# Review of Enzyme Hemagglutanin Esterase

Soham Samajpaty – Department of General Medicine, International Faculty, Russian National Research Medical University named after N I Pirogov

## INTRODUCTION

Although widely neglected in the world of medicine and research and development programs, enzyme hemaglutanin esterase is a domain of utmost importance in Virology and developmental biology of virus. Without this enzyme many virus are functionally cripple which amplifies its importance manifold. The enzyme facilitates binding of a huge range of virus to its host receptor, change in kinetics of this grossly affects such function. The enzyme is widely found in *influenza C virus, toroviridae* and *coronaviridae* groups. Primarily consisting of three functional roles; receptor binding activity, receptor hydrolysis activity and membrane fusion activity, it's an indispensable part of a functional virus. The enzyme has wide-scale implications. It has a very complex structure, varying from one virus to another. Not enough light has been ushered on the chemistry of this enzyme but available literature indicates the complex chemical functioning of it as well. It's presence in wide range of viruses affecting animal worlds including humans and wild species alike, makes it a necessary topic for intense discourse of research. Structural complexity of each domain of HE is such that many could not have been resolved by spectroscopy as well. In brief, enzyme Hemagglutanin Esterase is the "energy shutting system" which enables virus to infect its host species and ensures its sustainability.

## KEYWORDS

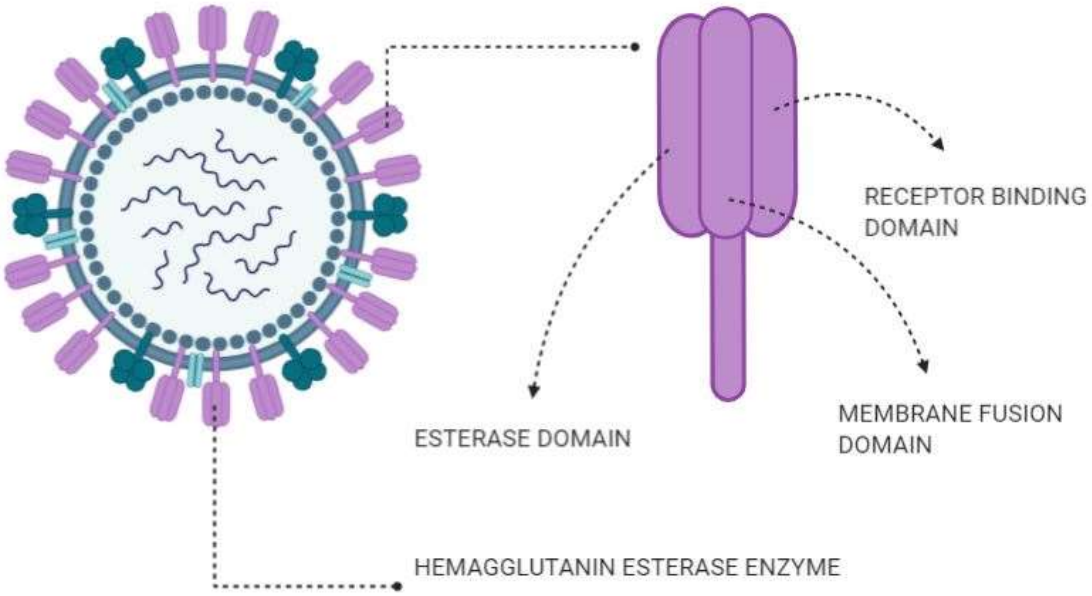
Heemagglutanin Esterase, Coronaviridae, COVID19, virology, developmental biology

## METHODS

The research is based on wide literature review on the chemistry of the enzyme Hemagglutanin esterase and the virology prospect of the same. Structure, evolution and significance is the main domain of the research. Attempt has been made to summarize the nature and chemistry of HE with special attention to the domain of virology and developmental or evolutionary biology.

## THE ENZYME

Hemagglutanin esterase is a fusion-glycoprotein (HEF) attached to the viral envelop of several virus such as *toroviridae, influenza C virus* and *coronaviridae*. Hemagglutanin and Neuraminidase proteins are found in many *influenza* strains but HEF is only seen in single spike virus. It is a trimer with three domains; membrane fusion domain, esterase domain and receptor-binding domain. The binding domain attaches to N-acetyl-9-O-acetylneuraminic acid of glycolipids and glycoproteins of the virus host cell. This is followed by receptor hydrolysis of the host cell by the esterase domain, helping escape of the viral particle. By the time the viral genome has been incorporated inside the host cell. For the purpose of the virus being infectious, the HEF is cleaved by trypsin-like proteases producing $HEF_1$ and $HEF_2$. The receptor-binding domain is present at the surface loop of esterase domain and the esterase domain is present in the loop of membrane domain




(https://www.ebi.ac.uk/interpro/entry/InterPro/IPR003860/). Although from normal perspective of molecular biology it might be assumed that acylation of HEF is determining factor for its action but it has been proven experimentally in *influenza virus* that HEF doesn't meet the criteria – acylation doesn't affect the cleavage of HEF into $HEF_1$ and $HEF_2$. Moreover it has shown better cleavage property at lower pH in fluorescence study, making acidic environment a more suitable ground for thriving of the virus carrying this enzyme. Although it is to be mentioned that acylation or de-acylation doesn't confer any change in the pH dependence for conformational change of HEF (https://onlinelibrary.wiley.com/doi/full/10.1111/cmi.12541).

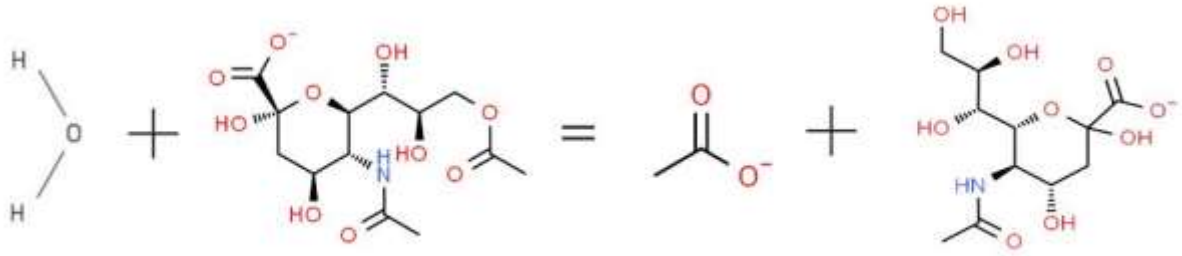
Evolutionarily RDEs of *orthomyxoviridae* are also related to hemagglutanin esterase, But it is to be noted that HEs are only found in positive stranded RNA virus like *influenza C and corona virus*. Despite, of biotechnical research discourse, the source of HE is coronavirus still remains a mystery. The HE is manifested significantly in bovine coronavirus (BCoV), SARS-nCoV and COVID19. Influenza C and torovirus maintains the same virological characteristics and even has been reported in infectious salmon anemia virus (ISAV). Apart from minor sequential dissimilarities, HE structure among the *coronaviridae* family is maintained. But differences are noteworthy when compared with *nidoviridae* and most of the changes are seen in receptor domain R with amino acid substitutions, frame-shifts, insertions and deletions. In nidovirus HE protein – the single cystein residue $Cys^6$ at $HEF_1$'s N-terminus which made disulfide bond with $Cys^{137}$ of $HEF_2$ is absent. Again in coronavirus HE protein, bonds between $Cys^{106}$- $Cys^{151}$ and between $Cys^{196}$- $Cys^{296}$ are absent. Whereas, in torovirus, the C7-C10 bond has been replaced by disulfide bond between C10 and newly formed cystein residue C9a. Such conformational differences ensure difference in binding site specificity of different virus having this RDE (https://link.springer.com/article/10.1007/s10719-006-5438-8#rightslink). Just like any other chemical compound, enzyme HE is also morbid. Experimental data on enzyme kinetics of HE of influenza C virus has perfectly shown that the enzyme is susceptible to destruction by serine protease type of chemical compounds. Isocoumarins - 3,4- dichloroisocoumarin, 3-chloroisocoumarin, and 4-chloro-3- [(3-isothioureido)propoxy]-isocoumari, shows significant activity towards inhibition of HE in influenza C virus. The experimentally observed amino acid sequence of - (X-G-A-S-V-L-X-Q-S-X-X-I-G-F-X-X-X-X) was studied for time-dependent kinetics with these isocoumarins. The 3,4-dichloroisocoumarin appeared as the most potent inhibitor with half-life of

inactivation in second order reaction as 410 per s per M (https://jvi.asm.org/content/63/5/2056.short). In influenza C virus, the 4th segment among the 7 segments of genes code for hemagglutinin esterase. Hence, there has to be a specific codon in RNA of virus which codes for HE. Apart from absence of HEF subunit 2, in HEF subunit 1 there is complete functional similarity with 2019-nCoV. Studies based on electron microscopy and X-ray crystallography has shown on influenza that the HEF has a mushroom like structure with globular head and a stalk (https://www.pnas.org/content/105/26/9065). In the Bovine coronavirus it is a 140kDa structural glycoprotein. The in-vitro expression sequence of HE is BCoV begins with CTAAAC consensus sequence, located 15bp upstream from the start codon ATG and ending 343bp downstream of TAG termination codon. On the basis of N-terminal signal peptide, each viron anchored HE predictably has N-terminal ectodomain and C-terminal anchor. It has also been observed that HE in BCoV when infected in host cell becomes rapidly disulfide-linked and glycosylated and most of it is associated inside the viron particle in the ER assembly line and *cis*-Golgi. HE although still remain highly under-evaluated as many virus contains it in vestigial form or currently in inactive state as in case of hepatitis coronavirus A59 (https://jvi.asm.org/content/64/4/1834.short). However, the HE in human coronavirus has undergone a significant change. Human beta1-coronavirus OC43 and HKU1 has undergone mutation in the HE lectin domain, resulting into loss of viron associated receptor destruction activity and thus, some clustered receptor population are not cleaved. Separation of deep hydrophobic pocket of $P_1$ from $P_2$ is done by side chain of $F^{211}$ which is a residue in $\beta_{12}$/ $\beta_{13}$ in the β-hairpin. Thus, ligand binding largely depends on shape complementarity and hydrophobic interactions, stabilized by protein-sugar interaction. But substitution in the $T^{114}N$ substitution in OC43 HE has generated a novel N-glycosylation site (TTS NRS). Experimental data have clarified that glycosylation of $N^{114}$ reduces the affinity of HE lectin domain for 9-O-Ac-Sias. Subsequent data of GenBank collected from USA, Western Europe and China in early 1990s have shown that both type A and type B HEs have undergone many mutations in addition to those present already in OC43. But surprisingly in type B HEs $N^{114}$ glycosylation is lost again due to $S^{116}A$ substitution (NRS NRA) (https://www.sciencedirect.com/science/article/pii/S1931312817300707). Already five coronavirus of zoonotic origin have established their hosts among humans – OC43, HKU1, SARS (2002), MERS and SARS-nCoV19. SARS-nCoV19 although has become a big headache for human survival and livelihood, the HE of it remains highly neglected from the research world and co-relation between S-protein and HE is very feebly understood. It's is an immense disaster of 21st century but there is huge lack of literature on HE of SARS-nCoV19. Although, gene of HE in this case has been identified as EYW02-gp3 (https://www.ncbi.nlm.nih.gov/gene/?term=Human+coronavirus+OC43+hemagglutinin-esterase). On the other hand, *toroviridae* which although have structural similarity with S-protein and viral envelope, has huge difference in sequence. But, HE-like sequence has been identified in torovirus Berne Virus (BEV). An isolated BEV in culture medium has been experimentally studied for this purpose. Translation of a non-functional open reading frame (ORF), termed as ORF4 yielded amino acid sequence with 30% HE similarity. Furthermore, evidence of complete presence of HE genomic sequence in bovine torovirus has been found. When, a sucrose-gradient purified BoTV strain Breda was prepared from experimentally infected cattle feces in an experiment and the obtained viral RNA was taken source for template RT-PCR with oligoneucleotide derived BEV sequence. cDNA amplification by covering 3'-most and 3kb of the BoTV genome was cloned into pGEM-T vectors and was sequenced. The result after ORF translation showed that torovirus contains a HE protein is a 65K protein in the experimental model (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC191764/). Enzyme Hemagglutanin esterase has been studied in laboratory setup by using GMOs and have been optimally found to be expressed in many virus apart from the mentioned ones. Murine Hepatitis virus and Murine Coronavirus are live examples of successful laboratory expression of HE under specific conditions.

**THE CHEMISTRY**

Enzyme Hemagglutanin esterase is a trimeric glycoprotein, essential for a functional viral infection in host cell. Although, in culture medium not an absolute necessity. As we have discussed, it has three domains - receptor binding domain, membrane fusion domain and esterase enzyme domain. The fusion protein HEF after infecting the host cell breaks into $HEF_1$ and $HEF_2$. It binds to the N-acetyl-9-O-acetylneuraminic acid of glycolipids and glycoproteins of the virus host cell. A generalized chemical reaction for the receptor binding purpose to clarify the catalytic activity of the enzyme, keeping the substrates on left hand side and products on the right hand side can be written as;-

**$H_2O$ + *N*-acetyl-9-*O*-acetylneuraminate = acetate + $H^+$ + *N*-acetylneuraminate**



(https://www.uniprot.org/uniprot/Q9Q9G3).

Analysis of the chemical dynamics of *N*-acetyl-9-*O*-acetylneuraminate offers a deep introspection into the functioning of HE, as it's the reactive substrate of the binding domain. A CVFF electric force field study to determine the conformational changes and behavior of $GM_3$ gangliosides of *N*-acetyl-9-*O*-acetylneuraminate has been studied experimentally in relation to NMR data, in this regard. Non-acetylated derivative $G_{D1\alpha}$ of *N*-acetyl-9-*O*-acetylneuraminate separately – (I) Neu5,9$Ac_2$α2-3Galβ1-3GalNAc; (II) Neu5,9$Ac_2$α2-3Galβl-3GalNAcβ1-4Gal; (III) Neu5,9$Ac_2$α2-3Galβ1-3GalNAcβ1-4('Neu5Acα2- 3)Gal and (IV) Neu5,9$Ac_2$α2-3Galβ1-3GalNAcβ1-4('Neu5Acα2-3)Galβ1-4Glc **;** were obtained. The Neu5,9Ac2a2-3Gal linkage in fragments (I) and (II) shows transition between two distinct confirmations as deduced from the two sets of Ψ and Φ segments of the same. From Ψ and Φ angle trajectories, it was found flexibility of Neu5,9Ac2a2-3Gal in fragment (II) is more than that in fragment (III). But such change in flexibility is insignificant between the fragments (III) and (IV). As far as the chemical reactivity is considered, the fragments (I), (II) and (III) does not show any significant change with solvent medium. But Ga1β1-4Glc linkage within fragment (IV) changes to various angles of Ψ and Φ depending on the type of solvent and salvation sphere. Fragment (IV) of $G_{D1\alpha}$ interacts between different spatially separated fragments of it (https://pubmed.ncbi.nlm.nih.gov/8922951/).

**THE PROSPECT NOVEL USE IN THERAPEUTICS**

As, we understand from the above discussion, enzyme hemagglutanin esterase can be of utmost importance in developing clinical therapeutics. Virus possessing this enzyme functional only by virtue of its enzyme activity. It is very precise that without reduction of activation energy no chemical reaction can be facilitated. The enzyme HE brings down activation energy of reaction of S-protein binding domain with ACE-2 receptor in coronavirus

(for example) to reaction feasible activation energy. Hence, complete destruction of HE might be a novel approach towards a new generation of antibody production. Monoclonal or polyclonal antibodies targeting HE and thus inhibiting its function will leave the virus in functionally crippled state. Mutations in HE are rarely visible in between different viral groups and even if noticed, then with very minute difference in structure, a wide discussion on this has also been done above. Hence, HE enzyme is a novel target for developing future therapeutics, not only against one but a wide range of virus and different mutant strains of the same virus.

**DISUCUSSION**

Enzyme Hemagglutanin Esterase is a huge domain of research in medical therapeutics, general virology and developmental biology of viral species. From this review literature we might confront the wide range of spectrum in which HE enzyme is a functional domain of immense importance. Although it is evident that researches conducted on this topic is very less and futile. Understanding of the chemical and biological character of it is highly underdeveloped. Considering the evolution side it might be stressed upon that the enzyme has manifested itself initially in *toroviridae* and have then taken over the infection mechanism of *coronaviridae* and *influenza C* virus. Although there are structural dissimilarities which are widely evident between classes of virus and even sub-species of it. There are noteworthy changes in inter-carbon linkage and disulfide-bridges in HE among different viruses. But from epidemiology point of inspection it is worthy to be declared that HE has an indispensible role to play in coronaviruses and influenza viruses to properly infect and manifest itself in host cell. Without HE they are functionally cripple. Few laboratory experiments have also shown probable presence of HE in influenza A and D viruses as well. Hence, deep research in laboratory scale is needed on HE from the virology and chemistry point of view. In future such research might harness mind-blowing developments in the field of medicine.

**REFERENCES**


- https://www.ebi.ac.uk/interpro/entry/InterPro/IPR003860/
- https://onlinelibrary.wiley.com/doi/full/10.1111/cmi.12541
- https://link.springer.com/article/10.1007/s10719-006-5438-8#rightslink
- https://jvi.asm.org/content/63/5/2056.short
- https://www.pnas.org/content/105/26/9065
- https://jvi.asm.org/content/64/4/1834.short
- https://www.sciencedirect.com/science/article/pii/S1931312817300707
- https://www.ncbi.nlm.nih.gov/gene/?term=Human+coronavirus+OC43+hemagglutinin-esterase
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC191764/
- https://www.uniprot.org/uniprot/Q9Q9G3
- https://pubmed.ncbi.nlm.nih.gov/8922951/