Tutora
Dra. Carme Rovira Virgili
Departament de Química Inorgànica i Orgànica

Grau de Química

Treball Final de Grau

Neuraminidase superfamily compendium.
Compendio de la superfamilia de las neuraminidasas.

Alba Recio Sierra

January 2022

Universitat de Barcelona
B:KC Barcelona Knowledge Campus
Campus d'Excel·lència Internacional

An expert is a person who has made all the mistakes that can be made in a very narrow field.

— Niels Bohr

En primer lloc, agrair-li a la meva Tutora Carme Rovira, tot el temps que ha dedicat a ajudar-me i guiar-me durant aquesta última etapa del grau.

Gràcies també a la meva família, amics i a la meva parella per animar-me i recolzar-me als moments bons i no tant bons que ha tingut la carrera. En especial a tu papa, per ensenyar-me a superar-me a mi mateixa i llevar-me encara més forta quan sembla que no hi ha sortida.

REPORT

CONTENTS

1. SUMMARY

The neuraminidase activity was described for the first time in 1942 by the virologist George Hirst in the influenza virus. Indeed, neuraminidase types define flu subtypes nowadays and are of major relevance in the development of modern flu vaccines. Although neuraminidases have been known for a long time, they continue to be the focus of much current research because of their potential for future drug development and their close relationship with many essential biological processes. As other carbohydrate-active enzymes, they are classified into families according to sequence similarity in the Carbohydrate-active enzymes database (CAZy). However, it is necessary to resort to further information sources to learn about many relevant aspects of each neuraminidase. Taking advantage of the large amount of information that has been collected in the literature in recent years, this project is a compendium of all neuraminidases from which a 3D structure is available and proposes a new classification based not only on structure but also function, location in Nature and organisms, as well as pathogenesis of each neuraminidase.

Keywords: Neuraminidase, sialidase, sialic acid, neuraminic acid, N-glycosylneuraminic acid, glycosyl hydrolase

2. RESUMEN

La actividad enzimática de las neuraminidasas se describió por primera vez en 1942 por el virólogo George Hirst en el virus Influenza. De hecho, hoy en dia los tipos de neuraminidasa constituyen los distintos subtipos de influenza virus y por tanto son de gran relevancia en el desarrollo de las vacunas modernas contra la gripe. A pesar de que hace ya mucho tiempo que se conocen estas enzimas, las neuraminidasas continúan estando en el punto de mira de muchas investigaciones actuales debido a su potencial para el desarrollo de futuros fármacos y su estrecha relación con una gran cantidad de procesos biológicos esenciales. De la misma forma que sucede con otras enzimas con actividad en los carbohidratos, las neuraminidasas estan clasificadas en distintas familias en la base de datos *Carbohydrate-active enzymes database (*CAZy) en función de la similitud de seqüencias proteinogénicas y se debe recurrir a otras clasificaciones a la hora de aprender a cerca de las funciones que realiza cada una de las neuraminidasas. Aprovechando la gran cantidad de información que se ha recogido en distintos artículos en los últimos años, en este proyecto se elabora un compendio de todas las neuraminidasas de las cuales se conoce la estructura cristalina y se propone una nueva clasificación no solo tiniendo en cuenta la estructura de la enzima, sino en base a la función de la neuraminidasa, el tipo de organismo del que forman parte y la relación con posibles enfermedades.

Palabras clave: Neuraminidasa, sialidasa, ácido siálico, ácido neuramínico, ácido N-acetilneuramínico, glicosidasa

3.INTRODUCTION

3.1. NEURAMINIC ACID

The main ligand (substrate) of neuraminidases is neuraminic acid, also named sialic acid. It is a nine-carbon monosaccharide that is present in the different branches of the tree of life. There are more than 50 sialic acid variants. The most common and the best-known compound is N-acetylneuraminic acid (Neu5Ac or NANA) or named using IUPAC nomenclature: 5-(acetylamino)-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranosonic acid (see Figure 1).

Other common derivatives, less abundant, are N-glycosylneuraminic acid and 5-hydroxyneuraminic acid.

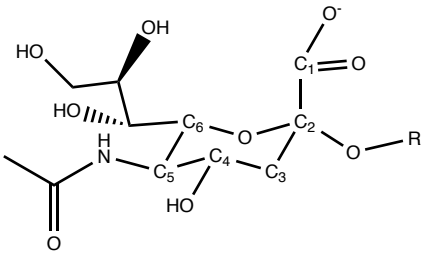


Figure 1: N-acetylneuraminic acid structure

Neuraminic acids are normally located terminally on surfaces of cells or macromolecules but can also be located within. These sugars are natural parts of lipids and glycoproteins but can also be found in membrane proteins such as insulin receptors, cell membranes, mucopolysaccharides, ... They are deeply engaged in fundamental biological processes for the maintenance of life but in a different way than other carbohydrates. The involvement of neuraminic acids in biological processes include the functions of mediators in intercellular communication, mediators in cell-cell adhesion, cell renovators, receptors of viruses and bacteria among others. Terminal neuraminic acids determine the survival time of erythrocytes, thrombocytes, and leukocytes.

An example of the importance of neuraminic acids refers to the inflammation process. The interaction between selectins (vascular adhesion molecules on endothelial surfaces) and neuraminic acids slow down the diffusion of leukocytes mediating this way the entrance of leukocytes into damaged tissues to initiate an immune response.

Another example is neuraminic acid roll into the cell renewal process. Erythrocytes (cells that form the bone marrow) pass to the blood developing an oxygen transport function to all body tissues. Neuraminic acid residues in the extremities of erythrocyte membranes protect these cells. The percentage of loss of the neuraminic acids through the maturation of erythrocytes, determine their age. Furthermore, the complete loss of them exposes galactose residues implying erythrocytes are ready for destruction and replacement.

Probably the most interesting biological process mediated by neuraminic acids refers to the infection of humans by viruses and bacteria. This issue has been the target of lots of studies, and most of them are still in progress. The carbohydrate is involved in the entrance and spread of the pathogens through cells. The better understanding of the role of neuraminic acids in the whole process has opened up new possibilities for the development of new drugs and progress into the study of diseases such as influenza virus for example.

In Table 1 there's an overview about some of the functions and locations of neuraminic acid in mammals.

Biological function	Location
Charge carrier	cells, blood cells, tumour cells, gangliosides (central nervous system), glycoproteins
Receptors	serotonin, alkaline polymers, tetanus toxin, viruses
Part of antigens	blood groups, bacteria
Masking of antigens	antigens of erythrocytes, transplantation antigen, tumour antigen, normal cell antigen
Biological activity of hormones	intrinsic factor (vitamin B12 absorption), choriogonadotropin (women infertility), erythropoietin (anaemia)

Table 1: Overview about some of the functions and locations of neuraminic acid in mammals

4. OBJECTIVES

Nowadays there is a large amount of scientific articles that study a single neuraminidase and discuss the importance of the enzyme in a biological process. It's easy to find information about

any neuraminidase, its structure and how the understanding of the enzyme opens up in most cases new opportunities in the world of research and development of new drugs or better understanding of a disease. There is many information on a given neuraminidase [44] [45] [46] [59] or a group of this enzymes [14] [41] [48] , but it's difficult to find a review or an article that gathers information not only on neuraminidase structures but also on their function, their mechanism, which biological kingdom they belong and which diseases are they related. In view of the relevance of neuraminidases in chemistry, biology, and medical sciences, we have built a compendium of all neuraminidases for which the molecular structure is available and propose a classification that gathers all structures of known neuraminidases and separates them according to the kingdom to which they belong and the biological process with which they are related. Also, at the same time this new classification is related with other well-known classifications in order to achieve a global vision of the neuraminidase superfamily and the importance it in life.

5. METHODS

Based on the scope and the objective of this project, a neuraminidase structure research was proposed based on two databases: Protein Data Bank [18], CAZy database [16]. These two databases had been used as a source crystal structures and has helped to acquire general information about the group of enzymes for the proposal of the new classification of neuraminidases. The crystal structures research had been carried out during the two last weeks of September 2021. Bearing in mind neuraminidases have got into the spotlight of current research, the information collected in the project could be outdated, so it is recommended to consult the updated databases mentioned in parallel with this project to contrast the information.

Also, the bibliographic research has been done using three scientific databases: Scifinder, PubMed and Web of Science. Each database mainly has information about a specific neuraminidase for a given organisms and their potential biomedical application. About 100 articles have been consulted and 62 references have been used in this project. The main keywords used were: neuraminidase, active centre, active site, sialidase and the and the specific name of the neuraminidase enzyme (e.g. influenza neuraminidase). The methodology employed is described in more detail in section 8.3.

6. WHAT ARE NEURAMINIDASES?

6.1. FUNCTION

Neuraminidases belong to a class of enzymes named as glycosyl hydrolases (GHs, also referred to as glycosidases). GHs catalyse the hydrolysis of the glycosidic bond of in carbohydrates and glycosides. Some studies reveal that the spontaneous hydrolysis of the glycosidic bond it's a surprisingly slow reaction. In fact, it is estimated that the half-life for the hydrolysis of a cellulose glycosidic bond is about 4,7 million years. [40] Glycosidases can accelerate the spontaneous reaction by rates of up to 10^{17} being this way one of the most powerful known enzyme catalysts. [19] [29]

Neuraminidases (also named sialidases) are a type of glycosyl hydrolase.[38] The hydrolysis of the glycosidic bond in neuraminidases results in the breakage of the C-O bond between neuraminic acid and another monosaccharide.[29] In general, neuraminic acid is the terminal sugar in a polysaccharide or complex carbohydrate [1] (see the Figure 2).

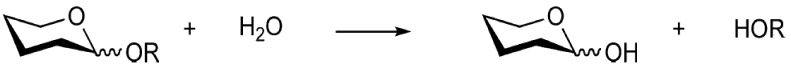


Figure 2: Glycosyl hydrolase function originally published in Cazypedia [17]

6.2. CLASSIFICATION OF GLYCOSIDASES

Neuraminidases and therefore glycosidases, can be classified in different ways:

6.2.1. Classification by cleaving ability: exo/endo

A glycosidase, also named glycoside hydrolase, is exo- or endo depending on the position of the cleavage done on the glycoside chain. Exo-glycosidases are glycoside hydrolase enzymes that cleave the glycosidic linkage of a terminal monosaccharide in an oligosaccharide or polysaccharide (the so-called non-reducing end). The endo ones cleave the chain between two residues [12] (see Figure 3). Since neuraminic acid is commonly found at the non-reducing end of carbohydrates, neuraminidases are in general exo-acting.

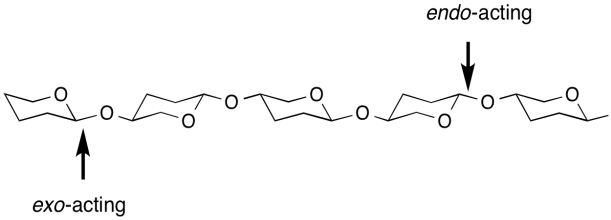


Figure 3: Cleaving ability of glycosidases *originally published in Cazypedia* [17]

6.2.2. Classification by mechanism: retaining/inverting

Depending on the way the scission of the glycosidic bond is done we can differentiate between retaining and inverting glycosidases. Retaining glycosidases keep the configuration (α or β) of the anomeric carbon. Inverting glycosidases change the configuration of the anomeric carbon transforming an alpha reagent in to a beta product and the other way around [17] (see Figure 4).

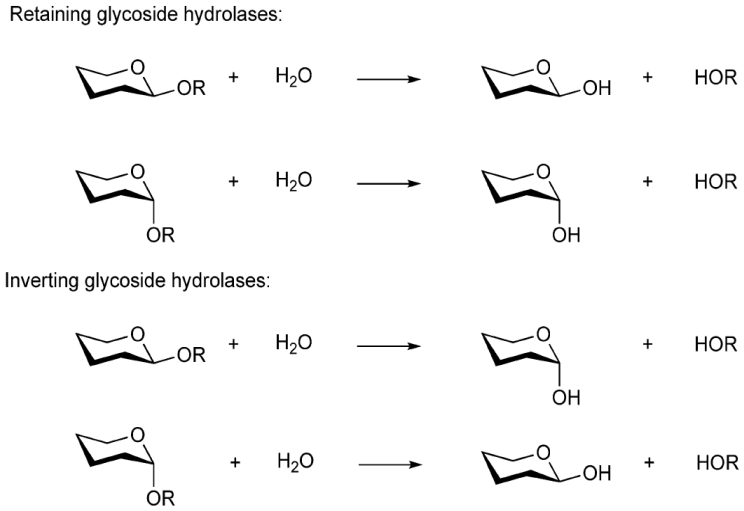


Figure 4: Mechanism retaining and inverting mechanism glycoside hydrolases, *originally published in Cazypedia* [17]

6.2.2.1. Glycosidases retaining mechanism

The retaining mechanism occurs via a two-step, double-displacement mechanism involving two oxocarbenium ion-like transition states. In the first step, named as glycosylation, a glycosyl-enzyme intermediate is formed through the attack of the nucleophile to the anomeric carbon of

the saccharide containing the scissile C-O bond. The reaction is assisted by the other residue acting as an acid catalyst that protonates the glycosidic oxygen as the bond cleaves.

In the second step, the deglycosylation step, the intermediate is hydrolysed by water molecule, which gets deprotonated by the acid/base residue - now acting as a base catalyst - as it attacks the anomeric centre, achieving the second transition state and resulting in the unbinding of the glycoside from the enzyme, forming the final product. Appendix 1 shows the retaining mechanism for an α- or β-glycoside:

A secondary reaction may take place when the acceptor of the second reaction step is not a water molecule but another sugar, which attacks the anomeric carbon via one of its hydroxyl groups. In this case, the reaction leads to the synthesis of a new carbohydrate. This reaction is named as transglycosylation. In most GHs this reaction leads to low product yields. However, some GHs (named as trans-glycosidases) have evolved to make transglycosylation as their main product (see Appendix 1).

6.2.2.2. Glycosidases inverting mechanism

The inverting hydrolysis occurs via a one-step mechanism. The nucleophile and acid/base residues are strategically positioned in the active site (see Appendix 1). One water molecule is positioned in such a way that a general acid and a general base assistance can succeed at the same time. The deprotonation of the water molecule, the formation of the new O_{water}-C1 bond and the cleavage of the C1-O glycosidic bond occurs simultaneously, resulting in a stereo-inversion of configuration upon hydrolysis. The general acid and general base residues are normally glutamate or aspartate, depending on the glycosyl hydrolase, but the identity of some of them is still unclear nowadays for some enzymes.[29] [38] See in Appendix 1 the inverting mechanism for an α- or β-glycoside.

6.2.2.3. Comparison between retaining and inverting mechanisms

In both retaining and inverting mechanisms, the enzymatic hydrolysis of the C-O bond takes place through general acid catalysis that demands two critical residues (the catalytic residues): a nucleophile or general base and a proton donor. In glycosidases these two roles are normally played by residues having carboxylate groups. In either case, retaining or inverting mechanisms, the orientation of the proton donor is identical. However, they differ in the distance between the two catalytic residues: [38][29] The two residues are at an average of 5.5 Å in the retaining enzymes, while they are separated 10 Å in the inverting ones (Figure 5). [29][38] This distance difference is consistent with what could be expected for both mechanisms. In retaining glycosidases, the nucleophile must be positioned sufficiently close to allow the formation of a covalent bond in the intermediate. However, in the inverting glycosidases the distance between the side chains must be wide enough to fit the glycoside and the water molecule [29].

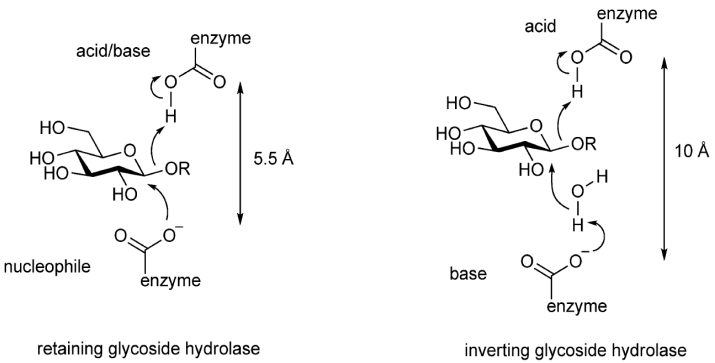


Figure 5: Comparison between residue distances in retaining and inverting glycosidases, *originally published in Carbohydrates: the essential molecules of life* [29]

7. NEURAMINIDASE MECHANISM

The neuraminidase catalysis mechanism is similar to the one originally proposed by Dan Koshland in 1953 for glycosidases. As mentioned above, most glycosidases have carboxylate groups as acid/base and nucleophile residues involved in the catalysis. However, neuraminidases use a tyrosine as the catalytic nucleophile, and this tyrosine is activated by a nearby base residue (normally an aspartate or a glutamate). [38] The tyrosine gives different properties to the enzyme that help catalysing reactions in which the use of a negatively charged carboxylate as the nucleophile would be unfavourable. For example, in a reaction in which the anomeric centre has a negative charge nearby (such as the carboxylate substituent of neuraminic acid), the negative charge of a carboxylate nucleophile would cause repulsion and disadvantage for catalysis. In this case a tyrosine, that is a neutral residue, supported by a suitably positioned general base to enhance it's nucleophilicity by deprotonating the OH group, will promote catalysis. [29] Neuraminidases cleave the glycosidic bonds from various sialoderivatives, such as glycoproteins, glycolipids and oligosaccharides, forming an hemiacetal or hemiketal sugar depending on the initial saccharide and leading the corresponding free aglycon.

Globally, all neuraminidase mechanisms (retaining, inverting, and intramolecular) follow the following four steps:

i) Binding of neuraminic acid to the active site by an arginine triad (a key feature of the active site pocket and highly conserved across all known neuraminidases) [36][37]

ii) Stabilization of the reaction transition state by employing a nucleophile pair Tyr/Glu or Tyr/Asp

iii) Acid/base catalysis assisted by the aspartic acid residue

iv) Deglycosylation by an active site water molecule.

7.1. NEURAMINIDASE RETAINING MECHANISM

When talking about catalytic mechanisms through which neuraminidases react, we can differentiate, as for the glycosidases, between retaining and inverting mechanisms. Most neuraminidases cleave terminal neuraminic acid saccharides using a similar set of active site residues. These include a tyrosine acting as the nucleophilic catalyst, two aspartate and/or glutamate residues that act as general acid or base (one protonates the glycosidic oxygen and the other one deprotonates the catalytic Tyr), and three arginines that are supposed to be associated to the binding of the neuraminic acid carboxylate, as explained before. [36] [37]

The hydrolysis reaction takes place in two steps, with one intermediate, as in classical glycosidases. In the first step, the glycosylation, a glycosyl-enzyme intermediate is formed through the attack of the tyrosine hydroxyl oxygen atom to the neuraminic acid anomeric carbon. Simultaneously, the tyrosine is deprotonated by its neighbour glutamic/aspartic acid. The reaction is assisted by the acid/base residue (either an aspartate or a glutamate), acting as an acid catalyst that protonates the glycosidic oxygen as the bond cleaves. The bond breaking energy could be compensated when the negatively charged carboxylate side chains introduce favourable electrostatic interactions with the positive anomeric charge developed at the transition state (with pyranose boat conformation). [47] In this step there is a first inversion of the configuration. In the example, the carboxylate that was in axial position in the first state is now in the intermediate in equatorial configuration line in an S_N2 reaction.

In the second step, the deglycosylation, the intermediate is hydrolysed by water with the other residue now acting as a base catalyst deprotonating the water molecule as it attacks the anomeric centre achieving the second transition state, again with boat configuration, and resulting in the unbinding of the enzyme and the glycoside, forming the final product. In this second step a second inversion of configuration occurs resulting in an overall retention of configuration at the anomeric centre. Between steps, we can see how the third carboxylate residue helps the reaction by stabilizing the oxocarbenium-ion in the transition states and enhancing the nucleophilicity of the tyrosine through the reaction. As with retaining glycosidases, when the acceptor of the second reaction step is not a water molecule but another sugar, the reaction leads to the synthesis of a new carbohydrate and the reaction is named as transglycosylation, in this case trans-sialylation (see Appendix 1).

7.2. NEURAMINIDASE INVERTING MECHANISM [28]

Inverting neuraminidases are a minority group within all the neuraminidase superfamily, and have been discovered later than most of the retaining examples. They are all in the family GH58. Some studies reveal that the discovered neuraminidases that catalyse inverting reactions also are endo-acting. Therefore, instead of terminal neuraminic acids, the enzymes hydrolyse polyneuraminic acids.

When talking about important residues in their active centres, there is still a lot of questions, but some studies reveal that there are some significant differences between retaining exo-acting neuraminidases and these relatively new endo-acting group. The inverting neuraminidases that

have been characterized until now do not keep the entire active centre, missing some of the most representative residues of the mechanism such as the nucleophilic tyrosine. [28] [47] So, there are still some questions to answer such as which residue takes the nucleophile role or which amino acids are acting as the nucleophile assistant, assuming that the analogy with retaining neuraminidases would keep sense in this case. Some studies have speculated about Glu 581 being a key residue in endoNF neuraminidase. Mutagenesis experiments in this position show a significant drop in the enzymatic activity, being the mutant AE581A catalytically inactive. In spite of these findings, it's still unclear how catalysis of the inverting mechanism takes place at the molecular level.

7.3. NEURAMINIDASE INTRAMOLECULAR MECHANISM [41] [46]

These neuraminidases are named anhydro-sialidases. Their mechanism shares many features with the retaining mechanism. It also starts with the glycosylation and the formation of the intermediate. In contrast with retaining and inverting neuraminidases, in the second step the source of the nucleophilic hydroxyl is not a water molecule but the neuraminic acid 7-hydroxyl group, resulting in 2,7-anhydro-Neu5Ac instead of Neu5Ac. Another difference is the position of a Trp residue, with a voluminous side chain located at the active site. This tryptophan is near the tyrosine and could present hydrophobic interactions, which could be responsible for the altered reaction pathway that results in the interaction between the hydroxyl group of the neuraminic acid glycerol and the anomeric carbon. Also, this hydrophobic interaction could be the reason for the anhydro-sialidase specificity for the α-2,3 linked substrates (see Appendix 1).

8. CLASSIFICATION OF NEURAMINIDASES

8.1. CAZY CLASSIFICATION

The Carbohydrate Active Enzymes database (CAZy database) classifies structurally-related catalytic and carbohydrate-binding modules (or functional domains) of enzymes that degrade, modify, or form glycosidic bonds in carbohydrate-active enzymes. Glycosidases are the most abundant types of carbohydrate-active enzymes. Currently there exists 172 GH families plus a non-classified glycosidase family. This number is not static, updates as research progresses. The updated classification database is available through the CAZy website. [16]

A particularly useful feature of the sequence-based classification is that it almost always results in classification of glycosidases with identical enzymatic mechanisms (inverting or retaining) into the same family. Another advantage of this classification is that if specific functions can be assigned to individual residues in the active site (acid/base, catalytic nucleophile, or others), then, owing to sequence similarities inherent in that family, it is possible to find conserved or equivalent amino acids in other members of the same family, and it is therefore likely that similar catalytic roles are undertaken by these analogous residues. Predictions should always be made with care and remain faithful to experimental confirmation, but the CAZy classification can also be used in a predictive manner. In this way, sequence similarities can be used to suggest possible activities for unstudied enzymes that have been demonstrated previously for other members of the same family. [16] [35]

Neuraminidases are mainly classified in CAZy families GH33 (bacterial and eukaryotic neuraminidases), GH34 (influenza virus neuraminidases), GH83 (other viral neuraminidases) and GH58 (bacteriophage endoneuraminidases). [11] [19] [35] [29]

8.2. ENZYME COMMISSION (EC) CLASSIFICATION

The Enzyme Commission number (EC number) is a numerical classification for enzymes. The classification is based on the chemical reactions that enzymes catalyse. The EC number is associated with the catalysed reaction, and does not specify types of enzymes. If two enzymes (for example from different organisms and having a different sequence) catalyse the same reaction, then they are given the same EC number. The first digit of the EC number refers to the general reaction. Since neuraminidases are normally involved in hydrolysis, they are mainly classified with an EC number starting at 3-.[43]

Types of neuraminidases according to their EC classification:

Based on their substrate specificity and catalytic mechanism, exo-neuraminidases can be separated into three classes: hydrolytic, trans-neuraminidases and IT-sialidase.

- **Hydrolytic-neuraminidases:** Most of the neuraminidases are in this group, they cleave the glycosidic bond of terminal neuraminic acids and release free neuraminic acid. Hydrolytic-neuraminidases usually have wide substrate specificity and cleave α2-3-, α2-6- and α2-8-linked terminal neuraminic acids.

- **Trans-neuraminidases:** Retain the hydrolytic activity of usual neuraminidases, but with much less efficiency. This second group is mostly found in parasite enzymes. Unlike glycoconjugate sialyltransferases, they catalyse a reversible reaction cleaving the glycosidic bond of terminal neuraminic acids and transferring the cleaved neuraminic acid to other glycoconjugates. Trans-neuraminidase activity with specificity for α2-3-linked substrates was first discovered for the Trypanosoma cruzi neuraminidase. Trans-neuraminidases with other activities like α2-6- and α2-8-linked neuraminic acid substrates were discovered subsequently.

According to the Enzyme Commission, both classes (hydrolytic and trans-neuraminidases) are exo-α-neuraminidases (EC 3.2.1.18).

- IT-neuraminidase (intramolecular neuraminidases): Is the third class of exo-neuraminidases and the less common, they are also named as anhydro-neuraminidases. Currently, the discovered IT-neuraminidases are strictly α2-3-linkage specific and produce 2,7-anhydro-Neu5Ac instead of neuraminic acid (Neu5Ac).The 2,7-anhydro- Neu5Ac seems to be a transient compound as apparently it is slowly converted into free neuraminic acid in aqueous solution. However, the linkage and substrate specificity of neuraminidases is often unknown due to reliance on artificial substrates to their characterization. According to the Enzyme Commission (EC 4.2.2.15). [43] [46]

Also endo-neuraminidases constitute one additional group:

-**Endo-α-neuraminidases:** Hydrolyse viathe endo-cleavage of α2-8-sialosyl linkages of oligo- or polyneuraminic acids.

According to the Enzyme Comission (EC 3.2.1.129). [43]

8.3. PROPOSED CLASSIFICATION AND METHODOLOGY

Neuraminidases are widely distributed in nature. They are mostly found in virus but they are also common in other organisms such as bacteria, some invertebrates, mammals, protozoa or

fungi. Among organisms and in the same branches, neureminidases differ their biochemical properties, for instance in their binding affinity, substrate preference or kinetics. Nevertheless, they conserve structural similarities, domains and some key residues. In terms of biological activity, neuraminidases intervene in a large number of crucial processes throughout the different organisms. In mammals, neuraminidases are involved in pathomechanisms, nutrition and communication [14] [31]. Cytosolic human neuraminidase NEU2 in particular is related with human cancer. [13] Neuraminidases are also strongly related with other diseases such as influenza virus infection [10], mumps virus infection [57] or Chagas disease [33]. Some of these processes will be discussed in more detail in section 10 (biomedical interest of neuraminidases) in this project.

Given their crucial roles, neuraminidases are into the spotlight of current research and a large number of studies have been carried out to elucidate their functioning. The aim of this part of the project is to make a compendium of all structure-known neuraminidases and organize a personal classification of them starting from the biological organism.

The main source of neuraminidases that have been biochemically characterized is the CAZy database. In their website you can find an updated list of all discovered glycosyl hydrolases (GHs), even some recent candidates. There are different tabs where users can consult characterized or structured neuraminidases. Since there is still much to learn about these enzymes, as said before the classification is based on structured neuraminidases, characterized enzymes are not covered. CAZy database purpose for each neuraminidase a bunch of PDB codes of different investigations associated with that enzyme. During the project, the structure with better resolution of the set of the proposed codes by the database has been selected, ruling out inhibitor complexes, enzyme variants, different crystals used for the same structure, structures with low resolution, neuraminidase candidates (without sufficient experimental evidence) and obsolete or repeated codes. As mentioned before, neuraminidases are mainly classified in CAZy families GH33, GH34, GH58 and GH83. Therefore, the PDB codes belonging to these families were the analysed ones. [11] [29] [35] [39]

At the beginning of the classification, a total number of 501 PDB codes were found throughout the 4 families. Most of them are neuraminidases from virus distributed in families GH34 and GH83, 467 PDB codes had been excluded being complexation with inhibitors, drugs or other compounds the main reason for exclusion (see Table 2 and Figure 6). The remaining 34 PDB codes were selected for analysis.

These remaining 34 neuraminidases belong to the following types (see Figure 7 and Table 3) that will be explained in more detail in the following sections. As can be seen in table 3, neuraminidases and their sialyl substrates appear to be absent from plants and most other metazoans, but are present in a wide range of organisms. [19]

Excluded PDB codes	Number of excluded PDB codes
Other complexes (with receptors, other proteins, solvents or neuraminic acid)	133
Complexes with drugs	105
Complexes with inhibitors	71
Mutants	38
Same structures but less resolution	37
Repeated PDB codes	37
Other reasons (derivates, intermediates, Michaelis complex, different type of crystals…)	40
Neuraminidase candidates	6
Total number of excluded PDB codes	467

Table 2: Data analysis of excluded PDB codes

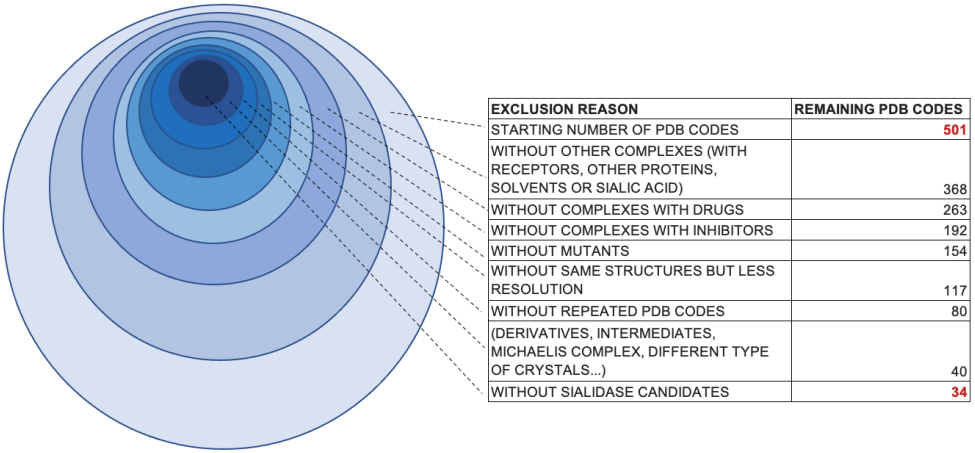


Figure 6: Representation of the remaining PDB codes before each exclusion stage

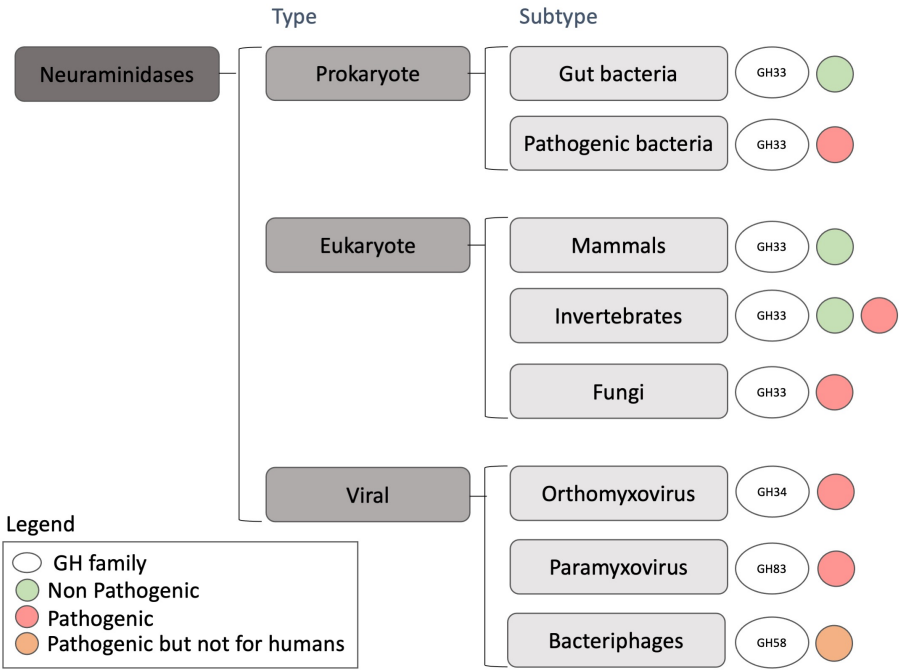


Figure 7: Proposed neuraminidase classification scheme. The complete list of enzymes can be found in Appendix 2.

Type	Subtype	Number of neuraminidases
Virus	Orthomyxovirus	12
	Paramyxovirus	5
	Bacteriophages	2
Prokaryote	Bacteria	10
Eukaryote	Mammals	1
	Invertebrates	3
	Fungi	1
TOTAL NUMBER OF NEURAMINIDASES		34

Table 3: Types and subtypes of organisms that own structure known neuraminidases

8.3.1. Viral neuraminidases

Viral neuraminidases represent 56% (19/34) of all the neuraminidases in Table 3. They are distributed in three families: GH34, which includes exclusively neuraminidases from Influenza A and B viruses (orthomyxovirus subtype), GH83 which include viruses capable of infecting humans (paramyxovirus subtype), and also a third family GH58 which is formed by bacteriophages, these viral neuraminidases infect non-human hosts, they infect bacteria. GH58 neuraminidases are endo-acting and catalyse via an inverting mechanism and due to its characteristic features are the only members of GH58.

8.3.1.1. Neuraminidases in Orthomyxoviridae

In orthomyxoviruses, the neuraminidase activity is only encountered in influenza viruses. There are four types of influenza viruses: A, B, C and D. Influenza C viral infections are not the most common type and generally cause mild illnesses in humans and are not thought to cause epidemics. Influenza D virus infections primarily affect cattle and are not known to infect or cause illness in humans. Influenza viruses A and B are the responsible for the most common infections causing every year winter epidemics (flu season). [22]

Influenza A viruses are divided into subtypes based on two proteins on the surface of the virus: hemagglutinin (HA) and neuraminidase (NA). There are 18 different hemagglutinin subtypes and 11 different neuraminidase subtypes (H1 through H18 and N1 through N11, respectively). [23]

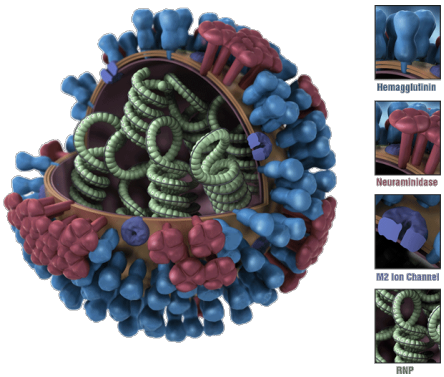


Figure 8: Influenza A virus neuraminidase, originally published by CDC [23]

These HA and NA proteins act as antigens. Antigens are molecular structures on pathogen surfaces that are recognized by the immune system and can trigger an immune response (for example antibody production). Both HA and NA play key roles in viral infection and spread. HA mission is to attach the virus to the host cell interacting with the cell surface neuraminic acids thus

initiating the viral entry. Once the virus has replicated, NA cleaves neuraminic acids from the cell and viral surfaces allowing the virus to leave the now infected cell and continue to spread through the organism. Noteworthy, neuraminidase inhibitors are one of the two major classes of antivirals available for the treatment and prevention of influenza [15] (see figure 8).

Since this project is focused on neuraminidases, the 11 types of neuraminidases have been considered in the table.

Influenza B viruses are not classified into subtypes, but instead lineages are used to divide them. There are 2 lineages: B/Yamagata and B/Victoria, being B the influenza type and the name following the bar the lineage. In this project we considered an example of each lineage discriminating PDB codes with worse resolution and regardless of the origin of the guest. [22]

8.3.1.2. Neuraminidases in Paramyxoviridae [20] [26]

In Paramyxoviridae, neuraminidases were identified in some genus such as rubulavirus, avulavirus or henipavirus. One exemple is the mumps virus, causing parotitis. There are other types of paramyxoviruses as morbilliviruses or pneumoviruses, for example, in which neuraminidase (NA) activity could not be detected, or at least there is no evidence on the CAZy database. Therefore, there is heterogeneity within the family. For most members of the paramyxoviridae family, the virus entry in the host cell, the first step of the viral infection cycle, is mediated by two viral glycoproteins:

-The attachment protein HN (hemagglutinin-neuraminidase), H (hemagglutinin) or G (glycoprotein) depending on the virus

-The fusion protein (F)

The glycoproteins that are involved in this studio are the ones with the attachment protein hemagglutinin-neuraminidase (HN) since they have neuraminidase activity.

In paramyxovirus, the HN protein plays different roles in viral entry and egress. These roles include binding to neuraminic acid receptors, cleaving neuraminic acid from carbohydrate chains and activating the F protein to activate membrane fusion and consequently the viral entry.

8.3.1.3. Neuraminidases in bacteriophages

Bacteriophages, also known as phages, are viruses that exclusively infect bacteria. The neuraminidases of these viruses differ from other NA on their mechanism and action. All the other neuraminidases have exo- or trans- activity and have a retaining mechanism, while this viral

neuraminidases does not follow the same pattern. These special neuraminidases make up the GH58 family. This family is a noteworthy exception to the other known neuraminidases since they have an active site that lacks some of the well-known conserved residues in other neuraminidases such as the nucleophilic tyrosine. They also follow a specific catalytic mechanism, an inverting mechanism. Also, endo neuraminidases are found to hydrolyse a neuraminic acid polymer since they are endo-acting and cleave non-terminal neuraminic acids. [28] There are actually two determined structures of endo neuraminidases forming the GH58 family:

One of them is endo neuraminidase endoNF derived from bacteriophage K1F and specific for Escherichia coli K1. The bacteriophage expresses a tailspike protein that degrades the polyneuraminic acid coat of the bacteria promoting the bacteriophage infection.

The other endo neuraminidase in this family is the endoN92 derived from the bacteriophage phi92 specific for Escherichia Coli K92. [47] These phages have the ability to digest the polyneuraminic acid coats specifically designed to protect bacteria against bacteriophage infection. Endo neuraminidases help hydrolysing the polyneuraminic coat. Moreover, these coats of some pathogenic bacteria are also designed to protect some pathogenic bacteria from the host immune system. The polyneuraminic coats made the bacteria surface practically identical to sugars produced by the host, mimicking this way the host's innate system and protecting the bacteria from the mammal immune system. [47]

8.3.2. Prokaryote neuraminidases (bacteria)

8.3.2.1. Neuraminidases in gut bacteria (non-pathogenic) [14] [31]

In the gastrointestinal tract, which is heavily colonized with bacteria, neuraminic acid is found in terminal locations of a type of glycoproteins named mucins. The proportions of the monosaccharide vary along the gastrointestinal tract with an increasing gradient from the ileum through the rectum in humans. There are two broad groups of mucins:

-The membrane-bound mucins: which are involved in cell interactions and signalling

-The secreted mucins: that are the main structural component of the mucus gel that covers the epithelium and plays an essential role in the maintenance of the equilibrium in our gut microbiota.

As happens in red blood cells, discussed before, the release of mucin terminal neuraminic acids (which normally protect mucins from the action of glycosidases) is the first step in the sequential degradation of mucins. The enzymes responsible of the neuraminic acid hydrolysis in

bacteria are near one another, forming a cluster known as the Nan cluster. The released free monosaccharide is used by the bacteria as a source of energy, nitrogen, and carbon. It seems like not all bacteria that have the ability to cleave host neuraminic acids also can use them as a nutrition source. The ones that have both capacities and are equipped with a Nan cluster are normally able to colonize mucus regions of the human body such as the gastrointestinal tract.

8.3.2.2. Neuraminidases in pathogenic bacteria

In contrast with gut commensals which seem to use neuraminidases primarily as a nutrition tool, some pathogenic bacteria use them to decrypt adhesion or toxin-binding sites. The same way as gut bacteria, pathogenic bacteria neuraminidases cleave the terminal mucin neuraminic acid causing a massive destruction of the host tissue and discovering possible binding sites in which the pathogenic bacteria will attach to continue spreading throughout the host organism. In general, in vitro studies showed that an increased neuraminidase expression improves bacterial adhesion to the epithelium. An example of this is what happens with Vibrio cholerae neuraminidase [31] which causes a life-threatening disease known as cholera or the neuraminidases found on Clostridium perfringens which also causes potentially fatal diseases such as gas gangrene or enterotoxemia in humans. NanI, NanH and NanJ are the neuraminidases that form the Nan cluster of Clostridium perfringens. NanI and NanH are listed in the proposed table, NanJ has been characterized but has no crystal structure. This last neuraminidase would be in the same group as the other two, and plays an important role in pathogenesis development. [25]

Another important group of pathogenic bacterial neuraminidases are the ones found in the bacteria Streptococcus pneumoniae (the pneumococcus), one of the most common human pathogens which colonize nasopharynx having the possibility to cause sinusitis, otitis media, pneumonia if it spreads to the lungs, and even bacteraemia or meningitis if it penetrates into the bloodstream and into the blood-brain barrier respectively. These bacteria express up to three neuraminidases, NanA, NanB and NanC that play different roles during the pathogenesis. [24] [32]

8.3.3. Eukaryote neuraminidases

Eukaryote neuraminidases are all in the family GH33, representing 15% (5/34) of all neuraminidases listed in the proposed classification. They are present in mammals, invertebrates and fungi.

8.3.3.1. Neuraminidases in Mammalia [41]

Four types of mammalian neuraminidases have been identified and characterized to date. They are also present in humans and are labelled as NEU1, NEU2, NEU3 and NEU4. They are encoded by different genes and differ in their subcellular localization and enzymatic properties including substrate specificity. Apart from these, each neuraminidase has been found to play a unique role and having individual properties. Their subcellular locations are:

NEU1 localized predominantly in lysosomes and expressed mainly in pancreas, followed by skeletal muscles. A low level of expression was also observed in the brain tissues.

NEU2 is located in cytosol and expressed mostly in skeletal muscle and faetal liver.

NEU3 is in plasma membranes, found on cell surfaces and expressed also in skeletal muscles, adrenal gland, pancreas, heart, liver. A lower expression of this gene is also found in placenta, lungs, kidney, and brain.

NEU4 lysosomes or mitochondria and intramembranous components. NEU4 occurs in two different forms that differ in the first 12 amino acids resulting in a short form (related with the endoplasmic reticulum membrane) and a longer form associated with mitochondria and lysosomes. Both longer and shorter forms are predominantly expressed in muscle, kidney and brain, and the short form is also found in colon and liver.

Even though 4 mammal neuraminidases have been characterized, only for one of them has been possible to obtain the structure, the cytosolic one (NEU2), therefore it's the only one that is listed in the proposed table. The reason why characterization of the other human neuraminidases is difficult could be due to the fact that they are membrane-bound proteins. [48]

8.3.3.2. Neuraminidases in invertebrates

In this group there are three listed neuraminidases. Two of them are from parasites, *Trypanosoma rangeli* and *Trypanosoma cruzi*. The first one is not related to any disease in humans, the second neuraminidase in this parasite is related to chagas, an infectious and inflammatory disease caused by the parasite Trypanosoma cruzi which comes from the faeces of different insects mostly located in South America. The neuraminidase present in the parasite has been strongly related with the process of invasion of the pathogen. The trans-sialidase present in the parasite catalyses the transference of neuraminic acid molecules from the host glycoconjugates to acceptor molecules placed on the parasite surface. Also, this trans-sialidase activity mediates pathogenic crucial activities such as the subversion of the host immune system,

helping the parasite survival and resulting in chronic infection.[34] [35] The other neuraminidase in this group is the anhydro-sialidase from the leech (Macrobdella decora), which was one of the first discovered intramolecular sialidases. [47]

8.3.3.3. Neuraminidases in fungi

The only fungal neuraminidase included in Table 3 comes from a filamentous fungi known as *Aspergillus fumigatus*, which is recognised to be the cause of severe diseases in patients with previous pathologies such as cystic fibrosis or immunosuppression.[43] This neuraminidase, in contrast with all the other listed ones, has a low activity hydrolysing Neu5Ac from glycan substrates. Kinetic studies of *Aspergillus fumigatus* neuraminidase revealed that it has greater hydrolytic activity with 2-keto-3-deoxy-D-glycero-D'Galacto-nonanoic acid (KDN) than with Neu5Ac. KDN is a naturally occurring neuraminic acid derivative that has the OH group in place of the N-acetyl group at C-5 (see Figure 9). KDN is found in almost all types of glycoconjugates often in place of Neu5Ac and predominantly found in bacteria and lower vertebrates. The catabolism of the KDN remains an unknown process and the catalytic mechanism of this neuraminidases has not been fully investigated. However, some studies show that some bacteria (Sphingobacterium multivorum) that has this neuraminidase (KDNase) use KDN acids as a carbon source just like some gut bacteria do with Neu5Ac. This KDNase has also been identified in several species of molluscs and rainbow trout.

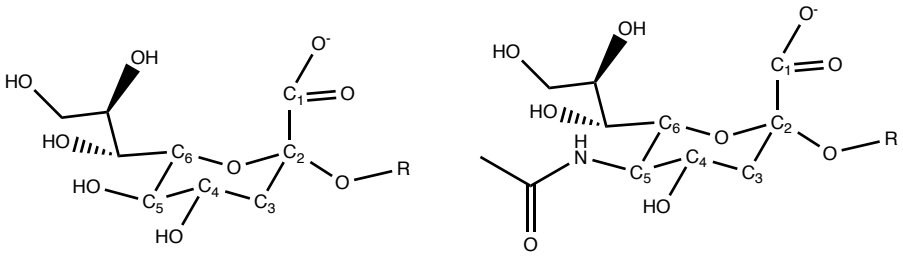


Figure 9: Comparison between KDN Acid structure and Neuraminic acid

9. COMPARISON OF NEURAMINIDASE STRUCTURES

All neuraminidases have a canonical six-sheet beta-propeller structure with a conserved catalytic site. The active site of neuraminidases contains residues forming hydrogen bonds and salt bridges among their side chains. The essential amino acid residues involved in binding and catalysis are highly conserved. [9] In general:

-An arginine triad, very preserved across all known neuraminidases which is a key feature of the active pocket. These arginines are involved in the coordination of the substrate. The three residues form a net positively charged environment that surround the negatively charged carboxylate group from the Neu5Ac and its derivatives thus binds the carboxylate group of the substrate via electrostatic interactions. [36] [37]

-The nucleophile pair Tyr/Glu, also highly conserved across all the neuraminidase family.

-An Aspartic acid that acts as a general acid/base catalyst, not present in all neuraminidase families.

In this part of the project, the VMD software (11 version) has been used to analyze the structures of one representative neuraminidase from each group of the proposed classification. The selected neuraminidases are listed in Table 4, together with the position of their most representative residues (residues numbers as in the sequence and PDB structure).

Type	Subtype	VMD representation	Tyr	Glu	Asp	Arg triad
Prokaryotes	Bacteria	NedA	370	92	291	68,276,342
Eukaryotes	Mammalia	NEU2	334	218	46	21,237,304
	Invertebrates	Trypanozoma cruzi	342	230	59	35,245,314
	Fungi	KDNASE	358	249	84	59,265,322
Virus	Orthomyxoviruses	Influenza N2	406	276	n.d.	118,292,371
	Paramyxoviruses	Mumps virus neuraminidase	540	407	n.d.	180,422,512
	Bacteriophages	ENDONF	**Thr**	**Ser**	**Gln**	**Arg**
			846	848	853	837

n.d. : non defined

Table 4: Neuraminidase representative for each family and important catalytic residues

9.1. PROKARYOTE NEURAMINIDASE STRUCTURES

9.1.1. Micromonospora viridifaciens neuraminidase (NedA) long form (68Kda)

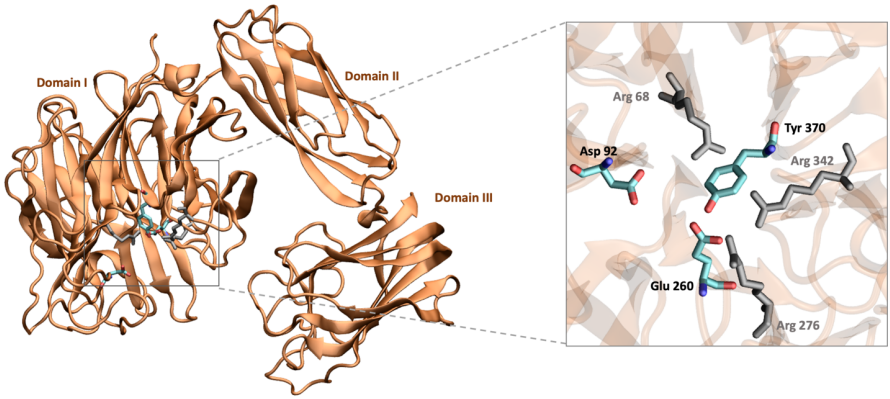


Figure 10: VMD Structure representation of multi-domain the long form of the *Micromonospora Viridifaciens* neuraminidase (NedA) from the PDB code 1EUU, with a 2.5 Å resolution (apo structure).

Bacterial neuraminidases are in general multi-domain enzymes. They have the beta-propeller folding domain with the active site and some additional domains that help with the carbohydrate recognition and give the enzyme function for specific substrates and environments. Normally the protein size differs between spices. In this case, the Micromonospora viridifaciens is secreted in two different forms, the short form of 41 Kda and the long form of 68 Kda. Bacterial neuraminidase's normally have an optimum pH range between 5 and 7.[58] Figure 11 shows the multi-domain protein with the conserved catalytic residues.

9.2. EUKARYOTA NEURAMINIDASE STRUCTURES

9.2.1. Cytosolic human neuraminidase (NEU2) – Mammals

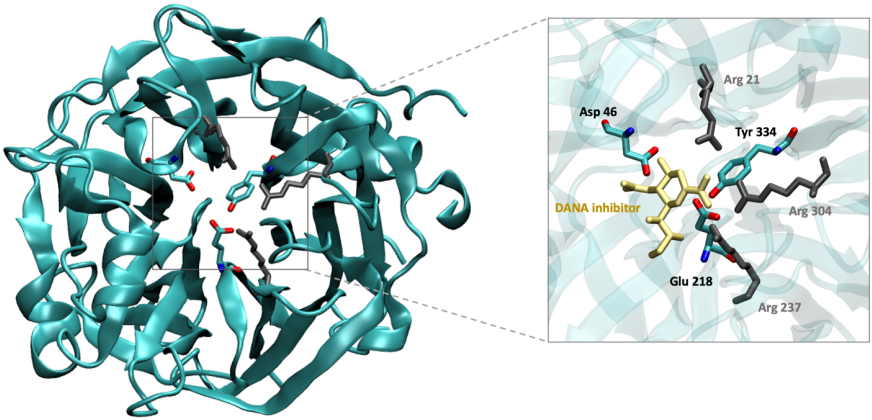


Figure 11: VMD structure representation of the cytosolic human neuraminidase NEU2 from PDB codes 1SNT, with 1.75 Å resolution (apo structure, left) and 1VCU, with 2.85 Å resolution (complex with the DANA inhibitor, right).

To date, among human neuraminidases only NEU2 could be crystallized to be structurally investigated. Based on the crystal form of NEU2, some structure models had been established for NEU1, NEU3 and NEU4. As shown in Figure 11, the six-sheet beta-propeller is conserved in the cathalytic centre which includes the also preserved catalytic residues. Moreover NEU1, NEU2, NEU3 and NEU4 have in common several motifs such as the Asp-boxes [61] [62], which are also shared by other neuraminidases. [48] The homology in the amino acid sequence between NEU1 and the other human neuraminidases is only 19-24%. The accordance between NEU2, NEU3 and NEU4 is 43-40% with each other. This differences between NEU1 and the other human neuraminidases might be associated with the subcellular location in the lysosomes. [48]

9.2.2. Trypanosoma Cruzi neuraminidase – Invertebrates

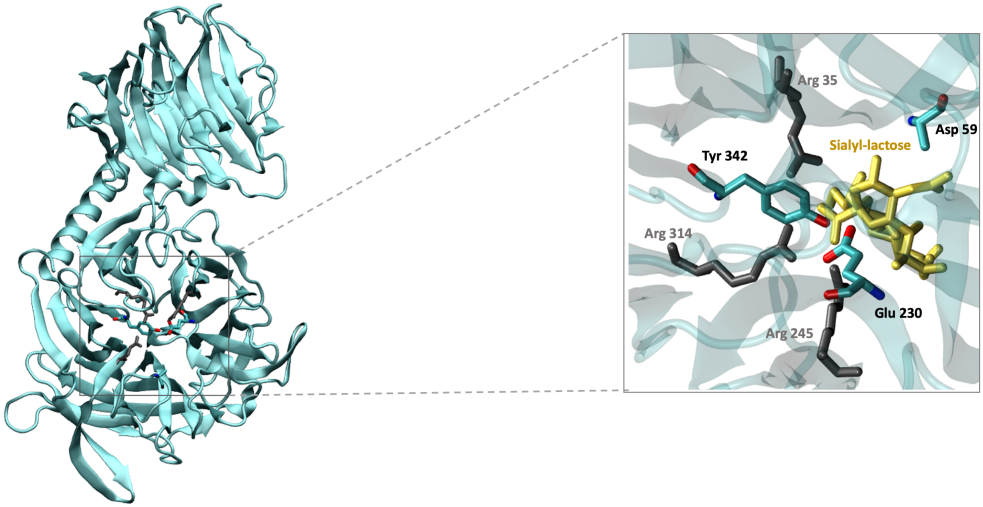


Figure 12: VMD Structure representation of the Trypanosoma Cruzi trans-neuraminidase from the PDB codes: [1MS3] with a 1.65 Å resolution of the apo structure and [1S0I] with a 1.6 Å resolution of the complex with the substrate sialyl-lactose.

These parasite proteins, contain the enzymatic domain on the N terminal extreme, which is the only one required for the trans-sialidase activity. They also have an antigenic domain on the C-terminal end.[34] Figure 12 shows the conserved structure of the neuraminidase active site.

9.2.3. Aspergillus fumigatus neuraminidase – Fungi

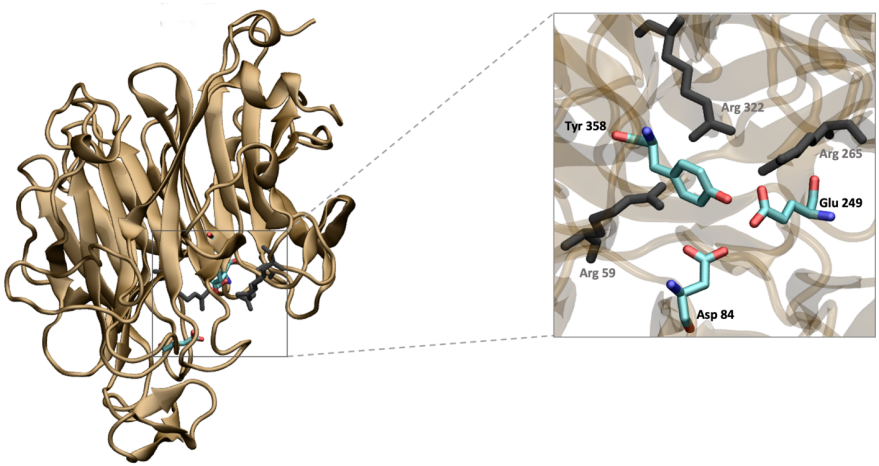


Figure 13: VMD Structure representation of the Aspergillus Fumigatus KDNase from the PDB code: [2XZI] with a 1.45 Å resolution of the apo structure.

The amino acid residues Tyr358 and Asp84 have been proposed to be the catalytic nucleophile and the general acid/base catalyst involved in the hydrolysis of the KDN acid by *Aspergillus fumigatus* KDNase, whereas Glu247 is the assisting residue that deprotonates the tyrosine. [56] Also, Arg171 (which is not highlighted in Figure 13) is believed to be the key amino acid reducing the hydrolysis effectivity of Neu5Ac by the KDNase. Mutation studies show that the substitution of the Arg171 with a leucine improved the catalysis toward Neu5Ac substrates. The wild neuraminidase with the arginine in the 171 position is primarily a KDNase and this might be because of steric clash with the N-acetyl group. [42]

9.3. VIRAL NEURAMINIDASE STRUCTURES

9.3.1. Influenza A Virus neuraminidase (N2) – Orthomyxovirus

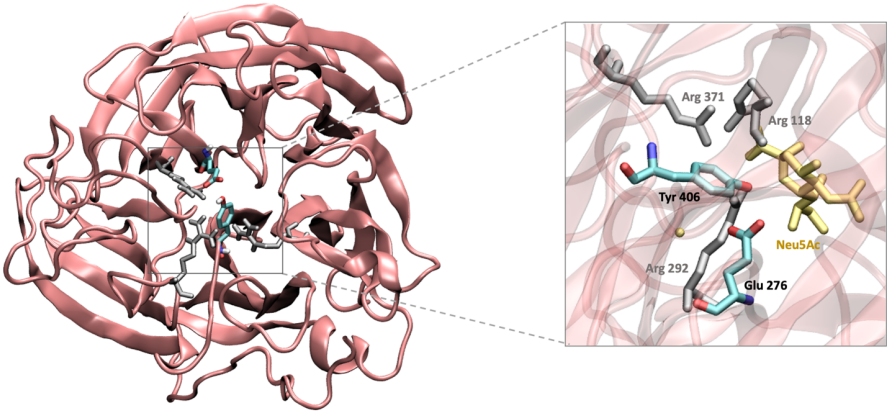


Figure 14: VMD Structure representation of the Influenza A/Tokyo/3/1967(H2N2) virus neuraminidase N2 from the the PDB codes: [1NN2] with a 2.20 Å resolution of the apo structure and [2BAT] with a 2.00 Å resolution of the complex with the substrate Neu5Ac.

The neuraminidases in influenza viruses are built by four identical monomers, which are associated in a mushroom-like tetramer. Figure 14 shows one of these monomers. Each monomer consists of six quadruple chain-, antiparallel beta sheets, whose arrangement leads to into the canonical beta propeller structure. As opposed to other structures, in this viral neuraminidase the catalytic Asp residue is missing. Some studies show that the replacement of the Asp151, the one that should be the analogue acid/base residue in the other active centres, with a glycine, does not affect significantly the activity of the enzyme or its ability to remove receptors for hemagglutinin. This residue, however, seems to have a restrictive role in the specificity of the neuraminidase interaction and its hydrolysing activity that complement the hemagglutinin binding function. [21] [60] Also, viral neuraminidases do not have Asp-boxes, a highly conserved motif in other neuraminidases. [46]

9.3.2. Mumps virus neuraminidase – Paramyxovirus

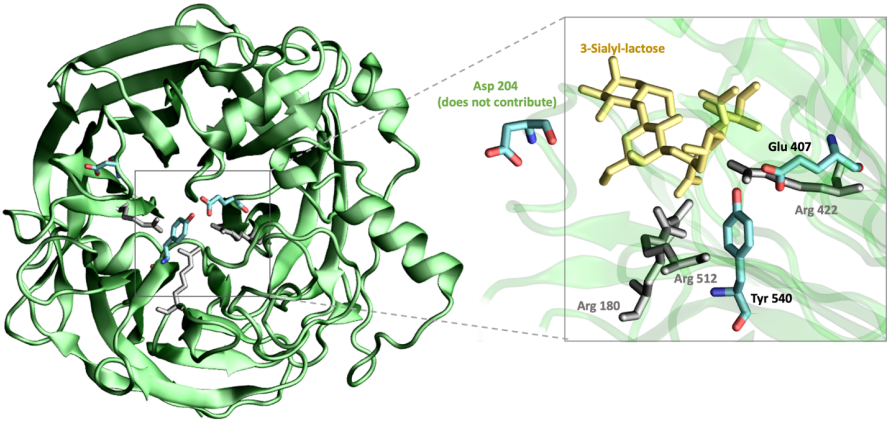


Figure 15: VMD Structure representation of the Mumps virus hemagglutinin-neuraminidase (MvN) from the the PDB codes: [5B2C] with a 2.24 Å resolution of the apo structure and [5B2D] with a 2.18 Å resolution of the complex with the substrate 3-Sialyl-lactose.

The hemagglutinin-neuraminidase (HN) is an oligomeric integral membrane protein consisting of: N-terminally a small cytoplasmatic region, followed by a short transmembrane domain, a stalk region, and a globular head which contains the active site domain. The active site is responsible for both the binding to the receptor (neuraminic acids) as well as the neuraminidase enzymatic activity (NA). The structure of HN and NA domains have been solved, there are different structure examples in the proposed table. However, the structure of the stalk region has remained unclear. The stalk region is thought to carry specificity determinants for F protein interaction and activation, which affects neuraminidase activity and contributes to the oligomerization of the protein. Mutational studies of the New Castle disease HN virus stalk revealed how they affect membrane fusion and consequently viral entry. [20] [26] Figure 15 shows that, as in influenza neuraminidases, the Asp residue that normally participates in the catalysis is pointing away from the substrate and the the active site. This could mean that, despite the function of the Asp residue is conserved in neuraminidases, in this case, as well that with influenza virus, the Asp residue does not contribute to catalysis. The other catalytic residues, Tyr540 and Glu407, are located as in other neuraminidases.

9.3.3. Escherichia phage K1F neuraminidase (EndoNF) – Bacteriphages

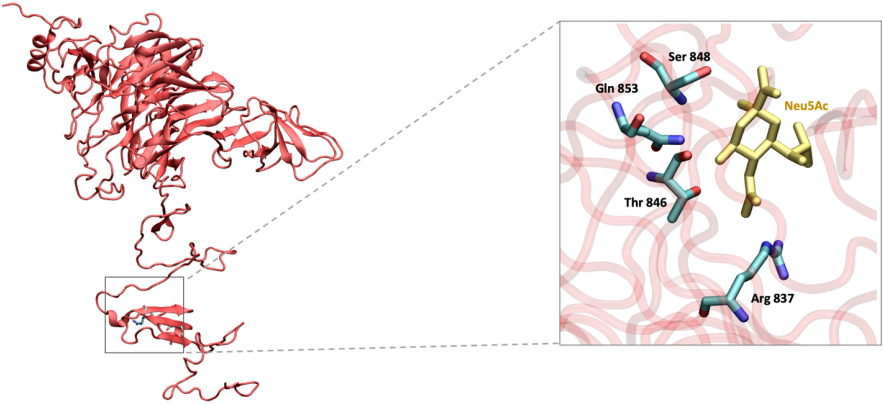


Figure 16: VMD Structure representation of the Esterichia phage K1F endo-neuraminidase (EndoNF) from the the PDB codes: [3JU4] with a 0.98 Å resolution of the apo structure and [1V0F] with a 2.55 Å resolution of the stalk domain with the substrate Neu5Ac.

The neuraminidase of the Escherichia coli K1 bacteriophage reveal a putative active site that is similar in geometry to those of exo-neuraminidases, but lacks three of the most characteristic neuraminidase residues: the tyrosine (nucleophilic catalyst), one of the two carboxylate catalysts, and one of the three arginines. This differing active site suggests a new endo-acting mode that results in to an inverting catalysis, which is probably the most differing of all the above structures. [47] The EndoNF is formed by a beta-barrel domain, a beta-propeller domain, and a stalk domain. The active site is located in the centre of the beta-propeller domain. Figure 16 shows the presence of a neuraminic acid molecule near the stalk domain, and the highlighted residues are the ones that interact with the substrate via hydrogen bonds. [59]

10. BIOMEDICAL INTEREST OF NEURAMINIDASES

Neuraminidases have been implicated in the pathogenesis of several diseases [2-4][9,10,14][21-27][31-33][41][50-53] and this has significant impact on human health. The inhibition of specific neuraminidases opens the possibility to fight bacterial or viral infections. [4][21-27] However, there's a problem with this therapeutic target: the similarity of the catalytic domain mostly conserved and shared by all types of neuraminidases [54][55], as well as the presence of these enzymes in a wide range of organisms, including humans. [41] [48] Also, the development of several pathological processes in the human body can be associated with either over-expression of neuraminidases or their lacking production. Therefore, we thought interesting to describe the biomedical interest of neuraminidases that are directly produced by the human body or related with pathogens that affect humans. [41]

10.1. NEURAMINIDASES IN BACTERIAL INFECTION

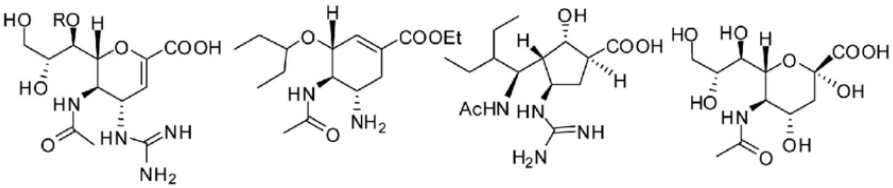
Neuraminidases are crucial for the survival of bacterial commensals, protozoa, and pathogens in the human body. [27] Neuraminidases play a key role in biofilm formation for example in *Streptococcus pneumoniae* bacteria. [24] [32] The released neuraminic acids can be used to stimulate as a source of nutrients and then stimulate the growth of some bacterial strains. [14] The neuraminidase production could be associated with the pathophysiology of bacterial infection and seems to promote the growth of pathogens like *Clostridium perfringens*, which causes intestinal infection and *Clostridial myonecrosis* [25] or *Escherichia coli* [14] in the gastrointestinal tract. Also, the neuraminidase activity of Vibrio cholerae is strongly related with the cholera disease, dues to the removal of the neuraminic acid from the gangliosides, the neuraminidase unmask the receptors of cholera toxin facilitating the attachment to the cells. [31] Neuraminidases are also related with dental health, for example in Tannerella forsythia. The neuraminidases produced by this bacteria help stimulate the growth of the bacteria in the oral cavity, which can lead to periodontitis. [41]

10.2. NEURAMINIDASE INHIBITORS AS ANTIVIRALS AND OTHER THERAPIES

It is worldwide approved that neuraminidase inhibitors such as oseltamivir (Tamiflu) and zanamivir (Relenza), or peramivir (Rapivab) can be used as effective drugs to suppress influenza A virus activity [4][6] (see Figure 17).

Moreover, inhibitors such as oseltamivir or the candidate siastatin B can be also used to put-down the growth of Tannerella forsythia bacteria, resulting in the reduction of periodontal plaque

biofilm. [41] There are some evidence showing that the influenza A virus inhibitor oseltamivir also inhibits the human neuraminidase NEU3, resulting in side effects. [54] Some investigations revealed that the candidate neuraminidase inhibitor siastin B also inhibits the cytosolic NEU2. [55] Influenza A inhibitors inhibit the bacterial neuraminidase NanA. [30] The exposed facts show the difficulties of the use of neuraminidase inhibitors as drugs. The choice of a highly selective and suitable inhibitor can be a crucial factor for the development of new effective treatments. The inhibitor must selectively reduce the activity of the desired neuraminidase related to the pathogen without affecting other human neuraminidase activity. [49]



R=H Zanamivir Oseltamivir Paramivir alfa-Neu5Ac
R=Me Laninamivir

Figure 17: Neuraminidase inhibitors for influenza A virus, *originally published by B. Richichi [49]*

10.3. NEURAMINIDASES IN CANCER

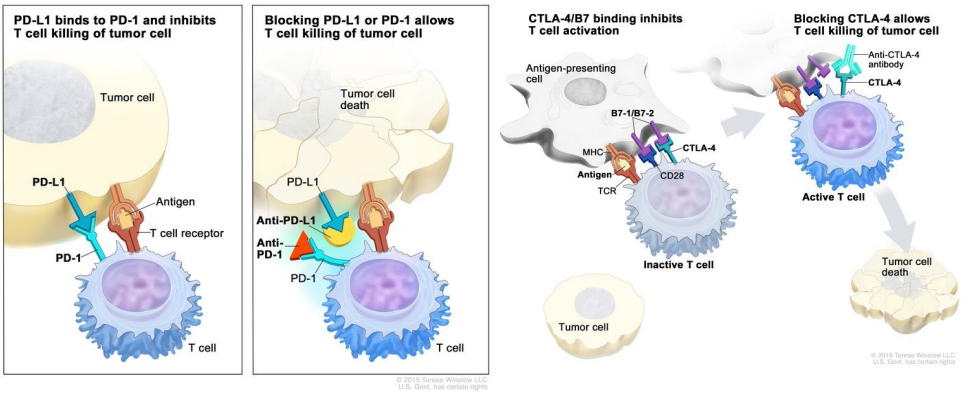
It is well known that conventional cancer treatments such as radiotherapy and chemotherapy, despite their effectiveness, can cause many side effects. These side effects occur when the treatment damages the healthy cells. These side effects differ depending on each person but the most common ones are: hair loss, neutropenia, lymphedema, nauseas, cancer pain, ... From a few years ago, extensive work has been carried out for the development of alternative immunotherapy treatments. The researchers James Allison and Tasuku Honjo won the 2018 Nobel Prize in Medicine for their pioneering work on cancer immunotherapy. [6] Immunotherapy treatments are based on avoiding the action of immune checkpoints. These are part of our system and prevent the immune system from acting too strongly and destroying cells that are healthy. The inhibition of these checkpoints allows a greater aversion against cancer cells by the immune system. There are two different ways to avoid the action of immune checkpoints:

-Interact with the immune checkpoint itself inhibiting its function

-Interact with the receptor that will attach to the immune checkpoint inhibiting the action of T-cells.

T cells, also called thymocytes, are a type of white blood cell and are part of the immune system. These cells are responsible for fighting tumour cells and protecting the body against infections. [7] [8] Sometimes, the action of the immune checkpoints can interfere in the action of the T-cells, resulting in the protection of the tumour cells.

Some of the control points that can be inhibited with these medicines are: the immune checkpoint PD-L1 and the CTLA-4, also the PD-1 ligand can also be inhibited (see figures 18 19).[5]



Figures 18 and 19: Immune Checkpoint inhibitors, originally published by Nacional Cancer institute [5]

Immunotherapies are only effective in less than 13% of the patients so that most of them suffering from this pathology cannot benefit from immunotherapy as a treatment. The key that could increase this effectiveness could be found in glycobiology and the action of neuraminidases.[3] Not only do cancer cells hide from the immune system through these immune checkpoints, but they also cause alterations in the glucan layer that lines their membrane in order to confuse the T cells. Figure 20 shows that a healthy cell is coated by a thin layer of ordered

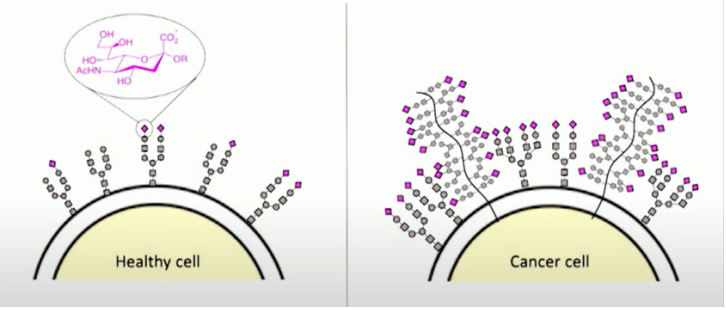


Figure 20: Healthy and cancer cell, YouTube screenshot originally published by Carolyn Bertozzi [2]

glucans while a cancerous cell is coated in a different (much denser) and disordered way. Also, cancer cell presents has a higher number of neuraminic acid molecules. [2]

A low production of neuraminidases (NEU1, NEU2 and NEU4) was found in cancer cells, leading to the accumulation of sialoglycans and tumour cell growth. [41] This higher content is also attributed to sialyltransferase upper-regulation. The increased expression of neuraminic acid on cell surfaces is a specific characteristic for tumour cells. Also, this accumulation is linked to immune evasion, blocking vital signalling pathways, avoiding the natural apoptosis, and reducing the efficacy of chemotherapy and radiotherapy. There is still much to discover about the process and the role of glycosylation in cancer and metastasis but it's clear that the understanding of it will be enlightening for the development of better cancer treatments. [2] [3]

10.4. NEURAMINIDASES IN CENTRAL NERVOUS SYSTEM DISEASES

The tissues of the brain are the tissues of the human body with a higher content of sialoglucoconjugated substrates. Changes in the neuraminidase activity, can modulate the sialylation level of glycans and affect the hippocampal memory and the synaptic plasticity. Some studies on rats observed that dessialylation of glycoconjugates by neuraminidases might be involved in memory processing. [50] The memory processing is related with a change on cell surfaces and can affect sialyl signalling-dependent neural activities. Abnormal sialyl signalling is known to cause Alzheimer's disease, epilepsy, or other diseases. So a decrease in NEU1 neuraminidase activity, promotes an excessive sialylation and might lead to abnormal accumulation of amyloid precursor proteins which are involved in the production of neurotoxic peptides relevant for the Alzheimer's disease. Then, the stimulation of neuraminidase production or the addition of non-human neuraminidases can be seen as a potential tool for neurodegenerative disease treatment. [51]

Another neurodegenerative disease related to neuraminidase activity is Parkinson disease. Neuraminidases can be used to achieve neuroprotective effects on cells. The use of neuraminidases to increase the GM1 ganglioside level in the brain by the degradation of polysialogangliosides could potentially become a new strategical way for the treatment of Parkinson's disease. The progressive reduction of the GM1 biosynthesis along the lifespan, is considered one of the underlying causes for neuronal loss in aged people and neuronal decline in patients with neurodegenerative diseases such as Parkinson. Some Parkinson disease model

studies have demonstrated the neuroprotective efficacy of the neuraminidase from Vibrio cholerae in mice. [52]

10.5. NEURAMINIDASE-RELATED GENETIC DISEASES

As explained before, either over-expression or missproduction of neuraminidases could trigger human pathogenesis. NEU1 dysfunction is the responsible of two rare but severe genetic diseases: sialidosis (OMIM #256550) and galactosialidosis (OMIM #256540). Both of them are genetic diseases that cause the reduction of the lysosomal neuraminidase activity and therefore the accumulation of gangliosides and other sialylated molecules in different tissues and organs, resulting in severe cytotoxicity and cell death. [53]

11. CONCLUSIONS

The bibliographical study carried out here shows that neuraminidases in general share the same active site, with some exceptions collected in viral neuraminidases, and catalyse the chemical reaction via a retaining mechanism. Many mechanistic questions remain unsolved, which were also briefly described. Among these, the details of the less abundant neuraminidases using intramolecular or inverting mechanisms. The bibliographic study also evidenced that there is no compendium of neuraminidases in which the most relevant chemical and biological information on each neuraminidase is collected. Therefore, we tried to do it ourselves. To do so, a total number of 501 neuraminidases were analysed and 34 of them were filtered for a more thoughtful structural, mechanistic and functional study. This resulted in a new classification according to their structure, function, pathogenesis and the organism they belong. We hope that this classification can be useful for future research in neuraminidases.

12. REFERENCES AND NOTES

[1] *Ácido siálico*. (2016, November 16). EcuRed. Retrieved September 2021, from https://www.ecured.cu/%C3%81cido_si%C3%A1lico

[2] Bertozzi, C. [Stanford Alumni]. (2019, November 1). *Sugar Coating of cancer cells* [Video]. Youtube. https://www.youtube.com/watch?v=S_p-ngKgFB4

[3] Landhuis, E. (2019, July 20). Glycoscience fuels a new breed of cancer immunotherapy. *Chemical & Engineering News*. https://cen.acs.org/pharmaceuticals/oncology/Glycoscience-fuels-new-breed-cancer/97/i29?ref=search_results

[4] Thayer, A. M. (2009, September 28). Flu fighters. *Chemical & Engineering News*. https://cen.acs.org/articles/87/i39/Flu-Fighters.html?ref=search_results

[5] National Cancer Institute. (2019, September 24). *Immune checkpoint inhibitors*. Retrieved September 2021, from https://www.cancer.gov/about-cancer/treatment/types/immunotherapy/checkpoint-inhibitors

[6] Allison, J. P., & Honjo, T. (2018, October 1). *The novel price in physiology or medicine 2018*. The Novel Price. Retrieved September 2021, from https://www.nobelprize.org/prizes/medicine/2018/press-release/

[7] National Cancer Institute. (2007). *T cell*. Retrieved September 2021, from https://www.cancer.gov/publications/dictionaries/cancer-terms/def/t-cell

[8] Wikipedia. (2021b, September 19). *Siglec*. Retrieved September 2021, from https://en.wikipedia.org/wiki/Siglec

[9] Chan, J., Watson, J. N., Lu, A., Cerda, V. C., Borgford, T. J., & Bennet, A. J. (2011). Bacterial and Viral Sialidases: Contribution of the Conserved Active Site Glutamate to Catalysis. *Biochemistry*, 51(1), 433–441. https://doi.org/10.1021/bi201019n

[10] Centers for disease control and prevention. (2021c, September 28). *Temporada de influenza*. Retrieved September 2021, from https://espanol.cdc.gov/flu/about/season/flu-season.htm

[11] Giacopuzzi, E., Bresciani, R., Schauer, R., Monti, E., & Borsani, G. (2012). New Insights on the Sialidase Protein Family Revealed by a Phylogenetic Analysis in Metazoa. *PLoS ONE*, 7(8), e44193. https://doi.org/10.1371/journal.pone.0044193

[12] Wikipedia. (2021a, January 19). *Exoglycosidases*. Wikipedia. Retrieved September 2021, from https://en.wikipedia.org/wiki/Exoglycosidase

[13] Dobie, C., & Skropeta, D. (2020). Insights into the role of sialylation in cancer progression and metastasis. *British Journal of Cancer*, 124(1), 76–90. https://doi.org/10.1038/s41416-020-01126-7

[14] Juge, N., Tailford, L., & Owen, C. (2016). Sialidases from gut bacteria: a mini-review. *Biochemical Society Transactions*, 44(1), 166–175. https://doi.org/10.1042/bst20150226

[15] Kim, J. H., Resende, R., Wennekes, T., Chen, H. M., Bance, N., Buchini, S., Watts, A. G., Pilling, P., Streltsov, V. A., Petric, M., Liggins, R., Barrett, S., McKimm-Breschkin, J. L., Niikura, M., & Withers, S. G. (2013). Mechanism-Based Covalent Neuraminidase Inhibitors with Broad-Spectrum Influenza Antiviral Activity. Science, 340(6128), 71–75. https://doi.org/10.1126/science.1232552

[16] *CAZy - Home*. (2021). Carbohydrate-Active Enzymes. Retrieved September 2021, from http://www.cazy.org/

[17] *Glycoside hydrolases*. (2017, April 27). CAZypedia. Retrieved September 2021, from https://www.cazypedia.org/index.php/Glycoside_hydrolases

[18] *Protein data bank*. (2000, January 1). The Protein Data Bank. Retrieved September 2021, from https://www.rcsb.org/

[19] Zechel, D. L., & Withers, S. G. (2010). ChemInform Abstract: Glycosidase Mechanisms: Anatomy of a Finely Tuned Catalyst. ChemInform, 31(15), no. https://doi.org/10.1002/chin.200015237

[20] Yuan, P., Swanson, K. A., Leser, G. P., Paterson, R. G., Lamb, R. A., & Jardetzky, T. S. (2011). Structure of the Newcastle disease virus hemagglutinin-neuraminidase (HN) ectodomain reveals a four-helix bundle stalk. *Proceedings of the National Academy of Sciences*, 108(36), 14920–14925. https://doi.org/10.1073/pnas.1111691108

[21] Shtyrya, Y. A., Mochalova, L. V., & Bovin, N. V. (2009). Influenza Virus Neuraminidase: Structure and Function. *Acta Naturae*, 1(2), 26–32. https://doi.org/10.32607/20758251-2009-1-2-26-32

[22] Centers for disease control and prevention. (2021b, August*). Antiviral drugs for seasonal influenza: Additional links and resources*. Retrieved October 2021, from https://www.cdc.gov/flu/professionals/antivirals/links.htm

[23] Centers for disease control and prevention. (2021a). *Types of influenza viruses*. Retrieved October 2021, from https://www.cdc.gov/flu/about/viruses/types.htm

[24] Janesch, P., Rouha, H., Badarau, A., Stulik, L., Mirkina, I., Caccamo, M., Havlicek, K., Maierhofer, B., Weber, S., Groß, K., Steinhäuser, J., Zerbs, M., Varga, C., Dolezilkova, I., Maier, S., Zauner, G., Nielson, N., Power, C. A., & Nagy, E. (2018). Assessing the function of pneumococcal neuraminidases NanA, NanB and NanC in in vitro and in vivo lung infection models using monoclonal antibodies. *Virulence*, 9(1), 1521–1538. https://doi.org/10.1080/21505594.2018.1520545

[25] Newstead, S. L., Potter, J. A., Wilson, J. C., Xu, G., Chien, C. H., Watts, A. G., Withers, S. G., & Taylor, G. L. (2008). The Structure of Clostridium perfringens NanI Sialidase and Its Catalytic Intermediates. *Journal of Biological Chemistry*, 283(14), 9080–9088. https://doi.org/10.1074/jbc.m710247200

[26] Takimoto, T., Taylor, G. L., Crennell, S. J., Scroggs, R. A., & Portner, A. (2000). Crystallization of Newcastle Disease Virus Hemagglutinin-Neuraminidase Glycoprotein. *Virology*, 270(1), 208–214. https://doi.org/10.1006/viro.2000.0263

[27] Zaramela, L. S., Martino, C., Alisson-Silva, F., Rees, S. D., Diaz, S. L., Chuzel, L., Ganatra, M. B., Taron, C. H., Secrest, P., Zuñiga, C., Huang, J., Siegel, D., Chang, G., Varki, A., & Zengler, K. (2019). Gut bacteria responding to dietary change encode sialidases that exhibit preference for red meat-associated carbohydrates. *Nature Microbiology*, 4(12), 2082–2089. https://doi.org/10.1038/s41564-019-0564-9

[28] Morley, T. J., Willis, L. M., Whitfield, C., Wakarchuk, W. W., & Withers, S. G. (2009). A new sialidase mechanism BACTERIOPHAGE K1F ENDO-SIALIDASE IS AN INVERTING GLYCOSIDASE. JBC *Papers in Press*. https://doi.org/10.1074/jbc.M109.003970

[29] Stick, R., V., & Williams, S. (2008). Carbohydrates: The Essential Molecules of Life (2nd ed.). Elsevier Science.

[30] Gut, H., Xu, G., Taylor, G. L., & Walsh, M. A. (2011). Structural basis for streptococcus pneumoniae NanA inhibition by influenza antivirals zanamivir and oseltamivir carboxylate. Journal of Molecular Biology, 409(4), 496–503. https://doi.org/10.1016/j.jmb.2011.04.016

[31] Crennell, S., Garman, E., Laver, G., Vimr, E., & Taylor, G. (1994). Crystal structure of Vibrio cholerae neuraminidase reveals dual lectin-like domains in addition to the catalytic domain. Structure, 2(6), 535–544. https://doi.org/10.1016/s0969-2126(00)00053-8

[32] Hsiao, Y. S., Parker, D., Ratner, A. J., Prince, A., & Tong, L. (2009). Crystal structures of respiratory pathogen neuraminidases. Biochemical and Biophysical Research Communications, 380(3), 467–471. https://doi.org/10.1016/j.bbrc.2009.01.108

[33] Schenkman, S., & Eichinger, D. (1993). Trypanosoma cruzi trans-sialidase and cell invasion. Parasitology Today, 9(6), 218–222. https://doi.org/10.1016/0169-4758(93)90017-a

[34] Amaya, M. F., Watts, A. G., Damager, I., Wehenkel, A., Nguyen, T., Buschiazzo, A., Paris, G., Frasch, A., Withers, S. G., & Alzari, P. M. (2004). Structural Insights into the Catalytic Mechanism of Trypanosoma cruzi trans-Sialidase. Structure, 12(5), 775–784. https://doi.org/10.1016/j.str.2004.02.036

[35] Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., & Henrissat, B. (2013). The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Research, 42(D1), D490–D495. https://doi.org/10.1093/nar/gkt1178

[36] Klumpp, K., & Graves, B. (2006). Optimization of Small Molecule Drugs Binding to Highly Polar Target Sites: Lessons from the Discovery and Development of Neuraminidase Inhibitors. Current Topics in Medicinal Chemistry, 6(5), 423–434. https://doi.org/10.2174/156802606776743138

[37] Gema, L. R. S., Tolentino-Lopez, L. E., Martínez-Ramos, F., Padilla-Martínez, I., García-Machorro, J., & Correa-Basurto, J. (2015). Targeting a cluster of arginine residues of neuraminidase to avoid oseltamivir resistance in influenza A (H1N1): a theoretical study. Journal of Molecular Modeling, 21(1). https://doi.org/10.1007/s00894-014-2525-9

[38] Davies, G., & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. Structure, 3(9), 853–859. https://doi.org/10.1016/s0969-2126(01)00220-9

[39] Roggentin, P., Schauer, R., Hoyer, L. L., & Vimr, E. R. (1993). The sialidase superfamily and its spread by horizontal gene transfer. Molecular Microbiology, 9(5), 915–921. https://doi.org/10.1111/j.1365-2958.1993.tb01221.x

[40] Wolfenden, R., Lu, X., & Young, G. (1998). Spontaneous Hydrolysis of Glycosides. Journal of the American Chemical Society, 120(27), 6814–6815. https://doi.org/10.1021/ja9813055

[41] Lipničanová, S., Chmelová, D., Ondrejovič, M., Frecer, V., & Miertuš, S. (2020). Diversity of sialidases found in the human body – A review. International Journal of Biological Macromolecules, 148, 857–868. https://doi.org/10.1016/j.ijbiomac.2020.01.123

[42] Yeung, J. H. F., Telford, J. C., Shidmoossavee, F. S., Bennet, A. J., Taylor, G. L., & Moore, M. M. (2013). Kinetic and Structural Evaluation of Selected Active Site Mutants of the Aspergillus fumigatus KDNase (Sialidase). Biochemistry, 52(51), 9177–9186. https://doi.org/10.1021/bi401166f

[43] ExplorEnz: Search the Official IUBMB Enzyme List. (2009). The Enzyme Database. Retrieved September 2021, from https://www.enzyme-database.org/

[44] Lentz, M. R., Webster, R. G., & Air, G. M. (1987). Site-directed mutation of the active site of influenza neuraminidase and implications for the catalytic mechanism. Biochemistry, 26(17), 5351–5358. https://doi.org/10.1021/bi00391a020

[45] Neal Waxham, M., & Aronowski, J. (1988). Identification of amino acids involved in the sialidase activity of the mumps virus hemagglutinin-neuraminadase protein. Virology, 167(1), 226–232. https://doi.org/10.1016/0042-6822(88)90072-4

[46] Luo, Y., Li, S. C., Li, Y. T., & Luo, M. (1999). The 1.8 å structures of leech intramolecular trans - sialidase complexes: evidence of its enzymatic mechanism 1 1Edited by R. Huber. Journal of Molecular Biology, 285(1), 323–332. https://doi.org/10.1006/jmbi.1998.2345

[47] Schwarzer, D., Browning, C., Stummeyer, K., Oberbeck, A., Mühlenhoff, M., Gerardy-Schahn, R., & Leiman, P. G. (2015). Structure and biochemical characterization of bacteriophage phi92 endosialidase. Virology, 477, 133–143. https://doi.org/10.1016/j.virol.2014.11.002

[48] Magesh, S., Suzuki, T., Miyagi, T., Ishida, H., & Kiso, M. (2006). Homology modeling of human sialidase enzymes NEU1, NEU3 and NEU4 based on the crystal structure of NEU2: Hints for the design of selective NEU3 inhibitors. Journal of Molecular Graphics and Modelling, 25(2), 196–207. https://doi.org/10.1016/j.jmgm.2005.12.006

[49] Richichi, B., Lunghi, C., Papakyriakou, A., Francesconi, O., & Nativi, C. (2013). Sialylexoenitols as precursors for new analogues of sialidase inhibitors. Pure and Applied Chemistry, 85(9), 1803–1811. https://doi.org/10.1351/pac-con-12-11-08

[50] Minami, A., Meguro, Y., Ishibashi, S., Ishii, A., Shiratori, M., Sai, S., Horii, Y., Shimizu, H., Fukumoto, H., Shimba, S., Taguchi, R., Takahashi, T., Otsubo, T., Ikeda, K., & Suzuki, T. (2017). Rapid regulation of sialidase activity in response to neural activity and sialic acid removal during memory processing in rat hippocampus. Journal of Biological Chemistry, 292(14), 5645–5654. https://doi.org/10.1074/jbc.m116.764357

[51] Annunziata, I., Patterson, A., Helton, D., Hu, H., Moshiach, S., Gomero, E., Nixon, R., & D'Azzo, A. (2013). Lysosomal NEU1 deficiency affects amyloid precursor protein levels and amyloid-β secretion via deregulated lysosomal exocytosis. Nature Communications, 4(1). https://doi.org/10.1038/ncomms3734

[52] Schneider, J. S., Seyfried, T. N., Choi, H. S., & Kidd, S. K. (2015). Intraventricular Sialidase Administration Enhances GM1 Ganglioside Expression and Is Partially Neuroprotective in a Mouse

Model of Parkinson's Disease. PLOS ONE, 10(12), e0143351. https://doi.org/10.1371/journal.pone.0143351

[53] OMIM - online mendelian inheritance in man. (2021). Online Mendelian Inheritance in Man. Retrieved November 2021, from https://www.omim.org/

[54] M. Itzstein, M., Wu, W. Y., Kok, G. B., Pegg, M. S., Dyason, J. C., Jin, B., van Phan, T., Smythe, M. L., White, H. F., Oliver, S. W., Colman, P. M., Varghese, J. N., Ryan, D. M., Woods, J. M., Bethell, R. C., Hotham, V. J., Cameron, J. M., & Penn, C. R. (1993). Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature, 363(6428), 418–423. https://doi.org/10.1038/363418a0

[55] Rahman, M. M., Hirokawa, T., Tsuji, D., Tsukimoto, J., Hitaoka, S., Yoshida, T., Chuman, H., & Itoh, K. (2015). Novel pH-dependent regulation of human cytosolic sialidase 2 (NEU2) activities by siastatin B and structural prediction of NEU2/siastatin B complex. Biochemistry and Biophysics Reports, 4, 234–242. https://doi.org/10.1016/j.bbrep.2015.09.017

[56] Telford, J. C., Yeung, J. H., Xu, G., Kiefel, M. J., Watts, A. G., Hader, S., Chan, J., Bennet, A. J., Moore, M. M., & Taylor, G. L. (2011). The Aspergillus fumigatus Sialidase Is a 3-Deoxy-d-glycero-d-galacto-2-nonulosonic Acid Hydrolase (KDNase). Journal of Biological Chemistry, 286(12), 10783–10792. https://doi.org/10.1074/jbc.m110.207043

[57] Kubota, M., Takeuchi, K., Watanabe, S., Ohno, S., Matsuoka, R., Kohda, D., Nakakita, S. I., Hiramatsu, H., Suzuki, Y., Nakayama, T., Terada, T., Shimizu, K., Shimizu, N., Shiroishi, M., Yanagi, Y., & Hashiguchi, T. (2016). Trisaccharide containing α2,3-linked sialic acid is a receptor for mumps virus. Proceedings of the National Academy of Sciences, 113(41), 11579–11584. https://doi.org/10.1073/pnas.1608383113

[58] Gaskell, A., Crennell, S., & Taylor, G. (1995). The three domains of a bacterial sialidase: a β-propeller, an immunoglobulin module and a galactose-binding jelly-roll. Structure, 3(11), 1197–1205. https://doi.org/10.1016/s0969-2126(01)00255-6

[59] Schulz, E. C., Neumann, P., Gerardy-Schahn, R., Sheldrick, G. M., & Ficner, R. (2010). Structure analysis of endosialidase NF at 0.98 Å resolution. Acta Crystallographica Section D Biological Crystallography, 66(2), 176–180. https://doi.org/10.1107/s0907444909048720

[60] Lin, Y. P., Gregory, V., Collins, P., Kloess, J., Wharton, S., Cattle, N., Lackenby, A., Daniels, R., & Hay, A. (2010). Neuraminidase Receptor Binding Variants of Human Influenza A(H3N2) Viruses Resulting from Substitution of Aspartic Acid 151 in the Catalytic Site: a Role in Virus Attachment? Journal of Virology, 84(13), 6769–6781. https://doi.org/10.1128/jvi.00458-10

[61] Copley, R. R., Russell, R. B., & Ponting, C. P. (2001). Sialidase-like Asp-boxes: Sequence-similar structures within different protein folds. Protein Science, 10(2), 285–292. https://doi.org/10.1110/ps.31901

[62] Quistgaard, E. M., & Thirup, S. S. (2009). Sequence and structural analysis of the Asp-box motif and Asp-box beta-propellers; a widespread propeller-type characteristic of the Vps10 domain family and several glycoside hydrolase families. BMC Structural Biology, 9(1), 46. https://doi.org/10.1186/1472-6807-9-46

13. ACRONYMS

Asp - Aspartate

Arg - Arginine

CAZy – Carbohydrate-active enzymes database

DANA - 2,3-dehydro-2-deoxy-N-acetylneuraminic acid

EC – Enzyme Commission

F – Fusion protein

G - Glycoprotein

GH – Glycosyl hydrolase

Gln - Glutamine

Glu – Glutamate

GM1 - Monosialotetrahexosylganglioside

HA - Hemagglutinin

HN – Hemagglutinin-Neuraminidase

IT-sialidases – Intramolecular sialidases

IUPAC – Interational Union of Pure and Applied Chemistry

NA - Neuraminidase

NANA – Neuraminic acid

OMIM – Online Mendelian Inheritance in Man

PDB – Protein Data Bank database

Ser - Serotonine

Trp - Tryptophan

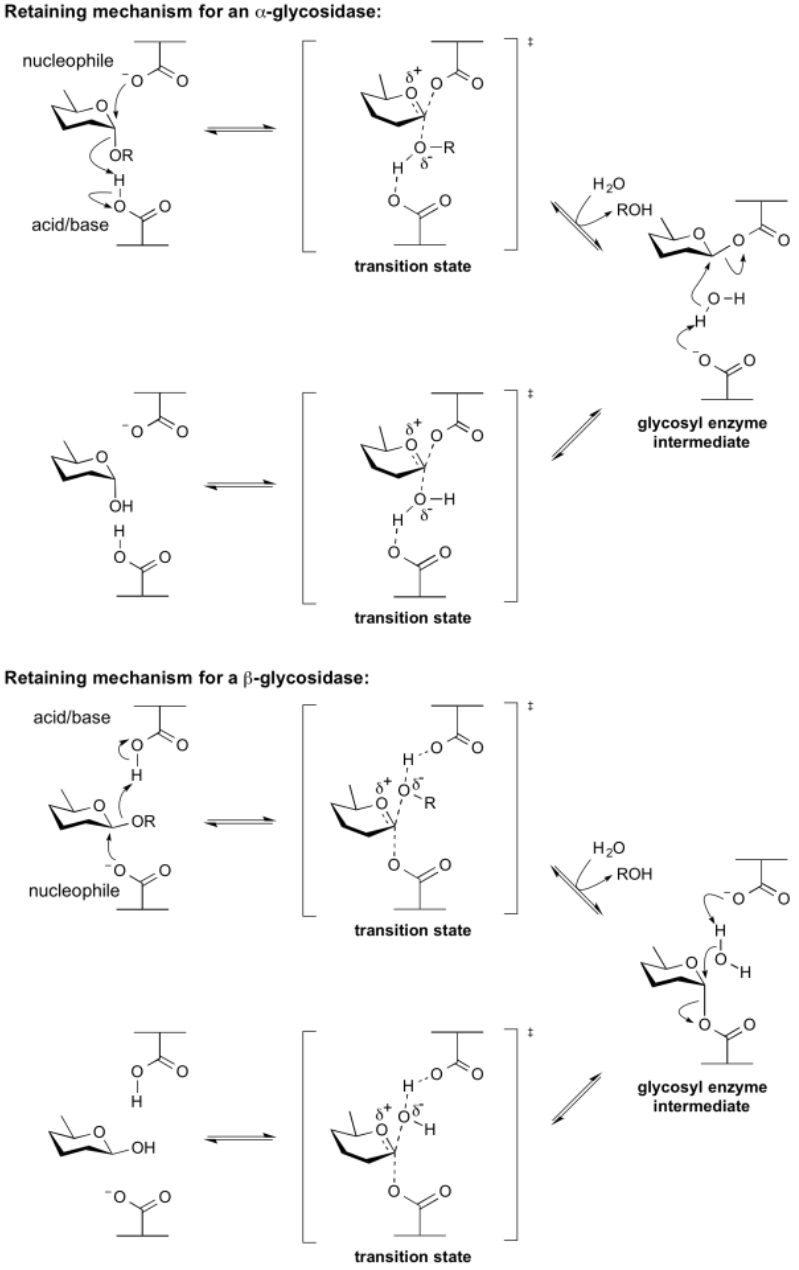
Tyr - Tyrosine

VMD – Visual Molecular Dynamics

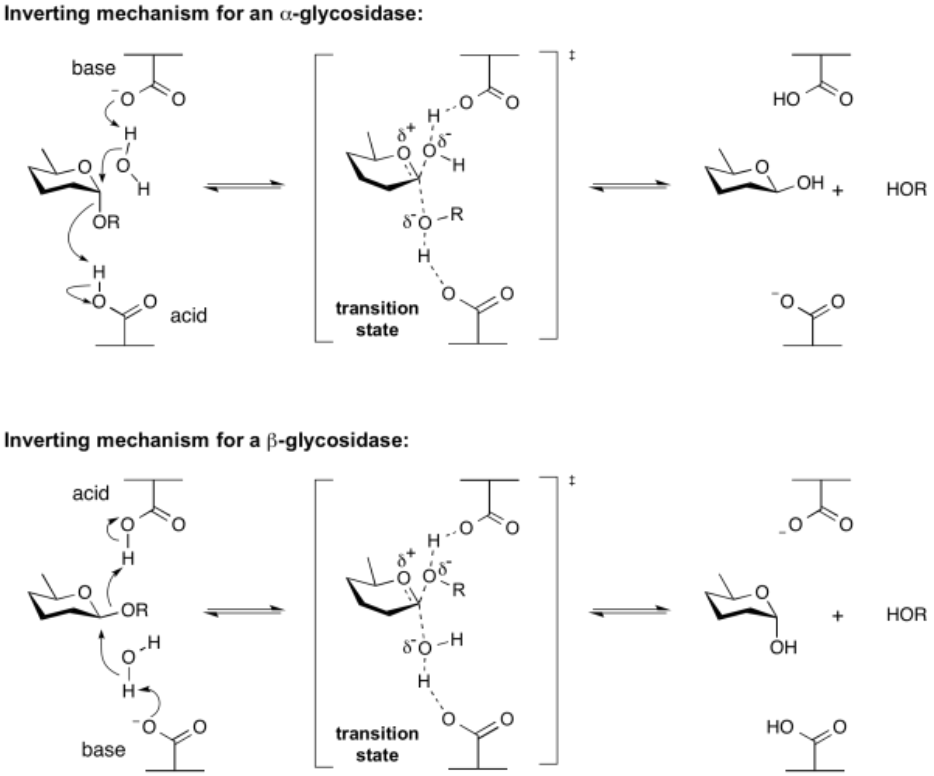
5NeuAc – Neuraminic acid

APPENDICES

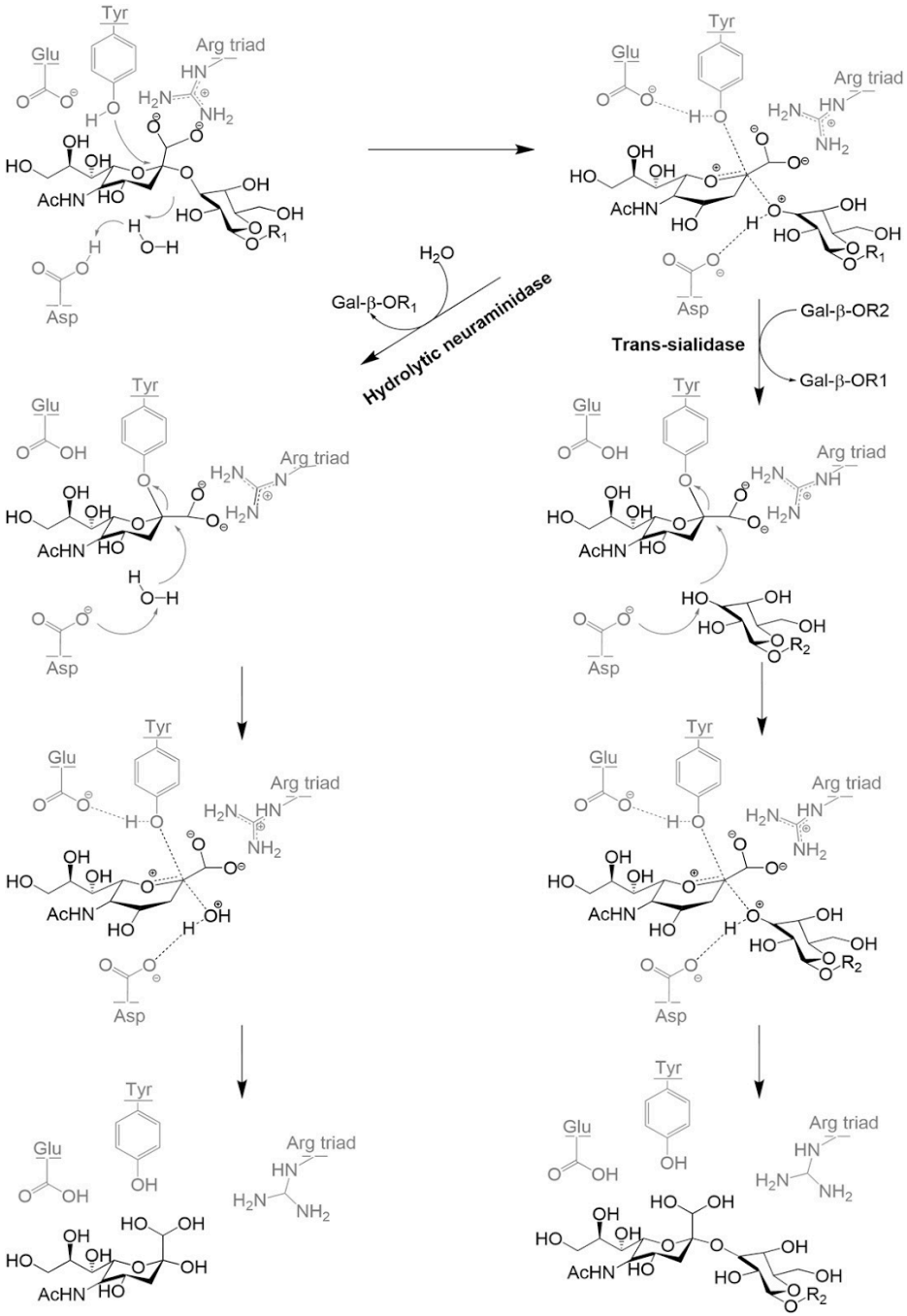
APPENDIX 1:MOLECULAR REACTION MECHANISMS



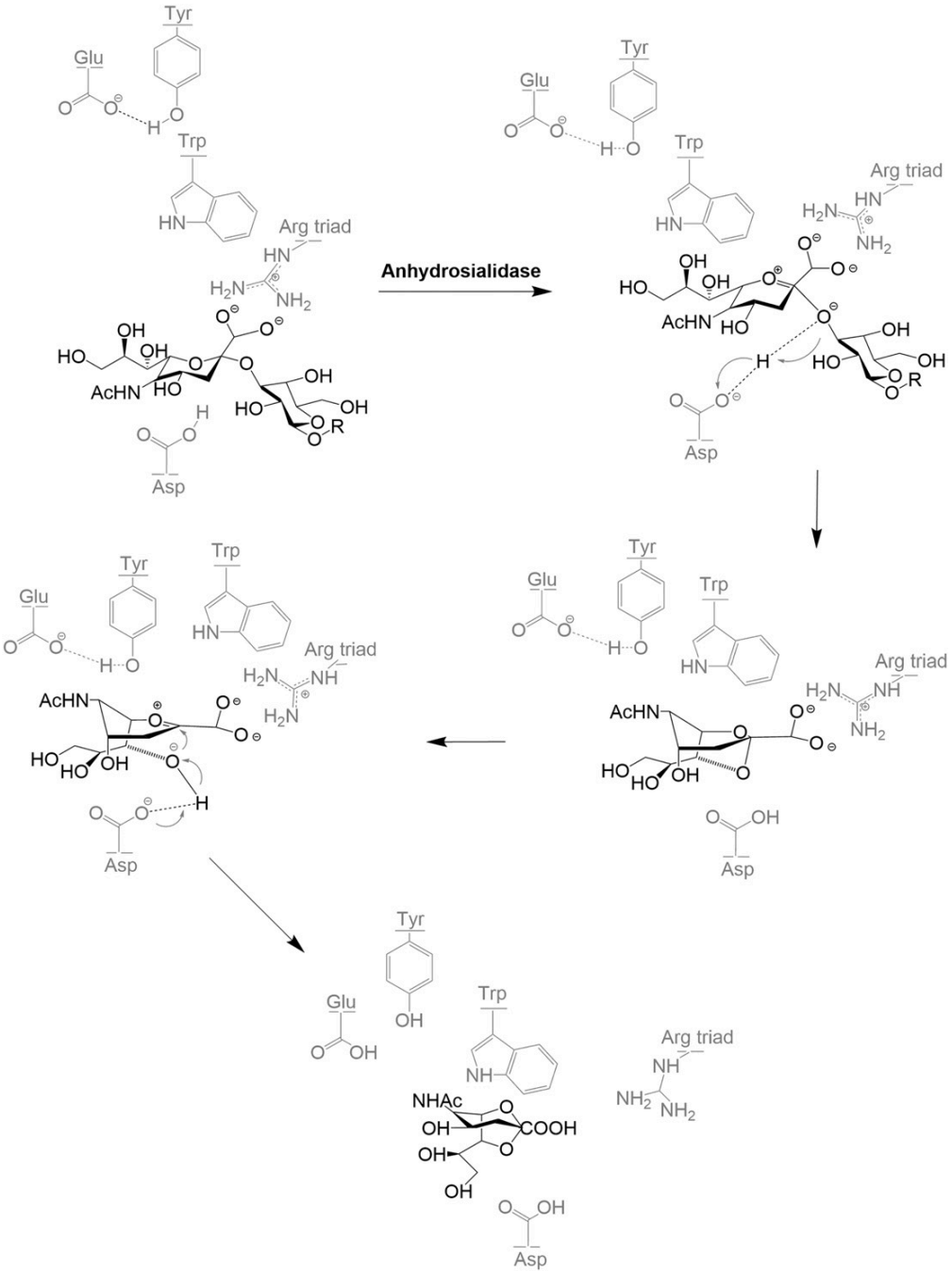
Glycosidase retaining mechanism, *originally published in Cazypedia* [17]



Glycosidase inverting mechanism, *originally published in Cazypedia* [17]



Retaining mechanism for hydrolytic and trans-neuraminidases, *originally published by S. Lipničanová [41]*



Intramolecular mechanism for anhydro-sialidases, *originally published by S. Lipničanová [41]*

APPENDIX 2: PROPOSED CLASSIFICATION TABLE

Family	Enzyme	Type	Subtype	Genus	Human related pathogenesis	EC number	Organism	PDB code	Mechanism
GH33	sialidase Hz136	PROKARYOTA	NON-PATHOGENIC BACTERIA	ALISTIPES	GUT BACTERIA	3.2.1.18	Alistipes	6MNJ 2.20	RETAINING
GH33	sialidase (BTSA)	PROKARYOTA	PATHOGENIC BACTERIA	BACTEROIDES	SEPSIS, ESTEROCOLITIS, PERITONITIS	3.2.1.18	Bacteroides thetaiotaomicron VPI-5482	4BBW 2.30	RETAINING
GH33	sialidase (NanI)	PROKARYOTA	PATHOGENIC BACTERIA	CLOSTRIDIUM	GAS GANGRENE, ENTEROCOLITIS	3.2.1.18	Clostridium perfringens str. 13	2VK5 0.97	RETAINING
GH33	sialidase (NedA)	PROKARYOTA	NON-PATHOGENIC BACTERIA	ACTINOBACTERIA	GUT BACTERIA	3.2.1.18	Micromonospora viridifaciens short form (41KDa) Micromonospora viridifaciens long form (68KDa)	1EUR 1.82 1EUU 2.50	RETAINING
GH33	sialidase (NanH) INTRAMOLECULAR TRANS-SIALIDASE	PROKARYOTA	PATHOGENIC BACTERIA	SALMONELLA	GASTROENTERITIS	4.2.2.15	Salmonella typhimurium TA262	3SIL 1.05	RETAINING
GH33	neuraminidase (NanC)	PROKARYOTA	PATHOGENIC BACTERIA	STREPTOCOCCUS	PNEUMONIA, OTITIS MEDIA	3.2.1.18	Streptococcus pneumoniae G54	4YW4 2.20	RETAINING
GH33	sialidase (NanA)	PROKARYOTA	PATHOGENIC BACTERIA	STREPTOCOCCUS	PNEUMONIA, OTITIS MEDIA	3.2.1.18	Streptococcus pneumoniae R6	3H71 1.70	RETAINING
GH33	intramolecular trans-sialidase (NanB)	PROKARYOTA	PATHOGENIC BACTERIA	STREPTOCOCCUS	PNEUMONIA, OTITIS MEDIA	4.2.2.15	Streptococcus pneumoniae TIGR4	2VW2 1.7	RETAINING
GH33	sialidase 26	PROKARYOTA	NON-PATHOGENIC BACTERIA	ESCHERICHIA	GUT BACTERIA	3.2.1.18	Escherichia coli	6MRX 2.00	RETAINING
GH33	Vibrio cholerae neuraminidase (VCNA)	PROKARYOTA	PATHOGENIC BACTERIA	VIBRIO	CHOLERA	3.2.1.18	Vibrio cholerae 569B 395	1W0O 1.9	RETAINING
GH33	2-keto-3-deoxynononic acid sialidase (KDNase)	EUKARYOTA	FUNGI	APERGILLUS	RESPIRATORY DISEASE IN IMMUNOCOMPROMISED INDIVIDUALS	3.2.1.-	Aspergillus fumigatus Af293	2XZI 1.45	RETAINING
GH33	cytosolic sialidase (Neu2)	EUKARYOTA	MAMMALIA	HUMAN	CANCER	3.2.1.18	Homo sapiens	1SNT 1.75	RETAINING
GH33	intramolecular trans-sialidase L (MDSA)	EUKARYOTA	INVERTEBRATES	LEECH	NON-PATHOGENIC	4.2.2.15	Macrobdella decora	1SLL 2.00	RETAINING
GH33	trans-sialidase (TcTS)	EUKARYOTA	INVERTEBRATES	TRYPANOSOMA	CHAGAS DISEASE	2.4.1.-	Trypanosoma cruzi	1MS3 1.65	RETAINING
GH33	sialidase (TrSA)	EUKARYOTA	INVERTEBRATES	TRYPANOSOMA	NON-PATHOGENIC	2.4.1.-	Trypanosoma rangeli	1MZ5 2.20	RETAINING
GH34	neuraminidase (NA) N2	VIRUS	ORTHOMYXOVIRIDAE	INFLUENZA A VIRUS	INFLUENZA A VIRUS	3.2.1.18	Influenza A virus (A/Tokyo/3/1967(H2N2))	1NN2 2.20	RETAINING
GH34	neuraminidase (NA) N6	VIRUS	ORTHOMYXOVIRIDAE	INFLUENZA A VIRUS	INFLUENZA A VIRUS	3.2.1.18	Influenza A virus (A/chicken/Nanchang/7-010/2000(H3N6))	4QN4 1.80	RETAINING
GH34	neuraminidase (NA) N5	VIRUS	ORTHOMYXOVIRIDAE	INFLUENZA A VIRUS	INFLUENZA A VIRUS	3.2.1.18	Influenza A virus (A/duck/Alberta/60/1976(H12N5))	3SAL 1.50	RETAINING
GH34	neuraminidase (NA) N8	VIRUS	ORTHOMYXOVIRIDAE	INFLUENZA A VIRUS	INFLUENZA A VIRUS	3.2.1.18	Influenza A virus (A/duck/Ukraine/1/1963(H3N8)) (A/DUCK/UKRAINE/1/63)	2HT5 2.40	RETAINING
GH34	neuraminidase (NA) N9	VIRUS	ORTHOMYXOVIRIDAE	INFLUENZA A VIRUS	INFLUENZA A VIRUS	3.2.1.18	Influenza A virus (A/Hangzhou/2/2013(H7N9))	4MWJ 1.80	RETAINING
GH34	neuraminidase (NA) N10	VIRUS	ORTHOMYXOVIRIDAE	INFLUENZA A VIRUS	INFLUENZA A VIRUS	3.2.1.18	Influenza A virus (A/little yellow-shouldered bat/Guatemala/060/2010(H17N10))	4GDJ 2.00	RETAINING
GH34	neuraminidase (NA) N7	VIRUS	ORTHOMYXOVIRIDAE	INFLUENZA A VIRUS	INFLUENZA A VIRUS	3.2.1.18	Influenza A virus (A/mallard/ALB/196/1996(H10N7))	4QN3 2.09	RETAINING
GH34	neuraminidase (NA) N4	VIRUS	ORTHOMYXOVIRIDAE	INFLUENZA A VIRUS	INFLUENZA A VIRUS	3.2.1.18	Influenza A virus (A/mink/Sweden/E12665/84(H10N4))	2HTV 2.80	RETAINING
GH34	neuraminidase (NA) N3	VIRUS	ORTHOMYXOVIRIDAE	INFLUENZA A VIRUS	INFLUENZA A VIRUS	3.2.1.18	Influenza A virus (A/swine/Missouri/2124514/2006(H2N3))	4HZV 1.80	RETAINING
GH34	neuraminidase (NA) N1	VIRUS	ORTHOMYXOVIRIDAE	INFLUENZA A VIRUS	INFLUENZA A VIRUS	3.2.1.18	Influenza A virus (A/Viet Nam/1203/2004(H5N1))	2HTY 2.50	RETAINING

Family	Enzyme	Type	Subtype	Genus	Human related pathogenesis	EC number	Organism	PDB code	Mechanism
GH34	sialidase (NA) INFLUENZA VIRUS B	VIRUS	ORTHOMYXOVIRIDAE	INFLUENZA B VIRUS	B/VICTORIA	3.2.1.18	Influenza B virus (B/Brisbane/60/2008)	4CPL 2.00	RETAINING
GH34	sialidase (NA) INFLUENZA VIRUS B	VIRUS	ORTHOMYXOVIRIDAE	INFLUENZA B VIRUS	B/YAMAGATA	3.2.1.19	Influenza B virus (B/Perth/211/2001)	3K36 2.20	RETAINING
GH58	endosialidase (endoN92)	VIRUS	BACTERIOPHAGE	INVERTING SIALIDASE	PHAGE PHI92	3.2.1.129	Enterobacteria phage phi92 ATCC 35860-B1	4HIZ 1.60	INVERTING
GH58	endosialidase NF (EndoNF)	VIRUS	BACTERIOPHAGE	INVERTING SIALIDASE	PATHOGENIC FOR BACTERIA	3.2.1.129	Escherichia phage K1F	3JU4 0.98	INVERTING
GH83	hemagglutinin-neuraminidase avian orthovulavirus	VIRUS	PARAMYXOVIRIDAE	AVULAVIRUS	NO, IN ANIMALS NEW CASTLE DISEASE VIRUS	3.2.1.18	Avian avulavirus 1 (KENSAS)	1E8U 2.0	RETAINING
GH83	hemagglutinin-neuraminidase (hPIV-3 HN)	VIRUS	PARAMYXOVIRIDAE	RESPIROVIRUS	PARAINFLUENZA VIRUS TYPE 3	3.2.1.18	Human parainfluenza 3 virus (strain NIH 47885)	4MZA 1.65	RETAINING
GH83	hemagglutinin-neuraminidase Mumps virus	VIRUS	PARAMYXOVIRIDAE	RUBULAVIRUS	MUMPS VIRUS	3.2.1.18	Mumps virus Miyahara vaccine (MIYAHARA VACCINE)	5B2C 2.24	RETAINING
GH83	hemagglutinin-neuraminidase (SV5 HN)	VIRUS	PARAMYXOVIRIDAE	RUBULAVIRUS	PARAINFLUENZA VIRUS TYPE 5	3.2.1.18	Parainfluenza virus 5 and Simian virus 5 (strain W3) 5 (W3)	1Z4Y 2.60	RETAINING
GH83	hemagglutinin-neuraminidase (SosV HN)	VIRUS	PARAMYXOVIRIDAE	RUBULAVIRUS	SOSUGA VIRUS	3.2.1.18	Sosuga virus	6SG8 2.50	RETAINING