HA2 that can affect the 75 Å-long α -helix normally known as the HA2 stalk (Lorieau et al., 2010; Sui et al., 2009; Weber et al., 1994).

The HA2 stalk has previously been shown to play a significant role in stabilizing the whole HA protein (DuBois et al., 2011b). The relationship between protein intrinsic disorder and influenza virulence was earlier reported and a high level of predicted disorder is found at the H1N1 HA2 stalk base (Goh et al., 2009), which predominantly contains asparaginyl (Asn) residues.

Asn is a polar amino acid reported to play an important role in maintaining protein structure among certain amino acid interactions found within a protein (Worth and Blundell, 2010). In addition, Asn was shown to form carbonyl–carbonyl interaction between the side-chain carbonyl and local backbone carbonyls within itself or with a neighboring Asn (Deane et al., 1999), allowing it to form a stacking conformation that is energetically more favorable (Deane et al., 1999; Tsai et al., 2005). Surprisingly, the significance of Asn residues at the H1N1 HA2 stalk base has not been elucidated. Here, we utilized a computational approach to determine the significance of the twin Asn-residue found at the H1N1 HA2 stalk base.

Materials and Methods

Influenza A H1N1 HA amino acid sequences of human, swine, and avian strains from 1918–2012 were obtained from the NCBI website. A total of 2830 amino acid sequences were analyzed. Similarly, H1N1 HA crystal structures were obtained from both NCBI and RCSB Protein Data Bank. In labeling the amino acid residues in the predicted H1N1 HA structures, we followed the numbering scheme as previously published (Wilson et al., 1981) and we designated subscript 1 and 2 to refer to HA1 and HA2, respectively. Predicted HA structures were made using Phyre 2 (Kelley and Sternberg, 2009). Briefly, the Phyre server scans user-submitted sequences against the nonredundant sequence database and, afterwards, a profile is constructed. The secondary structure is predicted by three independent secondary structure prediction programs and the confidence values of each program are averaged. All predicted HA structures were compared with available HA crystal structures for verification and consistency.

The twin Asn-residue (N145₂–N146₂) can be structurally traced to the salt bridge between the 110-helix and B-loop (Rachakonda et al., 2007). We determined the measurements of the salt bridge (R103₁ and E85₂) and the 110-helix:B-loop distance (D98₁ and L89₂) using all H1N1 strains studied. Moreover, we simulated an HA structure with either or both residues 145₂ and 146₂ substituted with Ala to determine how it affected both the salt bridge and 110-helix:B-loop distance measurements. All measurements were obtained by measuring the space between pre-identified amino acid residues in both the salt bridge and 110-helix:B-loop distance measurements using Jmol (Herraez, 2006). Briefly, Jmol is a free open

source applet developed for the interactive display of three-dimensional chemical structures. Moreover, Jmol can also read script files and allow instructions or commands to be applied to the predicted model.

Similarly, to determine the effects of substituting Ala in either or both residues 145₂ and 146₂, the empirical distribution of amino acid residues in the unaltered (N145₂–N146₂) and altered (A145₂–N146₂, N145₂–A146₂, A145₂–A146₂) HA structures were determined by establishing the Ramachandran plot using RAMPAGE (Lovell et al., 2003). Briefly, the RAMPAGE server utilizes density-dependent smoothing of amino acid residues that would show sharp boundaries at critical edges and would indicate a clear delineation between large empty areas and regions that are allowed but disfavored. This would allow us to establish any shift in the empirical distribution and, furthermore, emphasize the structural significance of the twin Asn-residue.

Results

Asn in residues 145₂ and 146₂ favor salt bridge formation between the 110-helix and B-loop

We compared predicted H1N1 HA structures generated from protein sequences obtained from 1918–2012 and found that the N145₂–N146₂ residue is structurally conserved. This was further confirmed using known H1N1 HA crystal structures. In both the HA predicted structures and HA crystal structures, we consistently found a conserved N145₂–N146₂ residue that can be structurally traced to the salt bridge between the 110-helix and B-loop (Fig. 1A), which would

FIG. 1. Conserved twin Asn-residue influences the salt bridge between the 110-helix and B-loop. (A) Representative influenza H1N1 hemagglutinin structure. Encircled region indicates the position of the conserved N145₂–N146₂ residue. Boxed region indicates the position of the salt bridge between the 110-helix and B-loop. (B–E) Salt bridge (R103₁ and E85₂) and distance measurements (D98₁ and L89₂) corresponding to the (B) conserved N145₂–N146₂ residue and amino acid substitutions made in (C) both or (D,E) either residue. 110-helix (violet) and B-loop (blue) are indicated.

Asparaginyl residues are found to occur frequently in irregular regions (Srinivasan et al., 1994) and, among H1N1 strains, this would seem to be a common feature at the HA2 stalk base (Goh et al., 2009). Throughout this study, we were

able to identify a conserved twin Asn-residue found at the H1N1 HA2 stalk base and that amino acid substitutions made in either or both residues 145₂ and 146₂ altered the measurements in both the salt bridge and 110-helix:B-loop distance. This highlights the structural significance of having Asn in both residues 145₂ and 146₂, that is, to maintain the salt bridge between the 110-helix and B-loop.

HA stability is shown to be regulated by salt bridges across the receptor binding domain (RBD) interfaces and that any alteration in HA stability is suggested to influence viral infectivity (DuBois et al., 2011a; Rachakonda et al., 2007; Xu and Wilson, 2011), emphasizing the importance of the salt bridge between the 110-helix and B-loop in stabilizing the RBD (DuBois et al., 2011b). A salt bridge is a combination of two noncovalent interactions (hydrogen bonding and electrostatic interactions) and the most common salt bridges occur from the anionic carboxylate of either aspartic acid or glutamic acid and the cationic ammonium from lysine or the guanidinium of arginine (Kumar and Nussinov, 2002). Salt bridges between the influenza HA 110-helix and B-loop have been correlated to both the stability of the ectodomain and coordination of cooperative motions essential for the fusion process (Rachakonda et al., 2007). This emphasizes the importance of the N145₂–N146₂ residue and, likewise, its role in maintaining the salt bridge between the 110-helix and B-loop.

Similarly, we were able to observe that the empirical distribution of HA amino acids shifted when Asn was substituted with Ala in either or both residues 145₂ and 146₂ favoring an increase in α_L -helix conformations and a decrease in partially-allowed α_R -helix conformations. Asn residues are known to adopt to structural conformations that favor an α_L -helix conformation and other partially-allowed conformations found in the Ramachandran plot more readily than any other nonglycyl amino acids (Srinivasan et al., 1994) ascribable to the carbonyl–carbonyl interaction occurring in their own backbone or with that of the previous residue (Deane et al., 1999). Consequentially, carbonyl–carbonyl interaction produces stabilization energy comparable with hydrogen bonds (Choudhary et al., 2009; Deane et al., 1999). The decrease in the number of partially-allowed α_R -helix conformation of HA amino acids would insinuate that there is less stabilization energy available. Consequentially, this may explain the increase in α_L -helix conformations of HA amino acids since this may partially compensate for the energy requirement needed for protein stability (Deane et al., 1999). However, we emphasize that even though α_L -helix conformations increased, it was not enough to maintain the salt bridge. We hypothesize that Asn in both residues 145₂ and 146₂ creates an energy-favorable HA conformation essential for salt bridge formation between the 110-helix and B-loop which, in turn, contributes to HA stability.

At the moment, there are several proposed anti-influenza therapies based on varying strategies (Palese and Garcia-Sastre 2002; Saladino et al., 2010). Based on our results, we propose that designing influenza vaccines and antiviral strategies that target the twin Asn-residue could have the potential to be another anti-influenza strategy.

Conclusion

In conclusion, we found that from 1918–2012 the N145₂–N146₂ residue is conserved in all strains of the influenza A

H1N1 subtype. We were able to trace the conserved N145₂–N146₂ residue to the salt bridge between the 110-helix and B-loop and, in addition, show that amino acid substitution in either or both residues 145₂ and 146₂ could alter the measurements in the salt bridge and 110-helix:B-loop distance. Furthermore, simulated amino acid substitutions in either or both residues 145₂ and 146₂ were also shown to change the empirical distribution of HA amino acids, resulting in a structural conformation that is less energetically favorable. Thus, we believe that both the N145₂ and N146₂ residues found in the conserved twin Asn-residue located at the H1N1 HA2 stalk base provide the necessary structural requirement needed to maintain and allow salt bridge formation between the 110-helix and B-loop. This, in turn, is necessary to stabilize the overall HA protein. Furthermore, this would insinuate that the twin-Asn-residue could be an ideal target for future anti-influenza therapies.

Acknowledgments

This work was supported by the following grants: Grants-in-Aid for Scientific Research, Grants-in-Aid for Young Scientist (2479210), Nihon University President's Grant for Multidisciplinary Research, Uemura Fund and Dental Research Center, Nihon University School of Dentistry (Tokyo), "Strategic Research Base Development" Program for Private Universities subsidized by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) 2010 (S1001024).

Author Disclosure Statement

The authors declare that no conflicting financial interests exist.

References

- Cherry JL, Lipman DJ, Nikolskaya A, and Wolf YI. (2009). Evolutionary dynamics of N-glycosylation sites of influenza virus hemagglutinin. *PLoS Curr* 1, RRN1001.
- Choudhary A, Gandla D, Krow GR, and Raines RT. (2009). Nature of amide carbonyl–carbonyl interactions in proteins. *J Am Chem Soc* 131, 7244–7246.
- Daniels R, Kurowski B, Johnson AE, and Hebert DN. (2003). N-linked glycans direct the cotranslational folding pathway of influenza hemagglutinin. *Mol Cell* 11, 79–90.
- Das SR, Puigbo P, Hensley SE, Hurt DE, Bennink JR, and Yewdell JW. (2010). Glycosylation focuses sequence variation in the influenza A virus H1 hemagglutinin globular domain. *PLoS Pathog* 6, e1001211.
- Deane CM, Allen FH, Taylor R, and Blundell TL. (1999). Carbonyl–carbonyl interactions stabilize the partially allowed Ramachandran conformations of asparagine and aspartic acid. *Protein Eng* 12, 1025–1028.
- DuBois RM, Aguilar-Yanez JM, Mendoza-Ochoa GI, et al. (2011a). The receptor-binding domain of influenza virus hemagglutinin produced in *Escherichia coli* folds into its native, immunogenic structure. *J Virol* 85, 865–872.
- DuBois RM, Zaraket H, Reddivari M, Heath RJ, White SW, and Russell CJ. (2011b). Acid stability of the hemagglutinin protein regulates H5N1 influenza virus pathogenicity. *PLoS Pathog* 7, e1002398.
- Goh GK, Dunker AK, and Uversky VN. (2009). Protein intrinsic disorder and influenza virulence: The 1918 H1N1 and H5N1 viruses. *Virol J* 6, 69.

- Herraez A. (2006). Biomolecules in the computer: Jmol to the rescue. *Biochem Mol Biol Educ* 34, 255–261.
- Kelley LA, and Sternberg MJ. (2009). Protein structure prediction on the Web: A case study using the Phyre server. *Nat Protoc* 4, 363–371.
- Kongchanagul A, Suptawiwat O, Kanrai P, Uiprasertkul M, Puthavathana P, and Auewarakul P. (2008). Positive selection at the receptor-binding site of haemagglutinin H5 in viral sequences derived from human tissues. *J Gen Virol* 89, 1805–1810.
- Kumar S, and Nussinov R. (2002). Close-range electrostatic interactions in proteins. *Chembiochem* 3, 604–617.
- Lorieau JL, Louis JM, and Bax A. (2010). The complete influenza hemagglutinin fusion domain adopts a tight helical hairpin arrangement at the lipid:water interface. *Proc Natl Acad Sci USA* 107, 11341–11346.
- Lovell SC, Davis IW, Arendall WB, 3rd, et al. (2003). Structure validation by C α geometry: Phi, psi and C β deviation. *Proteins* 50, 437–450.
- Palese P, and Garcia-Sastre A. (2002). Influenza vaccines: Present and future. *J Clin Invest* 110, 9–13.
- Rachakonda PS, Veit M, Korte T, et al. (2007). The relevance of salt bridges for the stability of the influenza virus hemagglutinin. *FASEB J* 21, 995–1002.
- Saladino R, Barontini M, Crucianelli M, Nencioni L, Sgarbanti R, and Palamara AT. (2010). Current advances in anti-influenza therapy. *Curr Med Chem* 17, 2101–2140.
- Skehel JJ, and Wiley DC. (2000). Receptor binding and membrane fusion in virus entry: The influenza hemagglutinin. *Annu Rev Biochem* 69, 531–569.
- Srinivasan N, Anuradha VS, Ramakrishnan C, Sowdhamini R, and Balaram P. (1994). Conformational characteristics of asparaginyl residues in proteins. *Int J Pept Protein Res* 44, 112–122.
- Sui J, Hwang WC, Perez S, et al. (2009). Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. *Nat Struct Mol Biol* 16, 265–273.
- Tsai HH, Reches M, Tsai CJ, Gunasekaran K, Gazit E, and Nussinov R. (2005). Energy landscape of amyloidogenic peptide oligomerization by parallel-tempering molecular dynamics simulation: Significant role of Asn ladder. *Proc Natl Acad Sci USA* 102, 8174–8179.
- Wang CC, Chen JR, Tseng YC, et al. (2009). Glycans on influenza hemagglutinin affect receptor binding and immune response. *Proc Natl Acad Sci USA* 106, 18137–18142.
- Weber T, Paesold G, Galli C, Mischler R, Semenza G, and Brunner J. (1994). Evidence for H(+)-induced insertion of influenza hemagglutinin HA2 N-terminal segment into viral membrane. *J Biol Chem* 269, 18353–18358.
- Wiley DC, and Skehel JJ. (1987). The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Annu Rev Biochem* 56, 365–394.
- Wilson IA, Skehel JJ, and Wiley DC. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* 289, 366–373.
- Worth CL, and Blundell TL. (2010). On the evolutionary conservation of hydrogen bonds made by buried polar amino acids: The hidden joists, braces and trusses of protein architecture. *BMC Evol Biol* 10, 161.
- Xu R, and Wilson IA. (2011). Structural characterization of an early fusion intermediate of influenza virus hemagglutinin. *J Virol* 85, 5172–5182.

Address correspondence to:

Marni Cueno

or

Kuniyasu Ochiai

Department of Microbiology

Division of Immunology and Pathobiology

Dental Research Center

Nihon University School of Dentistry

Tokyo 101-8310

Japan

E-mails: marni.cueno@nihon-u.ac.jp

or

ochiai.kuniyasu@nihon-u.ac.jp