



HAL
open science

Molecular and evolutionary bases of the biosynthesis of polysialylation : study of fish polysialyltransferases enzymatic specificities

Mathieu Decloquement

► To cite this version:

Mathieu Decloquement. Molecular and evolutionary bases of the biosynthesis of polysialylation : study of fish polysialyltransferases enzymatic specificities. Biochemistry, Molecular Biology. Université de Lille, 2023. English. NNT : 2023ULILS108 . tel-04691831

HAL Id: tel-04691831

<https://theses.hal.science/tel-04691831v1>

Submitted on 9 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITY OF LILLE – FACULTY OF SCIENCE AND TECHNOLOGY

Laboratory of Structural and Functional Glycobiology - UMR CNRS / USTL n°8576

DOCTORAL THESIS

Doctoral School of Health and Biology – Lille

To obtain the degree of PhD of Lille University by

Mathieu DECLOQUEMENT

Molecular and evolutionary bases of the biosynthesis of polysialylation:

Study of fish polysialyltransferases enzymatic specificities

Discipline: Molecular and cellular aspects of biology

Specialty: Biochemistry and molecular biology

PhD defense September, 6th 2023

Examination board

President	Dr François Foulquier	University of Lille, CNRS, FRA
Reporters	PD Dr Martina Mühlhoff	University of Hannover, DEU
	Dr Ramon Hurtado-Guerrero	University of Zaragoza, ESP
Examiners	Dr Alexandra Castilho	University of Vienna, BOKU, AUT
	PD Dr Sebastian P. Galuska	FBN Institute, Dummerstorf, DEU
Supervisor	Dr Anne Harduin-Lepers	University of Lille, CNRS, FRA



UNIVERSITE DE LILLE – FACULTE DES SCIENCES ET TECHNOLOGIES

Laboratoire de glycobiologie structurale et fonctionnelle UMR CNRS / USTL n°8576

THESE DE DOCTORAT

Ecole Doctorale de Biologie Santé – Lille

Pour l'obtention du titre de Docteur de l'Université de Lille par

Mathieu DECLOQUEMENT

**Bases moléculaires et évolutives de la biosynthèse de la polysialylation :
Etude des spécificités enzymatiques des polysialyltransférases de poissons**

Discipline : Aspects moléculaires et cellulaires de la biologie

Spécialité : Biochimie et biologie moléculaire

Soutenance le 6 Septembre 2023

Commission d'examen

Président :	Dr François Foulquier	Université de Lille, CNRS, FRA
Rapporteurs :	PD Dr Martina Mühlhoff	Université de Hanovre, DEU
	Dr Ramon Hurtado-Guerrero	Université de Saragosse, ESP
Examineurs :	Dr Alexandra Castilho	Université de Vienne, BOKU, AUT
	PD Dr Sebastian P. Galuska	Institut FBN, Dummerstorf, DEU
Directrice :	Dr Anne Harduin-Lepers	Université de Lille, CNRS, FRA



FOREWORD

This work was supervised by Dr **Anne Harduin-Lepers**

Within the Structural and Functional Glycobiology Unit (UGSF),

Mixed Research Unit (UMR) 8576 - National Center for Scientific Research (CNRS),

Located at the University of Lille in France and currently headed by Pr Yann Guérardel.

This thesis benefited from several financial supports from a grant from the French Ministry of Higher Education and Research (MESR), co-funding from the Hauts-de-France region. The work resulting from this thesis has been published in international scientific journals and presented as oral communications and posters during scientific meetings.

AVANT-PROPOS

Ce travail a été encadré par le Dr **Anne Harduin-Lepers**

Au sein de l'Unité de Glycobiologie Structurale et Fonctionnelle (UGSF),

Unité Mixte de Recherche (UMR) 8576 - Centre National de Recherche Scientifique (CNRS),

Basée à l'Université de Lille en France et actuellement dirigée par le Pr Yann Guérardel.

Cette thèse a bénéficié de plusieurs soutiens financiers provenant d'une allocation du Ministère de l'Enseignement Supérieur et de la Recherche (MESR), d'un co-financement de la région des Hauts-de-France. Les travaux résultant de cette thèse ont été publiés dans des journaux scientifiques internationaux et présentés sous forme de communications orales et posters au cours d'évènements scientifiques.

PUBLICATIONS AND COMMUNICATIONS

I. SCIENTIFIC PUBLICATIONS

Decloquement M, Venuto MT, Cogez V, Schulz C, Lion C, Noel M, Steinmetz A, Rigolot V, Biot C, Rebl A, Galuska SP, Harduin-Lepers A. **Salmonid polysialyltransferases to generate a variety of sialic acid polymers**. Nature Scientific Reports. 2023.

Cogez V, Vicogne D, Schulz C, Portier L, Venturi G, de Ruyck J, Decloquement M, Lensink M, Brysbaert G, Dall'Olio F, Groux-Degroote S, Harduin-Lepers A. **N-glycan on the non-consensus N-X-C glycosylation site impacts activity, stability and localization of the Sd^a synthase B4GALNT2**. International Journal of Molecular Sciences. 2023.

Fliniaux I, Marchand G, Molinaro C, Decloquement M, Martoriati A, Marin M, Bodart J-F, Harduin-Lepers A, Caillau K. Review : **Diversity of sialic acids and sialoglycoproteins in gametes and at fertilization**. Front Cell Dev Biol. 2022;10:982931. (Review).

Venuto MT, Decloquement M, Martorell Ribera J, Noel M, Rebl A, Cogez V, Petit D, Galuska SP, Harduin-Lepers A. **Vertebrate Alpha2,8-Sialyltransferases (ST8Sia) : A Teleost Perspective**. International Journal of Molecular Sciences. 2020;21(2):513.

Chang L-Y, Teppa E, Noel M, Gilormini P-A, Decloquement M, Lion C, Biot C, Mir A-M, Cogez V, Delannoy P, Khoo KH, Petit D, Guérardel Y, Harduin-Lepers A. **Novel Zebrafish Mono- α 2,8-sialyltransferase (ST8Sia VIII): An Evolutionary Perspective of α 2,8-Sialylation**. International Journal of Molecular Sciences. 2019;20(3):622.

II. ORAL COMMUNICATIONS

Decloquement M, Harduin-Lepers A. Enzymatic characterization of salmonid polysialyltransferases: substrate specificities contribute to diversity of fish polySia. Talk presented at: **UGSF Weekly Seminar** ; 2023 Mar 3 ; Villeneuve d'Ascq, FR.

Decloquement M, Venuto MT, Cogez V, Noel M, Lion C, Rigolot V, Biot C, Galuska SP, Harduin-Lepers A. Diversity of polysialylation machinery in fish species highlights exciting perspectives to generate original high therapeutic biomaterials. 3-min thesis video presented at: **6th Latin American Glycobiology Congress** ; 2021 Oct 5-8 ; Mexico City, MX (Online).

➔ **AWARD : BEST « 3-MIN THESIS » PRESENTATION.**

Decloquement M, Venuto MT, Cogez V, Noel M, Lion C, Rigolot V, Biot C, Galuska SP, Harduin-Lepers A. Diversity of polysialylation machinery in fish species highlights exciting perspectives to generate original high therapeutic biomaterials. Talk presented at: **INNOGLY-ECI-WG3-2021: Young Glyco-Scientists on stage** ; 2021 Sep 27 ; Firenze, IT (Online).

Decloquement M, Venuto MT, Cogez V, Noel M, Lion C, Rigolot V, Biot C, Galuska SP, Harduin-Lepers A. Diversity of polysialylation machinery in fish species highlights exciting perspectives to generate original high therapeutic biomaterials. 5-min thesis video presented at: **21^e Edition Journée André Verbert des Doctorants** ; 2021 Sep 22 ; Lille, FR (Online).

➔ **TOP18 BEST « 5-MIN THESIS » PRESENTATION.**

Decloquement M, Cogez V, Noel M, Lion C, Rigolot V, Biot C, Harduin-Lepers A. Innovative tools for the study of sialylation processes: Glycosyltransferases engineering and use of unnatural sialic acid donors. Talk presented at: **Congrès Annuel de la Société Française de Biochimie et Biologie Moléculaire** ; 2021 Jul 1-2 ; Paris, FR.

Decloquement M, Cogez V, Noel M, Lion C, Rigolot V, Biot C, Harduin-Lepers A. Innovative tools for the study of sialylation processes: Glycosyltransferases engineering and use of unnatural sialic acid donors. Talk presented at: **International Scientific CDG Symposium** ; 2021 Jun 23-24 ; Leuven, BE (Online).

Decloquement M, Cogez V, Noel M, Lion C, Rigolot V, Biot C, Harduin-Lepers A. Innovative tools for the study of sialylation processes: Glycosyltransferases engineering and use of unnatural sialic acid donors. Talk presented at: **12th Edition: Young Researchers in Life Sciences** ; 2021 Jun 16-18 ; Paris, FR (Online).

Decloquement M, Harduin-Lepers A. Molecular and evolutionary bases of the biosynthesis of sialylated determinants: Study of the fish sialylation machinery. Talk presented at: **UGSF Weekly Seminar**; 2021 Jan 29 ; Lille, FR (Online).

Decloquement M, Noel M, Lion C, Biot C, Cogez V, Guerardel Y, Harduin-Lepers A. Molecular phylogeny and exogenous sialylation approaches identify a novel α 2,8-sialyltransferase in zebrafish : ST8Sia VIII. Talk presented at: **30th Joint Glycobiology Meeting** ; 2019 Oct 27-29 ; Villeneuve d'Ascq, FR.

III. POSTER COMMUNICATIONS

Decloquement M, Venuto MT, Cogez V, Steinmetz A, Lion C, Rigolot V, Szydlowski N, Biot C, Galuska SP, Harduin-Lepers A. Diversity of polysialylation machinery in fish species highlights exciting perspectives to generate original high therapeutic biomaterials. **28^e Journées du Groupe Français des Glycosciences** ; 2022 May 30 – Jun 3 ; Branville, FR.

→ **AWARD : BEST POSTER PRESENTATION.**

Decloquement M, Venuto MT, Cogez V, Steinmetz A, Lion C, Rigolot V, Szydlowski N, Biot C, Galuska SP, Harduin-Lepers A. Diversity of polysialylation machinery in fish species highlights exciting perspectives to generate original high therapeutic biomaterials. **INNOGLY COST Action CA18103: Innovation with Glycans : New frontiers from synthesis to new biological targets** ; 2022 May 4-6 ; Lugano, CH.

→ **AWARD : BEST POSTER PRESENTATION.**

Decloquement M, Venuto MT, Cogez V, Noel M, Lion C, Rigolot V, Biot C, Galuska SP, Harduin-Lepers A. Diversity of polysialylation machinery in fish species highlights exciting perspectives to generate original high therapeutic biomaterials. **New Frontiers Symposium on Translational Glycoscience – Radboud umc** ; 2021 Nov 18-19 ; Nijmegen, NL (Online).

Decloquement M, Venuto MT, Cogez V, Noel M, Lion C, Rigolot V, Biot C, Galuska SP, Harduin-Lepers A. Diversity of polysialylation machinery in fish species highlights exciting perspectives to generate original high therapeutic biomaterials. **31st Joint Glycobiology Meeting** ; 2021 Oct 7-8 ; Hannover, DE (Online).

Decloquement M, Noel M, Cogez V, Lion C, Biot C, Guerardel Y, Harduin-Lepers A. Distinct specificities between human and zebrafish ST6Gal I. **GlycoT2020** ; 2020 Jun 21-23 ; Boston, US (Online).

COPYRIGHT STATEMENT AND REPRODUCTION

All figures taken from scientific publications and used in this manuscript have obtained permission to reproduce as reported on the last page and their authors have been cited.

All original works produced in this manuscript should obtain express consent before reproduction or public use.

DECLARATION DE DROITS D'AUTEUR ET REPRODUCTION

Toutes les figures tirées de publications scientifiques et utilisées dans ce manuscrit ont obtenu une autorisation de reproduction, rapportée en dernière page, et leurs auteurs ont été cités.

Toutes les oeuvres originales produites dans ce manuscrit doivent obtenir un consentement expressement formulés avant reproduction ou utilisation publique.

ACKNOWLEDGMENTS

First of all, I would like to thank all members of the jury comitee that accept to evaluate my PhD, it's a privilege to have such expert scientists involved in the examination of my thesis manuscript. I hope to have great discussions time with you during my PhD defense.

Au Dr **François Foulquier**, c'est un honneur pour moi que tu acceptes d'être le président de mon jury de thèse. Plus de 5 ans se sont écoulés depuis mon arrivée au laboratoire et nos premières discussions. Je partage bons nombres de tes avis et du regard extra-lucide que tu portes sur bien des situations de vie personnelles et professionnelles. Merci d'avoir autant contribué à mon développement avec tes conseils et tes remarques pointues de scientifique passionné et passionnant.

To Dr **Martina Mühlenhoff**, I would like to express my deepest gratitude that you accept to be reporter of my thesis. As a world expert involved in autopolysialylation discovery and development of tools to study polysialic acids, I hope my manuscript will be pleasant to read and will raise good discussions. Often collaborating with UGSF members, thank you again for material gifts as mAb735 and the various endoN versions and for your precious advices on their use.

To Dr **Ramon Hurtado-Guerrero**, thank you very much to be in my thesis committee, to read and evaluate my thesis manuscript. I have in mind our first exciting discussions about my poster during Lugano INNOGLY meeting which enabled us to develop longer conversations on strategies to optimize the purification of recombinant sialyltransferases and preserve their activity.

To Dr **Alexandra Castilho**, it's a real pleasure to count on your presence among the examiners of my thesis commitee. We had a great discussion during GFG meeting in Branville about my poster and thesis project and I hope that my progress since that moment will raise good discussions during my oral defense.

To Dr **Sebastian Galuska**, my PhD defense definitely could not happen without your presence. Our collaboration was a great opportunity to come in your lab during my thesis. Thank you very much for your hospitality, your support and your advices about my thesis and the future. I hope you will appreciate my manuscript and that my thesis work will contribute to strengthen the collaboration between our two laboratories.

Enfin dans ce comité, j'adresse un remerciement tout particulier au Dr **Anne Harduin-Lepers**, ma directrice de thèse, pour m'avoir guidé tout au long de cette sinieuse et passionnante aventure qu'est la thèse. Dès la fin de mon cursus de licence, vous avez décelé chez moi un potentiel et vous m'avez accompagné pour me permettre d'arriver jusque là. Pendant mon master, vous m'avez proposé ce défi original d'allier outils biochimiques et bioinformatiques pour tenter de répondre à

de grandes hypothèses phylogénétiques et faire le lien entre évolution moléculaires des espèces et spécificités enzymatiques des sialyltransférases. Dans cette thèse, vous m'avez à nouveau témoigné de votre confiance, encouragé à présenter mes travaux en congrès, à rencontrer des scientifiques et à faire un séjour dans un laboratoire étranger, permis de suivre des formations et de faire des enseignements et encadrements d'étudiants. Tous ces éléments m'ont été tellement bénéfiques. Votre rencontre m'a permis de me développer personnellement, je suis admiratif et essaie au mieux de m'inspirer de votre patience et votre mental, votre pédagogie, votre énorme culture scientifique et votre exigence du détail qui font de votre travail ce qu'il est. Le défi était de taille, et encore plus dans un contexte de pandémie mondiale, mais, ensemble, nous avons su aller au-delà des moments de doutes et de difficultés qui font partie d'une thèse. J'ai apprécié ce temps passé sous votre direction, de même que les discussions scientifiques et humaines lors de nos échanges quotidiens. Merci pour le temps, la passion et l'énergie que vous m'avez consacré. J'espère que vous avez apprécié à la fois encadrer ma thèse et ma compagnie au laboratoire.

Then, I would like to thank all people dedicated in the **technical supervision** of my thesis work. Thank you to Virginie and Marzia for molecular cloning, to Cédric for NMR following of donor substrates, to Céline for exogenous sialylation and confocal microscopy, to Anna, Marzia, Gesine and Kristina for DMB-HPLC preparation of samples and analysis in Germany, to Nicolas for APTS-FACE analysis and to Guillaume and Elin for the use of AlphaFold and biocomputational tools. Thanks to all **financial supports** for my thesis: French ministry of Research, Doctoral School of Lille « EDBSL » and Région Hauts-de-France and finally to PHC Procope and Collège Doctoral Hauts-de-France which have made possible my stay in Germany.

Merci au Dr **Isabelle André** et au Pr **Abderrahman Maftah** d'avoir fait partie de mon Comité de Suivi Individuel en distanciel et de m'avoir permis d'aiguiser ma réflexion scientifique et d'explorer, de manière cohérente, des pistes pendant ces années de thèse. Merci à vous aussi pour vos conseils personnels et professionnels.

Par ailleurs, j'aimerais remercier les membres qui ont composé et composent l'équipe « 017 ». Merci à **Philippe Delannoy**, récemment devenu professeur émérite pour l'ensemble de son dévouement à la communauté scientifique, pour m'avoir conseillé, fait part de ses critiques pointues et constructives lors de nos captivantes discussions. C'était un privilège pour moi d'avoir bénéficié de votre expérience au sein de l'équipe. Je remercie également **Sophie Groux-Degroote** pour m'avoir bien accueilli dans l'équipe dès le début, pour avoir toujours stimulé ma réflexion avec des remarques scientifiques constructives sur mes travaux et pour m'avoir permis de, parfois, prendre du recul. J'ai beaucoup de respect et d'estime pour l'énergie et l'abnégation que tu portes à tes travaux de recherche et je te souhaite le meilleur pour la suite. Merci à **Virginie Coge**, qui a été la première à me faire confiance et à m'ouvrir les portes du laboratoire en Licence. J'ai

commencé mon éveil scientifique à tes côtés et je ne l'oublierai pas. De même, merci pour le partage de ta « fibre verte », qui m'a permis d'être plus sensibilisé à l'environnement et à nos impacts en tant que chercheurs. Merci à nos deux ingénieures **Céline** et **Dorothee**, les mamans à tout faire de l'équipe. Dès mes premiers stages, vous m'avez formé, transmis votre rigueur scientifique et avez toujours été disponibles pour prodiguer vos conseils. Merci Céline pour ce temps dédié à optimiser les réactions de sialylation exogène et de microscopie confocale, j'aime à penser que mes sialyltransférases de poisson ont fini par te plaire. De même, merci Dorothee pour toute ton implication en culture cellulaire. Votre présence au sein du laboratoire et de l'équipe est précieuse. Parmi les plus ou moins jeunes, j'ai une pensée pour **Sumeyye** et **Maxence**, mes prédécesseurs doctorants que j'ai pris comme exemple, qui m'ont motivé et soutenu pour l'aventure de la thèse. Pendant ma thèse, j'ai passé des moments précieux avec **Ryma** et **Clairène**, on a formé une petite équipe ensemble, merci de votre soutien. J'ai une aussi pensée positive pour **Charlotte** et **Angéline**, deux battantes de l'équipe et bien accrochées à l'UGSF. Je vous souhaite à toutes les quatre de poursuivre votre route avec détermination et je suis persuadé que vous parviendrez à vos objectifs. Merci aussi à **Elin** pour son aide précieuse et ses conseils en bioinformatique. Enfin, je n'oublie pas les plus jeunes encadrés et conseillés: Lou, Morgane, Sabrina, Laebah, Lou ou encore Valentin.

Nombreuses sont les personnes avec qui j'ai partagé des moments et qui m'ont permis d'apprécier ces années au C9. Merci d'abord au directeur d'Unité, **Yann Guérardel**, pour m'avoir permis de réaliser cette thèse au sein du laboratoire. Merci aussi pour les moments et les discussions intéressantes malgré ton emploi du temps ultra-chargé. Je prendrai ma revanche au baby-foot lors d'un prochain congrès, promis ! Merci à toute l'**équipe administrative** pour leur aide et leur organisation malgré des tâches parfois bien complexes. Merci aussi au **personnel technique et d'entretien** pour nous permettre de travailler dans les meilleures conditions possibles.

Comment ne pas mentionner les **professeurs** et **enseignants-chercheurs** de l'UGSF qui m'ont instruit, guidé ma pensée scientifique et permis de partager leur quotidien et leur expérience. La liste est longue mais chacun d'entre vous a contribué à ma formation et mon épanouissement. Merci à Tony, un modèle humainement et scientifiquement à mes yeux, à Fred C pour sa pédagogie et son soutien (et à Guylaine pour ses innombrables tests PCR), à Isabelle, Florence, Elisabeth, Xavier, Mathieu, Ikram, Fabrice, Agnès, Cédric, Christophe B, Fred K, Sandrine et Christophe M. De même, merci à Patricia, Dominique, Manu, Marlène, Anne-Sophie, Nao, Clément, Corentin, Hammou, Sadia, Vincent, Jodie, Ferdinand, Louis-David, Zoé, Aurore, Baptiste, Mélissandre, Dimitri, Louis, Awatef, Sylvain, Yannick, Quentin, Thomas et aux anciens Charles, Ninon, Marine et Pierre-André pour les bons moments passés au et hors du laboratoire.

During my stay at **FBN lab** in Germany, I met also fascinating scientists and people. Firstly, I would like to adress special thanks to Sebastian and Marzia for their welcome and what they do to make a pleasant stay for me. I appreciated work with the team, Gesine, Kristina, Lisa, Franzi, Anna. You can rename your research team as « French Lab » for sure now ! Thanks to Alex for letting me do some works in your fishgenetic lab, I really appreciate time and discussions with you. Thanks to all people which have done sport meeting event, play football with me, shared good moments and made my stay enjoyable in Rostock and Dummerstorf. Finally, a special thanks to Marzia and Manuel for their welcome and sharing a good moment with me. I wish you all the best and I'm sure you are and will be great parents for Mia.

A titre **plus personnel**, j'aimerai prendre le temps de remercier les personnes impliquées dans mon équilibre entre vie professionnelle et personnelle. Merci d'abord à toutes celles et ceux investis dans les structures sportives de la MEL, les bénévoles et organisateurs des nombreuses courses à pied de la région auxquelles j'ai participé et qui sont essentielles à mon bien-être. Merci au personnel sportif universitaire (Fred, Nicolas, Maxime, Pascal) pour les séances passées en leur compagnie à se dépasser physiquement et mentalement.

Evidemment, j'aimerai remercier mes parents, ma sœur et mon frère pour m'avoir soutenu, aidé, conseillé ou simplement écouté sans rien attendre en retour. Chacun à votre manière, vous m'avez beaucoup donné, avez contribué à mon épanouissement et je mesure pleinement la chance d'avoir grandi avec vous : merci d'être la famille aimante que vous êtes.

Merci à mes proches amis hors labo, Anthoine, Amélie, V-E pour les super moments passés ensemble et qui m'ont permis de m'évader dans les bonnes et les moins bonnes périodes.

Enfin, j'aimerai adresser mes plus tendres remerciements à ma meilleure amie, ma compagne, mon soutien permanent, Audrey, celle avec qui je partage ma vie depuis 9 ans. Je resterai concis en disant simplement qu'on a vécu de sacrées épreuves tous les deux et j'espère qu'on continuera encore longtemps à vivre ces intenses moments ensemble. Merci tellement pour tout.

Je terminerai ces remerciements par une citation inspirante du regretté Pr. Axel Kahn, que j'ai eu le privilège d'écouter lors d'un séminaire fascinant à propos de « Science et éthique, penser l'humain » ayant eu lieu à Lille fin 2019.

« L'éthique, c'est de s'interroger sur la vie bonne et les valeurs qui la fondent [...] (telle)

une spiritualité agnostique nourrie et tournée vers l'Homme. »

Axel Kahn (1944-2021), L'éthique dans tous ces états, 2019.

LIST OF FIGURES, TABLES AND APPENDICES

I. FIGURES

FIGURE 1: VARIOUS REPRESENTATIONS OF GLUCOSE, THE MOST ABUNDANT MONOSACCHARIDE ON EARTH.....	1
FIGURE 2: SIMPLIFIED REPRESENTATION OF THE GLYCOSYLNUCLEOTIDES AND THE THREE TYPES OF GLYCOSYLATION DEVELOPPED BELOW.	4
FIGURE 3: GANGLIOSIDE BIOSYNTHESIS WITH THE SPECIFIC GLYCOSYLTRANSFERASES INVOLVED IN EACH ENZYMATIC REACTIONS.	9
FIGURE 4: THREE MAJOR FORMS OF SIALIC ACIDS IN VERTEBRATES.....	11
FIGURE 5: COMPLEXITY AND DIVERSITY OF SIALYLATION ON GLYCOCONJUGATES.	12
FIGURE 6: SIALIC ACID METABOLISM PATHWAY IN EUKARYOTIC CELLS.....	15
FIGURE 7: DIVERSITY IN OLIGO-/POLYSIA CHAINS AT DIFFERENT LEVELS.	23
FIGURE 8: REPULSIVE AND ATTRACTIVE PROPERTIES OF POLYSIA.....	28
FIGURE 9: PHYLOGENETIC TREE HIGHLIGHTING THE DIFFERENT TELEOST FISH ORDERS FROM VERTEBRATES EVOLUTION.	31
FIGURE 10: POLYSIA NANOPARTICLES ACCUMULATED HAVE A PROTECTIVE ROLE AGAINST HISTONE CYTOTOXICITY RELEASED DURING NETOSIS.	34
FIGURE 11: THE TWO LEVELS OF AMINO ACIDS CONSERVATION IN STS.	42
FIGURE 12: SIALYLATION REACTION MECHANISM.	43
FIGURE 13: REPRESENTATION OF SIALIC ACID TRANSFER CATALYSIS.....	45
FIGURE 14: GT FOLDS WITH (A) GT-A FOLD FOR CST-II ENZYME AND (B) GT-A-LIKE FOLD FOR PORCINE ST3GAL I.	51
FIGURE 15: 3D STRUCTURE OF THE CATALYTIC DOMAIN OF PORCINE ST3GAL I.....	52
FIGURE 16: OVERALL STRUCTURE AND DIMERIZATION OF HUMAN ST8SIA III.	54
FIGURE 17: ILLUSTRATION OF THE EVOLUTIONARY SCENARIO OF GT-29 SIALYLTRANSFERASES IN METAZOA.....	58
FIGURE 18: EVOLUTIONARY FATE OF ST GENE DUPLICATED AFTER WGD EVENTS.	59
FIGURE 19: CRYSTALLOGRAPHIC DATA OF TYR122 AND HIS122 RESIDUES FROM THE HUMAN ST6GAL I AND ST6GAL II IN INTERACTION WITH GAL AND GALNAC RESIDUES RESPECTIVELY.	60
FIGURE 20: DISTRIBUTION OF ST8SIA GENES DURING VERTEBRATES EVOLUTION.	62
FIGURE 21: DISTRIBUTION OF POLYST GENES IN VERTEBRATES AFTER WGD EVENTS.	63
FIGURE 22: MSA OF HUMAN POLYST ST8SIA II AND ST8SIA IV USING UNIPROT CLUSTALW.	67
FIGURE 23: CONCEPT OF POLYSIALYLATION: FROM PROTEIN-PROTEIN INTERACTION SWITCHING TO PROTEIN-GLYCAN INTERACTION.	70
FIGURE 24: PSTD-ELECTROSTATIC POTENTIAL SURFACES REPRESENTATIONS OF HUMAN AND COREGONE ST8SIA II/II-R AND ST8SIA IV.....	74

FIGURE 25: HOMOTRIMERIC STRUCTURE OF ENDO NF	87
FIGURE 26: SUGGESTED PROCESSIVITY MECHANISM OF ENDO NF BOUND WITH OLIGO/POLYSIA.	88
FIGURE 27: CRYSTAL STRUCTURE OF scFv735 IN COMPLEX WITH OLIGO SIA	91
FIGURE 28: MPSA ENZYMATIC ASSAY WITH DIFFERENT STEPS DESCRIBED TO DETECT SIALYLATION PROPORTIONAL TO COLORIMETRIC INTENSITY DETECTED AT 620 NM.	94
FIGURE 29: PROCESS OF DMB LABELLING OF SIA PRIOR ANALYTICAL APPROACHES.	96
FIGURE 30: SELECTIVE EXOENZYMATIC LABELING (SEEL) STRATEGY ADAPTED FOR MODIFICATION AND IMAGING OF BACTERIAL LIPOOLIGOSACCHARIDES.	98
FIGURE 31: STRUCTURE AND MODE OF ACTION OF METABOLIC SIALYLATION INHIBITORS AND METABOLIC LABELING REAGENTS.	100
FIGURE 32: MULTIPLE SEQUENCE ALIGNMENT SHOWING THE CONSERVED DOMAINS OF THE SALMONID AND HUMAN POLYSTs PROTEINS.	106
FIGURE 33: OPTIMIZED CONDITIONS FOR THE PRODUCTION OF <i>CMA</i> ST8 SIA IV.	108
FIGURE 34: QUALITY OF PRODUCTION FOR <i>CMA</i> Δ 28ST8 SIA IV IN TRANSFECTED HEK293 CELLS.	109
FIGURE 35: PRODUCTION OF THE RECOMBINANT <i>CMA</i> ST8 SIA II-R1 AND ST8 SIA II-R2 IN HEK293 CELLS.	111
FIGURE 36: ^{31}P NMR SPECTRA OF CHEMO-ENZYMATICALLY SYNTHESIZED ACTIVATED CMP-SIALIC ACIDS.	113
FIGURE 37: ANALYSIS OF CMP-SIALINAL STABILITY BY ^{31}P NMR AFTER 3 MONTHS OF STORAGE AT -80°C	114
FIGURE 38: LECTIN ASSAY TO DETERMINE ALCAM GLYCOSYLATION PROFILES ON MICROPLATES.	115
FIGURE 39: USE OF MPSA TO STUDY HUMAN POLYSTs ACTIVITIES.	117
FIGURE 40: USE OF MPSA TO STUDY <i>CMA</i> ST8 SIA IV ENZYMATIC ACTIVITY.	117
FIGURE 41: USE OF CHARACTERIZATION OF <i>CMA</i> ST8 SIA II-R1 AND <i>CMA</i> ST8 SIA II-R2 ON CD166/ALCAM USING THE MPSA.	118
FIGURE 42: OPTIMIZATION OF CONDITIONS FOR ST8 SIA IV ACTIVITIES IN MPSA.	119
FIGURE 43: DETERMINATION OF KINETIC PARAMETERS OF ST8 SIA IV TOWARDS DONOR AND ACCEPTOR SUBSTRATES.	120
FIGURE 44: FUNCTIONAL CHARACTERIZATION OF <i>HSA</i> AND <i>CMA</i> POLYSTs <i>CMA</i> ON VARIOUS ACCEPTOR SUBSTRATES USING THE MPSA.	121
FIGURE 45: POLYSIA STABILITY IS TEMPERATURE-DEPENDENT.	122
FIGURE 46: INITIAL POLYSIALYLATION STATUS OF POLYSTs AND ALCAM PRIOR SIALYLATION REACTION.	123
FIGURE 47: TIME COURSE POLYSIALYLATION WITH HUMAN AND FISH ST8 SIA IV SHOWING AN INCREASING POLYSIALYLATION STATUS OF ALCAM.	124
FIGURE 48: CRYSTAL STRUCTURE OF scFv735 (PDB: 3WBD) IN COMPLEX WITH α 2,8-LINKED OCTASIALIC ACIDS.	125
FIGURE 49: PROTEIN SEQUENCE OF MAb 735 AND LOCALIZATION OF HYDROGEN BOUND INTERACTIONS.	126
FIGURE 50: ALCAM POLYSIALYLATION ACHIEVED WITH HUMAN AND <i>CMA</i> POLYSTs AND DIFFERENT CMP-SIA DONORS VISUALIZED ON WB.	127

FIGURE 51: RELATIVE KINETIC PARAMETERS OF <i>CMA</i> ST8Sia IV TOWARDS CMP-NEU5Ac, CMP-NEU5Gc AND CMP-KDN DETERMINED FROM POLYSIALYLATION ON ALCAM.	128
FIGURE 52: INFLUENCE OF TIME AND TEMPERATURE ON ALCAM POLYSIALYLATION SYNTHETIZED BY <i>CMA</i> ST8Sia IV USING CMP-NEU5Gc AND CMP-KDN.	129
FIGURE 53: IMMUNOPRECIPITATION OF <i>CMA</i> ST8Sia IV ON ANTI-3xFLAG BEADS AND ITS USE ON MPSA.	131
FIGURE 54: PREPARATION OF POLYSIA ON BEADS AND N-GLYCANS RELEASE FOR ANALYSIS.	132
FIGURE 55: APTS DERIVATIZATION OF SIA MONOSACCHARIDES AND ANALYSES IN FACE.	132
FIGURE 56: OPTIMIZATION OF COLOMINIC ACID DETECTION ON FACE AFTER APTS DERIVATIZATION.	133
FIGURE 57: APTS-FACE ANALYSIS OF ALCAM N-GLYCANS AFTER POLYSIALYLATION WITH EACH ST8Sia IV.	134
FIGURE 58: DMB-POLYMERS OF COLOMINIC ACID ANALYZED BY ANION-EXCHANGE CHROMATOGRAPHY.	135
FIGURE 59: POLYSIA RESULTING FROM 10X SIALYLATION REACTION ON ALCAM PURIFIED ON IENDON-COLUMN. ...	136
FIGURE 60: CHAIN LENGTH ANALYSIS OF CD166/ALCAM AFTER POLYSIALYLATION BY <i>CMA</i> ST8Sia IV.	137
FIGURE 61: COMPOSITION ANALYSIS OF POLYSIA ON CD166/ALCAM.	138
FIGURE 62: POLYSIALYLATION OF HEK293 CELLS.	140
FIGURE 63: EXOSIALYLATION OF HEK293 CELLS WITH <i>CMA</i> ST8Sia IV.	141
FIGURE 64: EXOSIALYLATION OF HEK293 CELLS WITH <i>Hsa</i> ST8Sia IV.	142
FIGURE 65: SUMMARY OF EXOSIALYLATION ON HEK293 CELLS USING ST8Sia IV.	143
FIGURE 66: PRODUCTION AND SECRETION OF <i>CMA</i> ST8Sia IV IN DIFFERENT CELL LINES.	145
FIGURE 67: AUTOPOLYSIALYLATION STATUS OF EACH <i>CMA</i> ST8Sia IV PRODUCTION.	146
FIGURE 68: ENZYMATIC ACTIVITY OF EACH <i>CMA</i> ST8Sia IV PRODUCTION USING CMP-SiaNAL IN MPSA.	147
FIGURE 69: ENDSRIPT ANALYSES FOR STRUCTURAL SEQUENCE ORGANIZATION OF THE 5 POLYSTs.	150
FIGURE 70: 3D STRUCTURAL MODELING OF <i>CMA</i> ST8Sia IV.	152
FIGURE 71: POTENTIAL POLYSTs COOPERATION AFTER CO-INCUBATION IN MPSA.	153
FIGURE 72: SURFACE REPRESENTATION FOR HOMODIMERIC INTERACTION OF <i>CMA</i> ST8Sia IV.	154
FIGURE 73: ALPHAFOLD-MULTIMER MODELING FOR HOMO- AND HETERODIMERIZATION BETWEEN HUMAN POLYSTs.	155
FIGURE 74: ALPHAFOLD-MULTIMER MODELING FOR HETERODIMERS BETWEEN <i>CMA</i> POLYSTs.	156
FIGURE 75: CONTRIBUTION OF THIS WORK FOR BIOTECHNOLOGICAL APPLICATIONS.	168
FIGURE 76: MODEL THAT RESUME MAIN PROJECT EXPERIMENTS BASED ON SALMONID POLYSTs ENZYMATIC SPECIFICITIES TO EXPLAIN THE POLYSIA BIODIVERSITY OCCURRING IN SALMONID.	170

II. TABLES

TABLE 1: RETICULAR INITIATION FOR THE N-GLYCANS BIOSYNTHESIS.	5
TABLE 2: GOLGI MATURATION OF N-GLYCANS.	7

TABLE 3: BIOSYNTHESIS OF MUCIN TYPE O-GLYCANS IN THE GOLGI.....	8
TABLE 4: REPRESENTATION OF DIFFERENT SIALYLATED EPITOPES STRUCTURES REPORTED AS BIOMARKERS IN PATHOLOGIES AND DISEASES.	20
TABLE 5: LIST OF GLYCOPROTEINS DESCRIBED TO BE POLYSIALYLATED IN THE NATURE. ADAPTED FROM TEINTENIER- LELEVRE THESIS, 2006 ; SATO, 2013 ; BOYANCE THESIS, 2021.	24
TABLE 6: THE 20 HUMAN SIALYLTRANSFERASE GENES.....	37
TABLE 7: HUMAN PROTEIN SIALYLTRANSFERASES AND THEIR ENZYMATIC SPECIFICITIES.....	47
TABLE 8: CRYSTAL STRUCTURES OF SIALYLTRANSFERASES IN THE PROTEIN DATA BANK (PDB).	56
TABLE 9: IMPACT OF REPLACING ST8Sia IV AND ST8Sia II PBR RESIDUES ON SUBSTRATE POLYSIALYLATION.	68
TABLE 10: IMPACT OF N-GLYCOSYLATION ON ENZYMATIC ACTIVITY OF ST8Sia II AND ST8Sia IV.	71
TABLE 11: ADVANTAGES AND DISADVANTAGES OF CELL SYSTEM FOR GLYCOPROTEIN PRODUCTION.	76
TABLE 12: GLYCOMIC PROFILES AND SIALYLATED GLYCOCONJUGATES IN (A) HEK293 AND (B) CHO-K1 CELLS.	77
TABLE 13: DESCRIPTION OF SIALYLATED GLYCANIC STRUCTURES OF ACCEPTOR SUBSTRATES USED TO STUDY SIALYLTRANSFERASE ACTIVITY.	79
TABLE 14: LECTIN DESCRIPTION.....	85
TABLE 15: SIA EPITOPE ANTIBODIES AND THEIR SPECIFICITY.	90
TABLE 16: STABILITY OF THE RECOMBINANT <i>C.ma</i> POLYSTs SEQUENCES DESIGNED.	107
TABLE 17: SUMMARY TABLE OF THE PRODUCTION AND USE OF DIFFERENTS TRUNCATED STs DURING MY THESIS PROJECT.	112
TABLE 18: DMB-HPLC ANALYSIS AND SIA QUANTIFICATION FROM POLYSIALYLATION ON DNase I WITH VARIOUS SIA AND BOTH ST8Sia IV ENZYMES DETECTED ON WB.....	135
TABLE 19: SUMMARY OF ESPRIPT ANALYSES BETWEEN HUMAN AND COREGONE POLYSTs SEQUENCES.....	150
TABLE 20: IPTM+PTM SCORE PREDICTION FOR HOMODIMERIZATION OF ST8Sia IV SEQUENCES.....	154

III. APPENDICES

APPENDIX 1: BIOINFORMATIC TOOLS AND DATABASES TO STUDY GLYCANS AND GLYCOSYLTRANSFERASES FOR VERTEBRATES.	213
APPENDIX 2: PLASMID CONSTRUCTS USED FOR MOLECULAR CLONING OF STs.	219
APPENDIX 3: TABLE OF SENSE AND ANTISENSE OLIGONUCLEOTIDES USED TO GENERATE CDNA ENCODING TRUNCATED FORM OF THE <i>C.ma</i> ST8Sia2 R-1, ST8Sia2 R-2 AND ST8Sia4 LACKING THEIR FIRST 28, 35 OR 50 AA RESIDUES.	220
APPENDIX 4: FULL NUCLEOTIDE AND AA SEQUENCES OF SIALYLTRANSFERASES USED AND ACCEPTORS PRODUCED.....	221

LIST OF ABBREVIATIONS

Δ : deletion	CSP : Chemical Shift Perturbation
3D : Three dimensional	CSS : CMP-sialic acid synthetase
Å : Angstrom	CuAAC : Copper-catalyzed alkyne-azide cycloaddition
aa : Amino acid	Cys : Cysteine
ALCAM : Activated Leukocyte Cell Adhesion Molecule	Dol-P : Dolichol-Phosphate
AMP : Antimicrobial Peptides	<i>D.re</i> : <i>Danio rerio</i> , zebrafish
Ala : Alanine	DAPI : 4'-6'-diamino-2-phenylindole
APTS : 1-aminopyrene-3,6,8-trisulfonic acid	diSia : Disialyl motif
Asn : Asparagine	DMB : 1,2-diamino-4,5-methylene dioxybenzene dihydrochloride
Asp : Aspartic acid	DMEM : Dulbecco Modified Eagle Medium
ATCC : American Type Culture Collection	DNA : Deoxyribonucleotid acid
BDNF : Brain-Derived Neurotrophic Factor	Dol-P : Dolichol-Phosphate
BSA : Bovine Serum Albumine	DP : Degree of Polymerization
BSM : Bovine Submaxillary Mucin	dpm : Disintegrations per minute
BLAST : Basic local alignment search tool	ECL : Electrochemiluminescence
BTTAA : Bis-terbutyl-triazol-acetic acid	EDTA : Ethylenediamine-tetraacetic acid
CANX : Calnexin	endoN : Endoneuraminidase-N
CALR : Calreticulin	EPO : Erythropoietin
CAZy : Carbohydrate-Active enZymes	ER : Endoplasmic Reticulum
CBP : Carbohydrate-Binding Proteins	ESL-1 : E-selectin ligand-1
CDG : Congenital Disorder of Glycosylation	FACE : Fluorescent-Assisted Carbohydrate Electrophoresis
CDR : Complementary-Determining Regions	FACS : Fluorescence Activated Cell Sorting
CEA : Carcinoembryonic antigen	FAEC : Fluorescent Anion-Exchange Chromatography
Cer : Ceramide	FBS : Foetal bovine serum
CGE : Chemoenzymatic GlycoEngineering	FFP : Farnesyl-PyroPhosphate
CHO : Chinese Hamster Ovary cells	FGF2 : Fibroblast Growth Factor 2
Chr : Chromosome	Fig. : Figure
CMAH : CMP-Neu5Ac hydrolase	FITC : Fluorescein Isothiocyanate
<i>C.ma</i> : <i>Coregonus maraena</i> – salmonid	FL : Full Length
CMAS : CMP-Sia synthetase	Fuc : Fucose
CMP : Cytidine monophosphate	Gal : Galactose
CMP-Sia : CMP-Sialic acid	GalNAc : <i>N</i> -Acetylgalactosamine
CNS : Central Neuronal System	G-CSF : Granulocyte Colony-Stimulating Factor
COG : Conserv Oligomeric Golgi complex	GDP : Guanosine DiPhosphate
C^{ter} , CTD : C-terminal domain	
CTP : Cytidine triphosphate	

Glc: Glucose
Glu: Glutamic acid
Gly: Glycine
GNE: UDP-GlcNAc-2 epimerase/ManNAc kinase
GSL: Glycosphingolipid
GT: Glycosyltransferase
HCl: Chloride acid
HEK: Human Embryonic Kidney cells
HeLa: Human Cervical Cancer cells
HGNC: Hugo Gene Community Nomenclature
His: Histidine
HMO: Human Milk Oligosaccharides
HPLC: High Performance Liquid Chromatography
HRP: Horseradish Peroxidase
H.sa: *Homo sapiens*, human
HUGO: Human Gene Organization
iEndoN: inactive endoN
Ig: Immunoglobulin
II : Instability index
kDa: Kilodalton(s)
Kdn: 2-keto-3-déoxynonulosonic acid
K_m: Michaelis constant
KO: Knockout
LacdiNAc: GalNAc β 1-4GlcNAc
LacNAc: Gal β 1-4GlcNAc
LB: Luria-Bertani medium
LECA: Last Eukaryota Common Ancestor
Leu: Leucine
LLO: Lipid Linked-Oligosaccharides
MAH: Lectin *Maackia amurensis*
Man: Mannose
ManLev: *N*-levulinoylmannosamine
ManNAc: *N*-acetyl-*D*-mannosamine
ManNAI: *N*-(4-pentynoyl)-*D*-mannosamine
MCS: Multiple Cloning Site
MD: Molecular Dynamics
MFI: Mean Fluorescence Intensity
MGE: Metabolic Glycan Engineering
MOE: Metabolic Oligosaccharide Engineering
MPSA: MultiPlate Sialyltransferase Assay
MS: Mass Spectrometry
MSA: Multiple Sequences Alignment
MW: Molecular Weight
MYA: Millions Years Ago
NANS: Neu5Ac 9-phosphate synthase
NCAM: Neural Cell Adhesion Molecule
NCBI: National Center for Biotechnology Information
NETs: Neutrophil Extracellular Traps
Neu: Neuraminic acid
NEU: Neuraminidases
Neu5Ac: *N*-acetylneuraminic acid
Neu5Gc: *N*-glycolylneuraminic acid
NMR: Nuclear Magnetic Resonance
NRP-2: Neuropilin-2
NSCLC: Non-Small Cell Lung Carcinoma
N^{ter}: N-terminal side
OligoSia: oligo-sialic acid
OST: Oligosaccharyltransferase
PAE: Predicted Aligned Error
PAGE: Polyacrylamide gel electrophoresis
PBR: Poly Basic Region
PBS: Phosphate Buffer Saline
PDB: Protein Data Bank
PEG: PolyEthylene Glycol
PEP: Phosphoenolpyruvate
PFA: Paraformaldehyde
Phe: Phenylalanine
PL: Phospholipids
pLDDT: predicted Local Distance Difference Test
PolySia: Poly Sialic Acid
PolyST: Polysialyltransferase
PSGP: Polysialoglycoprotein
PFA: Paraformaldehyde
Pi: inorganic phosphate

Ppi: inorganic pyrophosphate
PSGL-1: P-Selectin glycoprotein ligand 1
PSGP: Salmonid polysialoglycoproteins
PSTD: PolySialylTransferase Domain
PTA: Phosphotungstic acid
pTM: predicted Template Modelling score
PTM: Post-translational modifications
RMSD: Root Mean Square Deviation
RNA: Ribonucleic acid
RPLC: Reverse Phase Liquid Chromatography
RPM: Round Per Minute
rtCSS: rainbow trout CSS
SaGD: Salmonid Genome Duplication
SAM: Sialic acid mimetic
SASD: Sialic Acid Storage Disorders
SCLC: Small Cell Lung Carcinoma
SDP: Specificity Determining Position
SDS: Sodium Dodecylsulfate
SEEL: Selective Exo-Enzymatic Labeling
Ser: Serine
Sia: Sialic acid
Siglec: Sia-binding Ig-type lectin
SiaLev: N-leulinoyl Sia
SiaMix: mixture of Neu5Ac, Neu5Gc and Kdn
SiaNAI: N-(4-pentynoyl)-D-neuraminic acid
sLe^a: Sialyl Lewis A
sLe^x: Sialyl Lewis X

SM: SialylMotif
SNA: Lectin *Sambucus nigra*
SNFG: Symbol Nomenclature For Glycans
SNP: Single Nucleotide Polymorphism
SPR: Surface Plasmon Resonance
ST: Sialyltransferase
SynCAM-1: Synaptic Cell Adhesion Molecule-1
TACA: Tumor Associated Carbohydrate Antigen
TBST: Tris Buffer Salin Tween
TFA: TriFluoroAcetic acid
TGD: Teleost Genome Duplication
TGN: Trans-Golgi Network
THF: TetraHydroFuran
Thr: Threonine
TMD, TM: Transmembrane Domain
TMB: Tetramethylbenzidine
Trp: Tryptophane
Tyr: Tyrosine
UDP: Uridine DiPhosphate
Val: Valine
V_H: Heavy chain
V_m: Maximum Velocity
V_L: Light chain
WB: Western-Blot
WGDR: Whole Genome Duplication Round
WFA: Lectin *Wisteria floribunda*
WT: Wild Type

SUMMARY

In contrast to humans that generate only a single type of polymer consisting of α 2,8-linked N-acetylneuraminic acid polyNeu5Ac, fishes are able to generate a wide range of different polysialic acids varying according to teleosts order (Salmoniformes, Cypriniformes, ...). While two human poly- α -2,8-sialyltransferases ST8Sia II and ST8Sia IV are responsible for polyNeu5Ac synthesis onto a limited number of glycoproteins, the fish biosynthetic machinery leading to the polySia biodiversity is currently not known. Indeed, phylogenetic studies of the α 2,8-sialyltransferases ST8Sia revealed a particular distribution of fish polysialyltransferases compared to their human homologs. Three polysialyltransferase genes *st8sia4*, *st8sia2-r1* and *st8sia2-r2* were identified in the salmonids genome that could be involved in the biosynthesis of these polySia.

During my PhD, I engineered soluble and tagged polysialyltransferases from the salmonid *Coregonus maraena* by molecular cloning and produced these recombinant enzymes in the cell culture medium of transiently transfected mammalian cells. In parallel, I have chemoenzymatically synthesized CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn activating the three main sialic acids reported in vertebrates (*i.e.* Neu5Ac, Neu5Gc and Kdn) to use them as natural donor substrates in sialylation reactions. I have also prepared various acceptor substrates as well as a functionalized sialic acid donor with a small alkyne group (*i.e.* CMP-SiaNA1) which allow bioorthogonal click-chemistry approaches. Using these specific biosynthetic tools and the recently developed MicroPlate Sialyltransferase Assay (MPSA), I optimized polysialylation assay conditions (temperature, pH, time-course) and compared enzymatic activities of the fish and human enzymes. Notably, I achieved the biochemical characterization of the fish ST8Sia IV on MPSA, demonstrated its preferential activity onto known acceptor substrates for the human polysialyltransferases like NRP-2 and DNase I and on other acceptors like PSGP, orosomucoid, BSM and ALCAM. Notably using ALCAM on MPSA, I determined kinetic parameters for the human and coregone ST8Sia IV showing differences of affinities towards CMP-SiaNA1.

Then, I carried out polysialylation reactions with CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn in cell-free and cell-based assays and structural analyses of polysialylated products using the anti-polySia monoclonal antibody 735 (mAb735) and endoneuraminidase-N (endoN). I could demonstrate unique substrate specificities of the fish enzyme compared to its human homologue. I showed that *C. maraena* ST8Sia IV is of particular interest to modify glycoproteins with notable differences in substrates recognition to generate *in vitro* and cell surface polymers made up of Neu5Ac, Neu5Gc and Kdn.

In an effort to understand how this polysialylation pathway is regulated in vertebrates, influence of autopolysialylation was studied. As reported for mammalian polysialyltransferases, I have also shown that autopolysialylation is possible for the fish ST8Sia IV but is not essential for its catalytic activity.

Finally, modeling of the 3D structure of polysialyltransferases using AlphaFold was initiated and has revealed disordered loops and specific amino-acid changes in conserved sialylmotifs, PBR and PSTD motifs. These structural differences between human and salmonid polysialyltransferases might be involved in potential dimeric interactions, flexibility of the catalytic domain and could explain the enzymatic specificities of salmonid polysialyltransferases.

These studies shed light on structure-function of vertebrate polyST and paved the way for the synthesis of innovative polySia-based materials which will be combined to antimicrobial biomolecules for the development of novel bioactive surfaces in the future.

Keywords : Polysialylation, polySia, sialyltransferases, ST8Sia, CAZy GT-29, recombinant proteins, polysialylated glycoproteins, enzymatic specificities, vertebrate evolution, click-chemistry, mAb735, endoN, glycobiology.

RÉSUMÉ

Contrairement aux humains qui ne génèrent qu'un seul type de polymère constitué d'acides N-acétylneuraminiques liés en α 2,8, les poissons sont capables de générer une large gamme d'acides polysialiques variant selon l'ordre des téléostéens (Salmoniformes, Cypriniformes, ...). Alors que deux poly- α -2,8-sialyltransférases humaines ST8Sia II et ST8Sia IV sont responsables de la synthèse de polyNeu5Ac sur un nombre limité de glycoprotéines, la machinerie de biosynthèse des poissons conduisant à la biodiversité de polySia n'est actuellement pas connue. En effet, des études phylogénétiques des ST8Sia ont révélé une distribution particulière des gènes de polyST de poisson par rapport à ceux humains. Trois gènes de polyST *st8sia4*, *st8sia2-r1* et *st8sia2-r2* ont été identifiés dans le génome des salmonidés et pourraient être impliqués dans la biosynthèse de ces polySia.

Au cours de mon doctorat, j'ai conçu des polyST solubles et étiquetées du salmonidé *Coregonus maraena* par clonage moléculaire et j'ai produit ces enzymes recombinantes dans le milieu de culture de cellules de mammifères transitoirement transfectées. En parallèle, j'ai synthétisé chimio-enzymatiquement du CMP-Neu5Ac, CMP-Neu5Gc et CMP-Kdn, à partir des trois principaux acides sialiques rapportés chez les vertébrés (*i.e.* Neu5Ac, Neu5Gc et Kdn), pour les utiliser comme substrats donneurs naturels dans les réactions de sialylation. J'ai également préparé divers substrats accepteurs ainsi qu'un donneur d'acide sialique fonctionnalisé avec un alcyne (*i.e.* CMP-SiaNA1) pour des approches bioorthogonales de chimie-clic. En utilisant ces outils de biosynthèse spécifiques et le test MPSA récemment développé, j'ai optimisé les conditions des essais de polysialylation (température, pH, durée) et comparé les activités des enzymes de poisson et humaines. J'ai notamment réalisé la caractérisation biochimique de ST8Sia IV de poisson en MPSA, démontré son activité préférentielle sur des substrats accepteurs connus pour les polySTs humaines (NRP-2 et DNase I) et sur d'autres accepteurs (PSGP, orosomucoïde, BSM et ALCAM). Notamment en utilisant l'ALCAM en MPSA, j'ai déterminé les paramètres cinétiques des ST8Sia IV humaine et de corégone et montré leurs différences d'affinités envers le CMP-SiaNA1.

Ensuite, j'ai effectué des réactions de polysialylation avec du CMP-Neu5Ac, CMP-Neu5Gc et CMP-Kdn dans des essais *in vitro* et cellulaires suivis d'analyses structurales de produits polysialylés en utilisant l'anticorps monoclonal anti-polySia 735 et l'endoneuraminidase-N. J'ai pu démontrer des spécificités de substrat uniques de l'enzyme de poisson par rapport à son homologue humaine. J'ai montré que *C.maraena* ST8Sia IV présente un intérêt particulier pour modifier les glycoprotéines avec des différences notables dans la reconnaissance des substrats pour générer des polymères *in vitro* et de surface cellulaire composés de Neu5Ac, Neu5Gc et Kdn.

Afin de comprendre comment cette voie de polysialylation est régulée chez les vertébrés, l'influence de l'autopolysialylation a été étudiée. Comme rapporté pour les polyST de mammifères, j'ai également montré que l'autopolysialylation est possible pour ST8Sia IV de poisson mais n'est pas essentielle à son activité catalytique.

Enfin, la modélisation de structure 3D de polyST a été initiée à l'aide d'AlphaFold et a révélé des boucles désordonnées et changements d'acides aminés dans les sialylmotifs conservés et les motifs PBR et PSTD. Ces différences de structure entre polyST humaines et de poisson pourraient être impliquées dans des interactions dimériques potentielles, une flexibilité du domaine catalytique et pourraient expliquer les spécificités enzymatiques des polyST de poisson. Ces études ont mis en lumière la structure-fonction des polyST de vertébrés et ouvrent la voie à la synthèse de matériaux innovants polysialylés qui seront combinés à des biomolécules antimicrobiennes pour le développement de nouvelles surfaces bioactives à l'avenir.

Mots-clés : Polysialylation, polySia, sialyltransférases, ST8Sia, CAZy GT-29, protéines recombinantes, glycoprotéines polysialylées, spécificités enzymatiques, évolution des vertébrés, chimie-click, mAb735, endoN, glycobiologie.

PREAMBLE

All life on Earth is composed by four types of macromolecules which are nucleic acids, lipids, proteins and carbohydrates. Also known as sugars or glycans, carbohydrates are the most abundant biological molecules on Earth, the main components of our diet and energy sources, and are essential building blocks present in every single cell in prokaryote and eukaryote organisms. Glycoscience, the science of carbohydrates, was considered as a lagging science until the 2000's due to the extraordinary complexity and diversity of existing glycans and a lack of tools and informations dedicated to their studies.

Although still difficult to study, glycoscience is now recognised worldwide as an emerging area since these two last decades with many governments, scientific societies and researchers which contribute to the spread and increase of glycoscience knowledge worldwide. Thanks to this scientific breakthrough, more and more scientific literature strengthen the idea that glycans are not only « sugars » as sucrose but could be considered as very complex macromolecules with a great structural diversity due to their conformation, anomery, linkage, size and chain lengths, physico-chemical properties, heterogeneity and their association to proteins, lipids or nucleic acids.

Glycans of glycoconjugates are key elements involved in many physio-pathological processes as modulation of cell-cell communication, helping sperm to recognize eggs during fertilization and host-pathogens interactions and for other striking examples such as ABO blood groups and glycan-based biomarkers, heparin polysaccharide in medicine, considered as a new hallmark of cancer when disregulated, in genetic diseases such as Congenital Disorders of Glycosylation (CDGs) or pathologies like in diabetes. Even R&D glycobiotechnology companies increased their glycomics markets to meet the social challenges as examples of the increasing use of cellulose in plants and hyaluronic acid, which is a glycosaminoglycan, and their bioindustry and biopharmaceutical applications in textile and in cosmetic, human milk oligosaccharides (HMO) with prebiotic properties or production of glycoengineered proteins and development of glycan-based vaccines, glycomimetics and antibodies with high specificities and enhanced properties.

These many challenges and perspectives encouraged me to achieved my thesis in health and biology, biochemistry and in glycoscience discipline. My scientific doctoral thesis work was done in our french monothematic unit focus on glycosciences, the UGSF lab (UMR 8576), to gain insights into structure/function relationships of glycans. Especially, within the “regulation of terminal glycosylation team”, our research topics are based on the fundamental mechanisms controlling terminal glycosylation in light of evolution in order to better understand its functional importance in Metazoa. We study terminal glycosyltransferases (GTs), the enzymes responsible for glycosylation, using integrative approaches with genomics and functional genomics and combining them with biochemical and cellular tools to have a better understanding of how changes in expression and function of GTs during evolution have influenced the diversification of animal species and the development of human diseases.

This thesis manuscript is divided into four chapters: a **bibliographic introduction** including references on glycosylation, sialic acids and sialylated structures, their biological roles and diversity. Then, I focused more precisely on polysialic acids (polySia), their diversity and importance in biology and I continued highlighting the enzymes responsible for sialylation : the sialyltransferases. After describing more particularly α 2,8-polysialyltransferases, I concluded this first part indicating the different tools currently used to study sialyltransferases and sialoglycoconjugates which led me to established my thesis objectives.

The second chapter is dedicated to the **results** obtained during my three years-thesis. Starting from the identification of a particular distribution of polysialyltransferase genes and polysialylation status in salmonid species in comparison to human, my thesis work was to clone, produce and characterize these sialyltransferases, in order to define their enzymatic activities and distinct specificities and to study polysialylated products formed with the use of own-produced enzymes, modified sugars donors and acceptors.

Then in the penultimate part, I **discussed** the results obtained, the contribution and perspectives of this work in the glycoscience field. The last part of this manuscript is dedicated to the **technical strategies** and **materials** used for the results obtained during my thesis.

At a personal level, this thesis gave me a lot on different aspects. I started to study terminal glycosylation out of curiosity and interest on sialyltransferases in terms of enzymatic activity and evolution since my Master. I realized the opportunity to use various approaches ranging from cell biology, molecular biology, biochemistry, chemistry and bioinformatic strategies, adapting molecular tools and organizing my ideas to answer biological and fundamental questions. I had the possibility to participate and present my research work in many international congresses, to travel and meet fascinating researchers actively involved in progress for glycosciences. I've done internship in our collaborator lab in Germany, I was involved in other scientific projects in the group as described in my results part and I've done teaching and supervision of students with passion and enthusiasm to conclude finally that this experience was a great human adventure with a social connection and scientific emulation.

TABLE OF CONTENTS

BIBLIOGRAPHIC INTRODUCTION	1
I. GLYCOSYLATION AND GLYCANS	1
A. Monosaccharides	1
B. Glycans, and glycoconjugates	2
II. BIOSYNTHESIS OF GLYCOCONJUGATES	3
A. Nucleotid sugar donors	4
B. Biosynthesis of N-glycosylproteins	5
C. Biosynthesis of mucin type O-glycans	7
D. Biosynthesis of gangliosides	9
III. SIALIC ACIDS AND SIALYLATED GLYCOCONJUGATES	10
A. Sialic acids discovery and structural diversity in Vertebrates	10
B. Sialic acids biosynthesis pathway	13
C. Sialylated glycoconjugates in physiology	16
D. Sialylated glycoconjugates in pathological states	17
E. Sialylated epitopes : biomarkers and recognitions	18
IV. POLYSIALIC ACID (POLYSIA): STRUCTURE AND BIOLOGICAL ROLES	23
A. Structure of polySia	23
B. Polysialylation : acceptors and biological roles	26
<i>i. PolySia in bacteria</i>	26
<i>ii. PolySia in human and vertebrates</i>	26
<i>iii. PolySia in fishes</i>	29
C. PolySia expression dysregulated in pathology	32

D.	Biotechnological and biotherapeutical use of polySia	32
V.	SIALYLTRANSFERASES AND THEIR FUNCTIONAL BIODIVERSITY	35
A.	Gene organization and regulation of ST genes expression	36
B.	ST genes expression in physiopathology	38
<i>i.</i>	<i>In physiology</i>	38
<i>ii.</i>	<i>In cancers and diabetes</i>	38
<i>iii.</i>	<i>In genetic and mental diseases</i>	39
C.	Protein organization of STs	40
D.	Mechanism action of sialylation reaction.....	43
E.	STs enzymatic specificity	44
F.	3D structures of sialyltransferases.....	51
<i>i.</i>	<i>Bacterial STs</i>	51
<i>ii.</i>	<i>Mammalian STs</i>	51
<i>iii.</i>	<i>Homo- and heterodimerization of mammalian STs</i>	53
G.	Molecular phylogeny and evolution of STs.....	57
VI.	POLYSIALYLTRANSFERASES	61
A.	Description of ST8SIA family	61
B.	Evolution of ST8Sia family	61
<i>i.</i>	<i>All ST8Sia enzymes</i>	61
<i>ii.</i>	<i>Polysialyltransferases</i>	62
C.	Role of vertebrates polySTs.....	64
<i>i.</i>	<i>Biosynthesis of polySia</i>	64
<i>ii.</i>	<i>Expression in mammals</i>	65

iii.	<i>Expression in zebrafish</i>	65
iv.	<i>Specificity towards acceptor substrates</i>	66
D.	Sequences of mammalian polyST and conserved motifs	67
i.	<i>Cysteines involved in polyST stability</i>	67
ii.	<i>PBR motif</i>	68
iii.	<i>PSTD motif</i>	69
iv.	<i>N-glycosylation and autopolySia of polySTs</i>	70
v.	<i>ST8Sia III : polyST or oligoST ?</i>	72
E.	Teleost fish polyST	72
VII.	METHODS TO STUDY SIALYLATION AND SIALYLTRANSFERASES	75
A.	Production of glycoengineered enzymes and acceptors.....	75
i.	<i>In bacteria system</i>	75
ii.	<i>In eukaryotic system</i>	76
B.	Lectins and neuraminidases	85
C.	Antibodies	89
D.	Enzymatic assays	92
i.	<i>Use of radiolabeled CMP-Neu5Ac</i>	92
ii.	<i>Detection of sialylation side products</i>	92
iii.	<i>Use of tagged-Sia and lectin detection</i>	92
iv.	<i>Detection of sialylation by analytical chemistry</i>	93
v.	<i>Use of unnatural modified Sia</i>	93
E.	Analytical approaches	95
F.	Sialylation labelling in cellulo and on cell surface.....	97

G.	STs Inhibitors development	99
H.	Bioinformatics tools and modeling	101
VIII.	SCIENTIFIC OBJECTIVES OF THESIS	102
RESULTS		104
I.	SALMONID POLYSIALYLTRANSFERASES SUBSTRATE SPECIFICITIES: CONTRIBUTION TO DIVERSITY OF FISH POLYSIA?	104
A.	Production of the Cma polysialyltransferases	104
<i>i.</i>	<i>Analysis of the Cma polySTs sequences</i>	<i>104</i>
<i>ii.</i>	<i>Molecular cloning of polySTs in expression vectors</i>	<i>106</i>
<i>iii.</i>	<i>Cellular transfection for Cma polySTs production</i>	<i>107</i>
B.	Preparation of donor and acceptor substrates for ST activity	112
C.	Development of enzymatic assay for polySTs activity	116
D.	Biochemical characterization of the Cma ST8Sia IV enzymatic activity	118
E.	PolySia assessment generated by polySTs on WB	122
F.	Structural assessment of polySia generated by polySTs.....	130
G.	Exosialylation of cultured cells using ST8Sia IV	139
II.	AUTOPOLYSIALYLATION OF POLYSIALYLTRANSFERASES : A REQUIREMENT FOR ENZYMATIC ACTIVITY OF CMA ST8SIA IV ?.....	145
A.	Production of Cma ST8Sia IV in different cell types.....	145
B.	Autopolysialylation of Cma ST8Sia IV produced in different cell types.....	146
C.	Enzymatic activity of Cma ST8Sia IV produced in different cell types.....	147
III.	ENZYMATIC SPECIFICITIES OF SALMONID POLYSIALYLTRANSFERASES: DIFFERENCES AT A STRUCTURAL LEVEL ?	148

A.	Modeling of monomeric polySTs and Comparison of 3D structures	148
B.	Homo- and heterodimerization of polyST for enzymatic activity	152
<i>i.</i>	<i>Co-incubation for enzymatic assay in MPSA</i>	152
<i>ii.</i>	<i>Homodimerization of polySTs</i>	153
<i>iii.</i>	<i>Heterodimerization of polySTs</i>	157
DISCUSSION		158
I.	PRODUCTION OF POLYSTS	158
II.	CHEMO-ENZYMATIC SYNTHESIS OF CMP-SIA DONOR SUBSTRATES	159
III.	BIOCHEMICAL CHARACTERIZATION OF POLYSTS	159
IV.	POLYSIA SYNTHESIS AND ANALYSIS	161
<i>i.</i>	<i>Detection on WB using mAb735 and endoN</i>	161
<i>ii.</i>	<i>Analytical methods to confirm polySia composition and size</i>	163
V.	POLYSIALYLATION REMODELING ON CELL SURFACE	164
VI.	AUTOPOLYSIALYLATION OF CMA ST8SIA IV	165
VII.	MODELING AND COOPERATION OF POLYSTS	166
VIII.	MID AND LONG-TERM PERSPECTIVES	167
IX.	CONCLUSION	169
MATERIALS AND METHODS		172
I.	MATERIALS AND ANIMALS	172
II.	BIOINFORMATICS ANALYSES	173
A.	In silico identification of polysialyltransferases sequences	173
B.	Modeling using AlphaFold 2	173
C.	Molecular docking for polySia tools	175

III. MOLECULAR BIOLOGY	175
A. RNA extraction, cDNA synthesis and PCR amplification of polySTs	175
B. Production and purification of plasmids	176
IV. CELLULAR BIOLOGY	177
C. Cell culture	177
D. Exogenous sialylation on cells: IF and FACS	177
V. BIOCHEMISTRY AND ANALYTIC APPROACHES	179
A. Western Blot	179
B. Plate lectin recognition assay	180
C. Enzymatic activity tests	180
D. Analytical approaches to study polymers of sialic acids	181
<i>i. Chain length analysis by HPLC.....</i>	<i>181</i>
<i>ii. Chain length analysis by FACE.....</i>	<i>182</i>
<i>iii. Quantitative analysis of polySia composition by HPLC.....</i>	<i>183</i>
VI. CHEMICAL SYNTHESIS	184
A. Chemo-enzymatic synthesis of CMP-activated sialic acids	184
B. Preparation of the different acceptor substrates	184
REFERENCES.....	185
APPENDIX	213

BIBLIOGRAPHIC INTRODUCTION

I. GLYCOSYLATION AND GLYCANS

A. MONOSACCHARIDES

Monosaccharides are the simplest basic units assembled to form carbohydrates. These essential building blocks are present in every single cell in prokaryotes and eukaryotes and represent the most abundant class of biological molecules on Earth. With a general molecular formula of $C_n(H_2O)_n$, monosaccharides exist in a linear and cyclic forms in solution and have a D or L configuration determined by the absolute configuration of the asymmetric carbon atom in C-4 for pentoses and C-5 for hexoses, where $n = 5$ or 6 respectively (Fig. 1A). However, the most common monosaccharides found in Vertebrates, with the exception of L-Fucose, have D configuration as for D-Glucose, D-Mannose, D-Galactose and sialic acids. They adopt a pyranose or furanose cyclic form, both α and β anomers depending on C-1 epimerization (Fig. 1B).

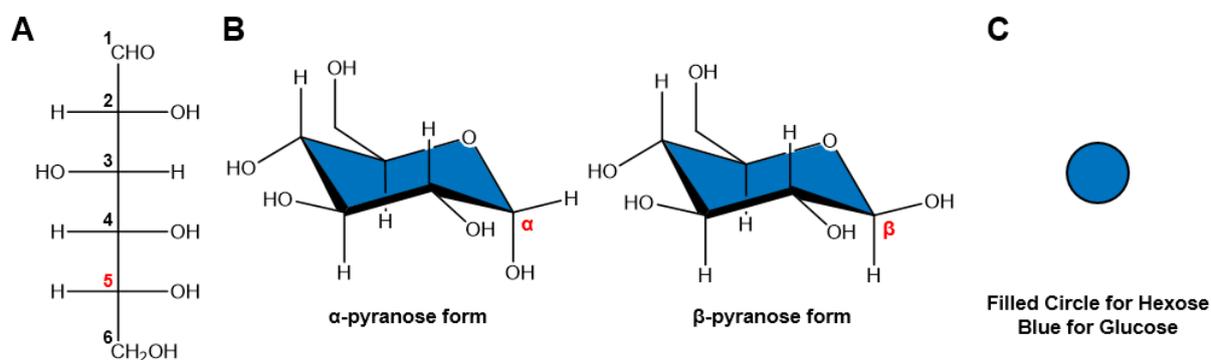


Figure 1: Various representations of Glucose, the most abundant monosaccharide on Earth. Represented according (A) Fischer projection in linear form where D configuration is established in C-5 (red), (B) Chair conformation of Haworth projection with α - and β -pyranose cyclic forms of the monosaccharide depending on the C-1 epimerization (red) in solution. Ring is fill with blue color according to (C) SNFG nomenclature and filled circle representation for hexose adopted by the glycoscientists community (Varki *et al.*, 2015a).

Following the rise of glycosciences in the beginning of this century, a Symbol Nomenclature for Glycans (SNFG) was adopted definitely adopted by the community in 2015 and regularly updated (available at: <https://www.ncbi.nlm.nih.gov/glycans/snfg.html>) to standardize and simplify the drawing glycan structures using symbol as described for the Glucose (Fig. 1C). These

symbols differ according to shapes as filled circle, square, triangle or diamond for hexose, hexosamine, deoxyhexose or 3-deoxy-nonulosonic acids respectively and according to colors as blue for Glucose, green for Mannose, yellow for Galactose, red for Fucose and/or Sialic acids family for the most common monosaccharides found in Vertebrates.

B. GLYCANS, AND GLYCOCONJUGATES

Associated to each other by glycosidic linkages, monosaccharides constitute complex macromolecules with a great structural diversity due to their conformation, anomery, linkage, size and chain lengths (di-, oligo- and poly-saccharide if their residues numbers are 2, between 2 and 20 or more than 20 respectively), physico-chemical properties, heterogeneity at different level (molecular, cellular, organisms) and either free or associated to form glycoconjugates with proteins or lipids. Indeed, glycoconjugates are composed of a glycan part, with the monosaccharides building blocks, associated either with a lipid to form glycolipids, with an asparagine (Asn), a serine/threonine (Ser/Thr) or a tryptophane (Trp) of a protein to form N-glycosylproteins, O-glycosylproteins or C-mannosylated proteins or even ribonucleic acids (RNA) to form the recently discovered glycoRNAs (Flynn *et al.*, 2021).

Essential for life, glycosylation is the enzymatic process by which a carbohydrate is covalently attached to a functional group of another molecule to form a glycoconjugate. Glycosylation is considered as the most complex co- and post-translational modification of proteins (>50% of all eukaryotic proteins) through the large number of key actors involved in all enzymatic steps. In parallel to amino acids, proteins and lipids, carbohydrates and glycosylation contribute to explain the enormous dynamic biological complexities of diverse organisms. One of the most striking examples is that the surface of any cell is decorated with a coat of glycoconjugates called glycocalyx and that there is no life without glycosylation (Varki *et al.*, 2011).

Glycoconjugates are extremely important for cellular life as supporting the molecular dialogue between cells (Varki, 2017). Involved in several crucial processes of cell mechanisms, glycosylation affects key protein functions, including folding, stability, solubility and immunogenicity. Likewise, glycans of glycoconjugates are also key elements involved in many physio-pathological

processes such as modulation of cell-cell communications, helping sperm to recognize eggs during fertilization, in host-pathogens interactions (Varki and Lowe, 2009). Glycosylation is considered as an evolutive mechanism which generate new structures without alteration of genetic information. The great diversity of glycoconjugates found in vertebrates results from a great complexity in the organization, composition and nature of the glycan structures formed. Notably, through their great potential for structural diversity, membrane surface glycans allow cells to escape pathogens. The Red Queen concept adapted by Pr A. Varki (Varki, 2011) illustrates their importance in host-pathogen interactions: during evolution, the host cell modifies its glycan epitopes (glycotopes) to escape a pathogen while the pathogen constantly adapts its glycome to recognize the host cell. Moreover, these glycotopes are recognized in most chronic and infectious diseases by receptors on the surface of bacteria, viruses and cells of the immune system (Nizet and Esko, 2009).

II. BIOSYNTHESIS OF GLYCOCONJUGATES

During their biosynthesis, glycoconjugates undergo a cascade of processes, where a multitude of enzymes, cofactors and donor substrates are involved in various cell compartments essential to their functions in cells. Particularly, N-glycosylation, described below, is crucial for the good structural conformation and function of N-glycosylproteins in cells (Figure 2).

This part of chapter will describe the glycosylation precursors and transporters required to initiate glycosylation pathways and will focus on the three major glycosylations that I was interested in my thesis: N-glycosylation, “mucin-type” O-glycosylation and gangliosides, a specific glycosylation which occurs on lipids (Figure 2). Nevertheless, many other glycosylations exist (Schjoldager *et al.*, 2020) such C-mannosylation, GPI-anchored proteins or 10 other types of O-glycosylation like O-GlcNAcylation mainly associated with emerging mechanism in cancer (El Yazidi-Belkoura and Lefebvre, 2022) or other glycosylation processes involved in polysaccharides chains synthesis with hyaluronan, glycosaminoglycans and also proteoglycans.

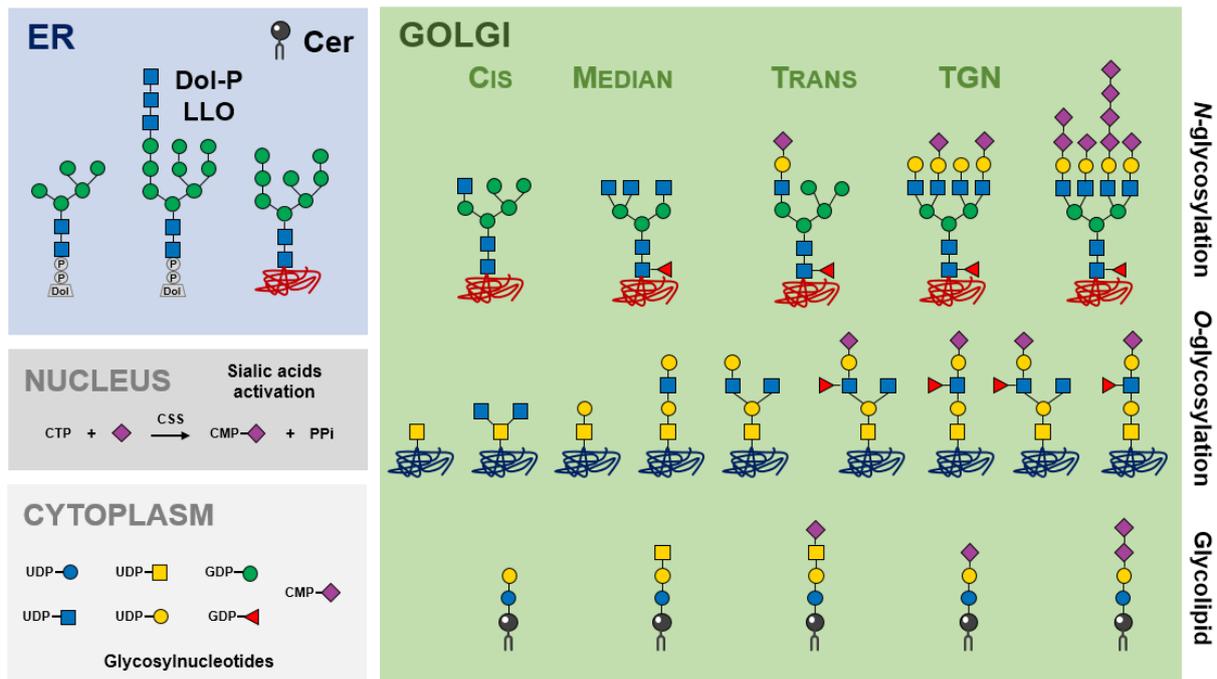


Figure 2: Simplified representation of the glycosyl nucleotides and the three types of glycosylation developed below. N-, O-glycosylation on proteins and glycosylation on lipids are represented through their different processes of initiation, elongation and maturation occurring in the various cell compartments. To be noted in vertebrates, sialic acids are mainly activated in the nucleus by the CMP-Sialic acid Synthase (CSS) to form CMP-Sia. ER : endoplasmic reticulum ; TGN : Trans-Golgi Network ; Cer : Ceramide ; PPi : inorganic pyrophosphate ; Dol-P LLO : Dolichol-Phosphate Lipid-Linked Oligosaccharide.

A. NUCLEOTID SUGAR DONORS

To be associated on glycoconjugates, monosaccharide precursors have to be imported from intracellular or external sugar sources (Freeze and Elbein, 2009), recognized and transported to different compartments with specific transporters for their activation and transfer (See Chapter III-B for Sialic acid transporters description). Most organisms can metabolize their own monosaccharides from diet or from recycling after a glycan degradation in the lysosomes. The main glycosyl-nucleotides for vertebrates are UDP- α -D-GlcNAc, UDP- α -D-Glc, GDP- α -D-Man, UDP- α -D-Gal, UDP- α -D-GalNAc, GDP- β -L-Fuc and CMP- β -D-Neu5Ac (Fig. 2). All glycosyl-nucleotides are activated in the cytosol except for CMP-Sialic acid (CMP-Sia) which is activated in the nucleus by the CMP-Sialic acid synthase (CSS) (Figure 2). Furthermore, only sialic acid is associated with a nucleotide monophosphate among all the monosaccharides in vertebrates (Münster-Kühnel *et al.*, 2004). After activation, nucleotid sugar donors are recognized by specific

glycosyltransferases located in the ER and in the different compartments of the Golgi apparatus to catalyze the transfer of their monosaccharide onto the nascent glycoconjugate.

B. BIOSYNTHESIS OF N-GLYCOSYLPROTEINS

N-glycosylation occurs on a majority of glycoproteins with a common core structure $\text{Man}\alpha 1-6(\text{Man}\alpha 1-3)\text{Man}\beta 1-4\text{GlcNAc}\beta 1-4\text{GlcNAc}\beta 1-\text{Asn}$. This structure could be completed up to 5 antennae leading to three types of N-glycans: oligomannose type only composed of Man residues, complex type constituted with $\text{Gal}\beta 1-3/4\text{GlcNAc}\beta 1-\text{R}$ lactosamine type 1/2 motif and hybrid which is a mix of both.

Initiated in the Endoplasmic Reticulum (ER) with dolichol phosphate (Dol-P) coming from farnesyl-pyrophosphate (FPP) of the mevalonate pathway, the synthesis of N-glycan core involves a succession of specific glycosyltransferases, the enzymes of glycosylation. Oriented firstly to the cytosolic side until $\text{Man}_5\text{GlcNAc}_2\text{-P-P-Dol}$ structure (Table 1A), the synthesis of Lipid Linked-Oligosaccharide (LLO) continues after translocation into the ER lumen with a myriad of glycosyltransferases which uses Dol-P-Mannose and Dol-P-Glucose as donor substrates to synthesize the $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2\text{-P-P-Dol}$ structures, named LLO (Table 1B).

Table 1: Reticular initiation for the N-glycans biosynthesis. Following successive actions of glycosyltransferases onto Dol-P (A), LLO was generated (B) and its transfer catalyzed by the oligosaccharyltransferase (OST) onto an amino acid sequence. Then, ER quality control system (C) is checkpoint and lead to the transport towards Golgi apparatus if the glycoprotein is well-folded (D). Nomenclature for glycosyltransferase name is used according to the Human Gene Organization (HUGO) (Eyre *et al.*, 2006).

Substrate donor	Glycosyltransferase name	Structure formed
A. ER membrane with cytoplasmic orientation		
GlcNAc-1-P	DOLK	Dol-P
	DPAGT1	GlcNAc-P-P-Dol
UDP-GlcNAc	ALG13/ALG14	GlcNAc ₂ P-P-Dol
	ALG1	Man ₁ GlcNAc ₂ P-P-Dol
GDP-Man	ALG2	Man ₂ GlcNAc ₂ P-P-Dol
	ALG11	Man ₅ GlcNAc ₂ P-P-Dol
B. Translocation with lumen orientation in ER membran		

	ALG3	Man ₆ GlcNAc ₂ P-P-Dol
Dol-P-Man	ALG9	Man ₇ GlcNAc ₂ P-P-Dol
	ALG12	Man ₈ GlcNAc ₂ P-P-Dol
	ALG9	Man ₉ GlcNAc ₂ P-P-Dol
Dol-P-Glc	ALG6	Glc ₁ Man ₉ GlcNAc ₂ P-P-Dol
	ALG8	Glc ₂ Man ₉ GlcNAc ₂ P-P-Dol
	ALG10	Glc ₃ Man ₉ GlcNAc ₂ P-P-Dol (LLO)

Glycosylation enzymes	Structure formed
-----------------------	------------------

C. ER membrane with lumen orientation

OST complex	Glc ₃ Man ₉ GlcNAc ₂ -Asn
MOGS and GANAB/PRKCSH	Glc ₁ Man ₉ GlcNAc ₂ -Asn
GANAB/PRKCSH	Man ₉ GlcNAc ₂ -Asn

Checkpoint with CANX/CALR

MAN1B1	Man ₈ GlcNAc ₂ -Asn
--------	---

D. Transport towards Golgi apparatus through ER-to-Golgi secretion pathway

Uncorrect folding of glycoprotein is supported by the ER quality control system with the soluble UDP-Glucose:glycoprotein glucosyltransferase (UGGT), Glucosidase II with its two subunits GANAB and PRKCSH and the calnexin/calreticulin (CANX/CALR) cycles (Table 1C). Misfolded glycoproteins have their glycan part firstly degraded by MAN1B1 and successive actions of EDEM1/2/3 and are translocated onto cytoplasm which lead to their poly-ubiquitinylation and proteasome degradation through the ERAD complex (Parodi *et al.*, 2014).

If there is correct folding, ER Alpha 1,2-Mannosidase (MAN1B1) removes a Man residue and the glycoprotein follows the ER-to-Golgi secretion pathway to be matured in the Golgi apparatus. Leaving ER compartment by anterograde traffic, glycoproteins are transported through vesicles formed by COP II coat to the Golgi apparatus (Table 1D). Then, they are matured by sequential enzymatic actions of glycosidases and glycosyltransferases in the different cis-/median-/trans-golgi compartments (Table 2A,B,C). Moreover, the cisternal organization of the Golgi apparatus is probably the main factor ensuring the sequential enzymatic process, particularly with Conserv Oligomeric Golgi complex (COG) (Reynders *et al.*, 2011) to form hybrid and complex types N-glycans with terminal monosaccharides as sialic acids or fucose (Fig. 2).

Table 2: Golgi maturation of N-glycans. N-glycosylproteins are matured in the Golgi apparatus through successive actions of glycoside hydrolasases and glycosyltransferases in the three different (A) cis-, (B) median- and (C) *trans*-Golgi compartments. HUGO nomenclature is used for enzymes name.

Substrate donor	Glycosylation enzymes	Structure formed
A. Cis-Golgi compartment		
-	MAN1A1/A2/C1	Man ₅ GlcNAc ₂ -Asn
UDP-GlcNAc	MGAT1	GlcNAc ₁ Man ₅ GlcNAc ₂ -Asn
B. Median-Golgi compartment		
-	MAN2A1	GlcNAc ₁ Man ₃ GlcNAc ₂ -Asn
	MGAT2	GlcNAc ₂ Man ₃ GlcNAc ₂ -Asn
UDP-GlcNAc	MGAT4A/B	GlcNAc ₃ Man ₃ GlcNAc ₂ -Asn
	MGAT5	GlcNAc ₄ Man ₃ GlcNAc ₂ -Asn
GDP-Fuc	FUT8	GlcNAc ₄ Man ₃ (Fuc ₁)GlcNAc ₂ -Asn
C. Trans-Golgi compartment		
UDP-Gal	Galactosyltransferases	Gal ₂ GlcNAc ₂ Gal ₄ GlcNAc ₄ Man ₃
UDP-GlcNAc	N-acetyl-glucosaminyltransferases	(Fuc ₁)GlcNAc ₂ -Asn
CMP-Sia	Sialyltransferases	Neu5Ac ₀₋₃ Gal ₂ GlcNAc ₂ Gal ₄
		GlcNAc ₄ Man ₃ (Fuc ₁)GlcNAc ₂ -Asn
GDP-Fuc	Fucosyltransferases	Neu5Ac ₀₋₃ Gal ₂ GlcNAc ₂ Gal ₄ (Fuc ₁)
		GlcNAc ₄ Man ₃ (Fuc ₁)GlcNAc ₂ -Asn

In summary, N-glycosylation leading to well-folded N-glycosylproteins is a pathway extremely complex and dependent of cell compartments, transporters and all these glycosylation enzymes and their enzymatic activity, availability of donor and acceptor substrates and all the other actors like cofactors involved in the glycosylation machinery.

C. BIOSYNTHESIS OF MUCIN TYPE O-GLYCANS

In contrast to N-glycosylation which is initiated in the ER with Dol-P and LLO transfer onto nascent protein, mucin type O-glycosylation, the most known and studied O-glycosylation, is initiated onto the Ser/Thr of a protein by addition of GalNAc. Firstly, the GalNAc residue is transferred from UDP-GalNAc in the cis-Golgi compartment by enzymatic catalysis of a large homologous Polypeptide-N-Acetylgalactosaminyltransferases (GALNT, also named ppGalNAcT) (Bennett *et al.*, 2012) generating the so-called Tn-antigen (Table 3A). This core structure serves as acceptor substrates for glycosyltransferases to synthesize core-1 (T-antigen) or

core-3 with C1GALT1 or B3GNT6 respectively (Bennett *et al.*, 2012), two major forms that could be extended in core-2 or core-4 mucin types as described below in Table 3A. Tn-antigen could be also extended in one of the minor core structures such as core-5, -6, -7 and -8 in the cis-Golgi.

Afterwards, core-1 O-glycosylproteins are directed to median-Golgi to undergo elongation with various N-acetylglucosaminyltransferases and Galactosyltransferases like B3GNT in GT31 CAZy family (Petit *et al.*, 2021), B4GALT and B6GNT (Table 3B). Finally, these Tn-antigen and different core structures go in the *trans*-Golgi compartment for sialic acids and/or fucose additions through sialyltransferases and fucosyltransferases activities (Table 3C), particularly sialyl T-antigen (sT) and sialyl Tn-antigen (sTn) as described in part III-E.

Table 3: Biosynthesis of mucin type O-glycans in the Golgi. Major core mucin types are represented in the cis-golgi compartment (A). Elongation and termination of mucin type O-glycans take place in the (B) median- and (C) *trans*-Golgi compartment respectively. HUGO nomenclature is used for enzymes name.

Substrate donor	Glycosyl transferase	Mucin type	Structure formed
A. Cis-Golgi compartment			
UDP-GalNAc	GALNT	Tn	GalNAc-Ser/Thr
UDP-Gal	C1GALT1	Core-1	Gal β 1-3GalNAc-Ser/Thr
UDP-GlcNAc	GCNT1	Core-2	Gal β 1-3(GlcNAc β 1-6)GalNAc-Ser/Thr
UDP-GlcNAc	B3GNT6	Core-3	GlcNAc β 1-3GalNAc-Ser/Thr
UDP-GlcNAc	GCNT3	Core-4	GlcNAc β 1-3(GlcNAc β 1-6)GalNAc-Ser/Thr
B. Median-Golgi compartment			
UDP-GlcNAc	B3GNT2		GlcNAc β 1-3Gal β 1-3GalNAc-Ser/Thr
UDP-Gal	B4GALT		Gal β 1-4GlcNAc β 1-3Gal β 1-3GalNAc-Ser/Thr
UDP-GlcNAc	B6GNT		GlcNAc β 1-3(GlcNAc β 1-6)Gal β 1-3GalNAc-Ser/Thr
UDP-Gal	B4GALT		Gal β 1-4GlcNAc β 1-3(GlcNAc β 1-6)Gal β 1-3GalNAc-Ser/Thr
C. Trans-Golgi compartment			
CMP-Sia	Sialyl transferases		Neu5Ac α 2-6GalNAc-Ser/Thr
			Neu5Ac α 2-3Gal β 1-3GalNAc-Ser/Thr
GDP-Fuc	Fucosyl transferases		Neu5Ac α 2-3/6Gal β 1-4GlcNAc β 1-3(GlcNAc β 1-6)Gal β 1-3GalNAc-Ser/Thr
			Neu5Ac α 2-3/6Gal β 1-4(Fuc α 1-3/4)GlcNAc β 1-3(GlcNAc β 1-6)Gal β 1-3GalNAc-Ser/Thr

D. BIOSYNTHESIS OF GANGLIOSIDES

Glycosphingolipids (GSL) are a specific type of glycoconjugates made of a lipid part (ceramide (Cer)) resulting from its *de novo* synthesis in the cytosolic side of ER, and a glycan part associated to ceramide in the Golgi through glycosyltransferases activity. Gangliosides are acidic GSL mainly expressed in the brain during development and overexpressed in different diseases such as neurodegenerative diseases and cancers.

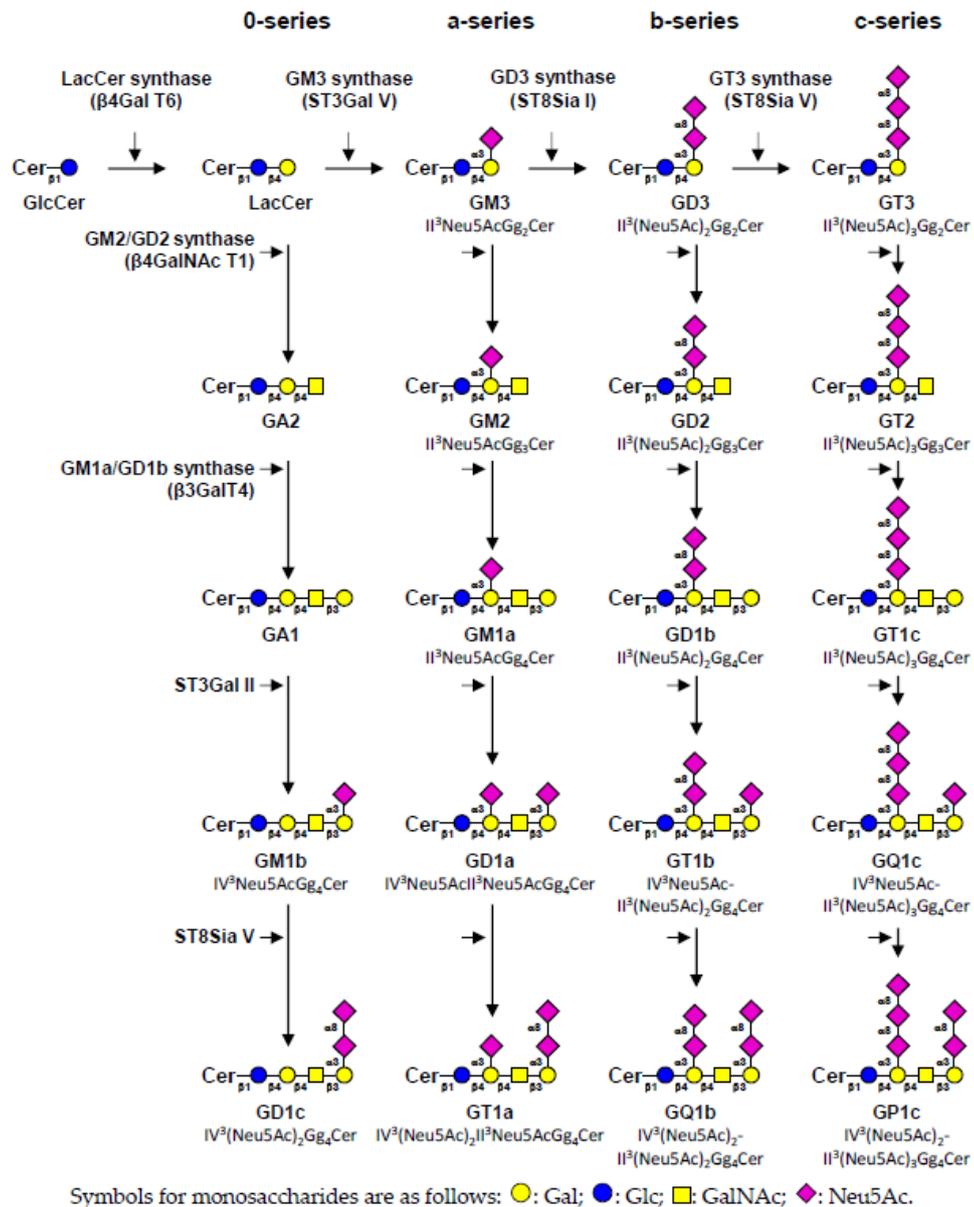


Figure 3: Ganglioside biosynthesis with the specific glycosyltransferases involved in each enzymatic reactions. To be noted, various sialyltransferases ST3Gal II, ST3Gal V, ST8Sia I and ST8Sia V are involved in the transfer of sialic acids with various specificity onto structures in the four 0-, a-, b- and c-series.

Illustrated from Cavdarli *et al.*, 2019a.

To initiate their biosynthesis, Cer is transported to the Golgi apparatus, recognized by UDP-Glucose Ceramide Glucosyltransferase (UGCG) which catalyze the transfer of Glc residue from UDP-Glc to generate GlcCer, then translocated to the Golgi lumen and converted in LacCer by B4GALT6 enzyme. At this point, LacCer represents a metabolic crossroads for the biosynthesis of various gangliosides by subsequent actions of glycosyltransferases, as ST6GalNAc III/IV/V/VI involved in the biosynthesis of α -series gangliosides or especially the three sialyltransferases ST3Gal V, ST8Sia I and ST8Sia V involved in the synthesis of all ganglioside precursors for a-, b-, c- series as described in Figure 3.

III. SIALIC ACIDS AND SIALYLATED GLYCOCONJUGATES

Among all monosaccharides, sialic acids (Sia) were the favorite molecules I worked with and studied during my thesis. This part will describe their discovery, incredible structural diversity and biosynthesis in Eukaryota.

A. SIALIC ACIDS DISCOVERY AND STRUCTURAL DIVERSITY IN VERTEBRATES

Sia were firstly discovered by Klenk and Blix in rat salivary mucin and on gangliosides in the 1940's and described by Gottschalk resulting from the chemical condensation between pyruvic acid and N-acetyl hexosamine. Also named neuraminic acids, Sia belong to a 9-carbon backbone monosaccharides family derived from 5-amino-3,5-dideoxy-D-glycerol-D-galacto-non-2-ulosonic acid (Neu) and are the most frequent terminal residues of glycans carried by glycoconjugates. Located at this terminal position and negatively charged on its carboxylate C1-group at physiological pH, Sia play a fundamental role in all biological processes (Schauer and Kamerling, 2018).

Described for a long time in various animal species (Warren, 1963) and in Bacteria, Archea and Eukaryota (Lewis *et al.*, 2009), Sia occur ubiquitously in nature with the exception of plants (Tiralongo *et al.*, 2013). Over 50 different Sia are reported in nature (Angata and Varki, 2002) characterized by a high structural and functional diversity. Among them, N-acetylneuraminic acid (Neu5Ac), N-glycolylneuraminic acid (Neu5Gc) and 2-keto-3-deoxy-nonulosonic acid (Kdn) (Fig.

4) are the 3 major Sia found in vertebrates except in humans, which show only residues of Neu5Ac (Schnaar *et al.*, 2014).

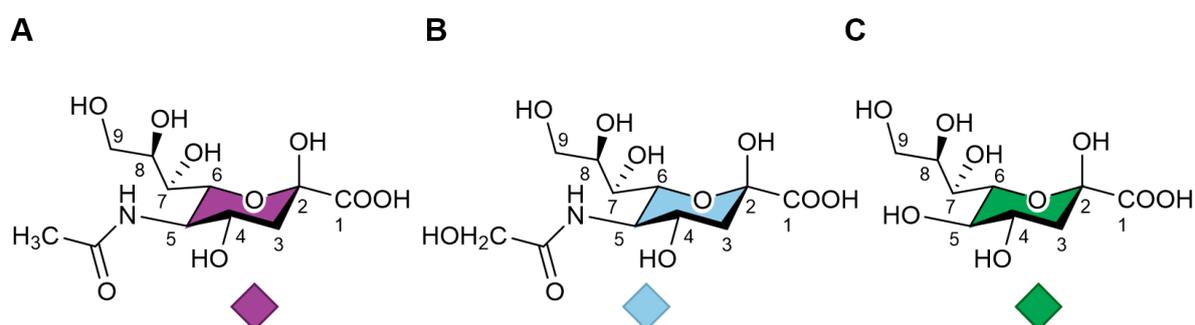


Figure 4: Three major forms of sialic acids in vertebrates. (A) Neu5Ac. (B) Neu5Gc. (C) Kdn.

As mentioned above, Sia residues include an ionized carboxylate group at C1-position, a keto group at C2-position often involved in linkage after its transfer catalyzed by a glycosyltransferase and a glycerol non-flexible side chain (Schauer and Kamerling, 1997) oriented in equatorial position from C6 to C9 (Fig. 4). Moreover, Sia residues can undergo various modifications such as acetylation, phosphorylation, methylation, sulfation and lacylation occurring on hydroxyl group at C4, C7, C8 and/or C9 positions (Fig. 5) (Sato *et al.*, 1993; Kitazume *et al.*, 1996 ; Schauer and Kamerling, 1997). Particularly, the Sialyl-O-Acetyltransferase CASD1 has been recently highlighted for O-Acetylation of GD2 ganglioside in breast cancer cells (Cavdarli *et al.*, 2021). Even lactonization is described as fine reaction to prevent Sia recognitions and cleavage by sialidases (Cheng *et al.*, 2004).

Interestingly, humans are not able to synthesize Neu5Gc because of inactivation of the *CMAH* gene about 2.8 million years ago (Irie *et al.*, 1998 ; Chou *et al.*, 1998 ; Varki, 2009) that is present in most vertebrates as cytosolic form to convert Neu5Ac into Neu5Gc (Shaw and Schauer, 1988).

The loss of CMAH could be explained by host-pathogen interactions hypothesized by Varki in the beginning of the century. Although in rat and mouse, two species models commonly used in laboratory, showed an inverse sialic acid ratio with 95% Neu5Ac and 5% Neu5Gc in rats and 11% Neu5Ac and 89% Neu5Gc in mice liver glycoproteins due to cytosolic activity of CMAH in mice (Lepers *et al.*, 1990). In this sense, a large glycomics study in several tissues of the zebrafish *Danio rerio* underlined the sialic acid biodiversity in several tissues, notably with high levels of Kdn in the

intestines and of Neu5Gc in the ovaries (Yamakawa *et al.*, 2018). Likewise, mixed diSia structures composed with Neu5Ac and or Neu5Gc linked with an $\alpha 2,8$ -linkage were found on zebrafish embryos (Chang *et al.*, 2009).

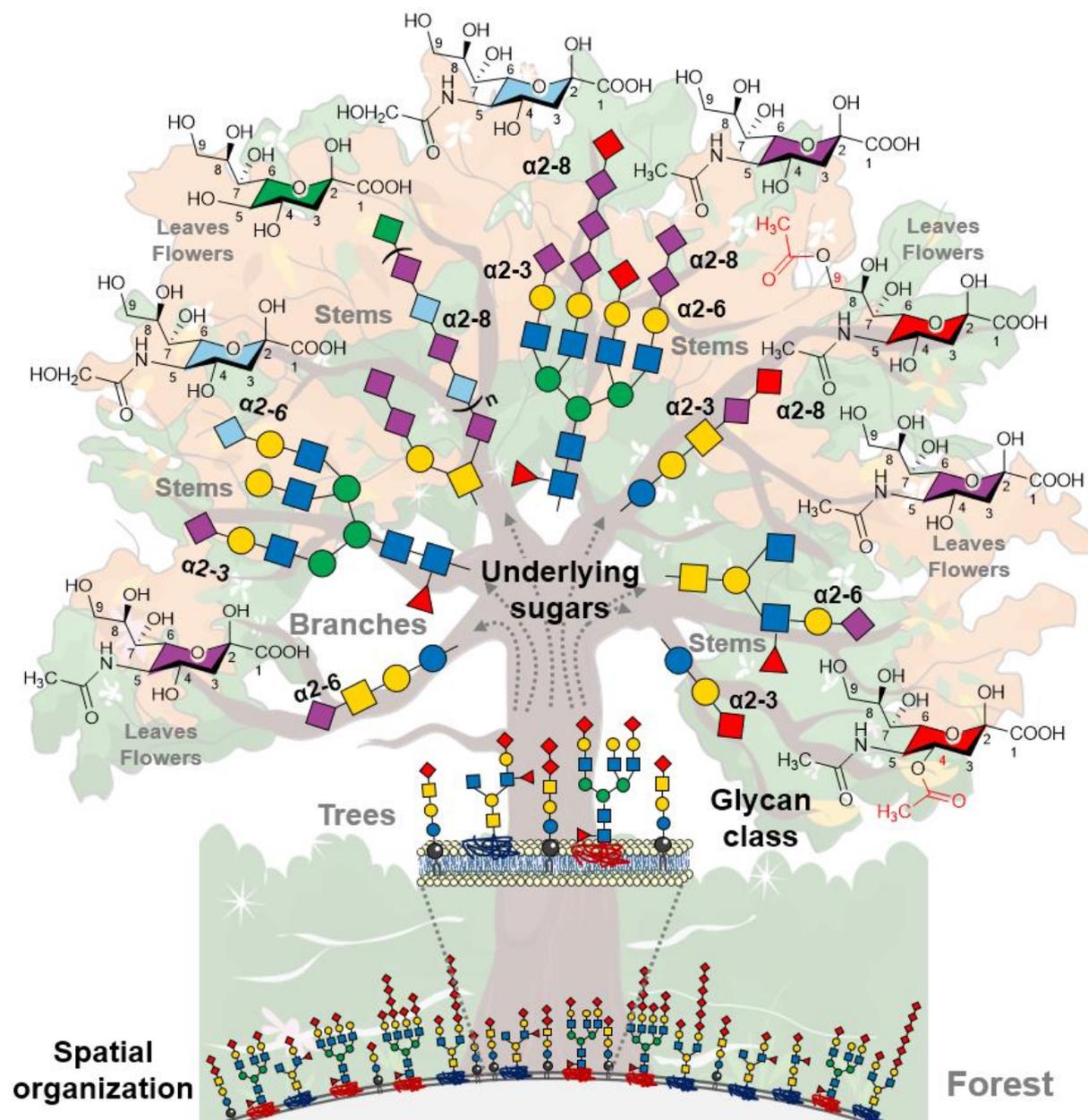


Figure 5: Complexity and diversity of sialylation on glycoconjugates. Graphical parallel from forest to leaves and flowers is done to summarize all the diverse structures that could be sialylated depending on the nature of glycoconjugates, glycans part, monosaccharides and linkages and finally, the nature and modifications occurring on Sia residues. Acetylation on hydroxyl groups on C-4 or C-9 of Sia (red) are highlighted in red. SNFG nomenclature is adopted for representation. Inspired by Cohen and Varki, 2010.

In response to inflammation caused by a pathogen, the salmonid *Salvelinus alpinus* displayed an increase of Neu5Gc levels in intestinal mucins (Venkatakrishnan *et al.*, 2019). Atlantic salmon O-glycans appear also heavily sialylated with Neu5Ac, Neu5Gc and Kdn detected in skin mucins (Jin *et al.*, 2015). To be noted, mono- and di-sialylated N-glycans are described as signature motifs in salmonid sera (Aoki *et al.*, 2021) and Neu5Gc is reported on glycophorin in red blood cells membranes of the carp *Cyprinus carpio* and of two perciformes *Seriola quinqueradiata* and *Pagrus major* (Aoki *et al.*, 2014). These sialo-oligosaccharide structures have bacteriostatic activity towards two pathogens found in aquaculture, *Vibrio anguillarum* and *Edwardsiella tarda* (Xu and Zhang, 2013), through interactions with their flagella (Aoki, 2022). Moreover, the quantity and nature of sialic acid polymers chains (polySia) in fish serum are much more important than in mammals (Zlatina *et al.*, 2018) and their presence on fish eggs suggests a protective role against pathogens infection (Sato *et al.* 1993).

In summary, this huge structural complexity and wide variety of sialic acids reported at different levels (Fig. 5) depending on the tissues and species considered, although dynamic, could explain complex biological molecular interactions. This area of investigation tend to be more and more explored by the scientific community.

B. SIALIC ACIDS BIOSYNTHESIS PATHWAY

Many complex pathways are involved in glycosylation machinery with series of enzymatic reactions to lead in Sia formation and sialylated molecules. Summarized in Figure 6, two pathways lead to cytosolic Neu5Ac in eukaryotic cells (Petit *et al.*, 2018).

From glucose incorporation, glycolysis and hexosamine pathway, UDP-GlcNAc can be used as an activated sugar donor substrate by the UDP-GlcNAc 2-epimerase/ManNAc kinase (GNE) to convert it in the epimeric N-Acetylmannosamine (ManNAc). GNE is a cytosolic bifunctional enzyme, which act as a master regulator of Sia synthesis (Hinderlich *et al.*, 2015) with its kinase activity responsible to phosphorylate ManNAc into ManNAc-6-P. Then, this pathway involves two enzymes: firstly, Neu5Ac 9-phosphate synthase (NANS) to condense phosphoenolpyruvate

(PEP) and ManNAc to form N-acetylneuraminic acid 9-phosphate (Neu5Ac-9-P). This molecule is used by N-acetylneuraminic acid 9-phosphatase (NANP) leading to cytosolic Neu5Ac (Fig. 6, part 2).

In parallel, free Sia can be recycled from degraded sialoglycoconjugates present at the cell surface or in the lysosome after cleavages by neuraminidases NEU1 and NEU4 (Monti *et al.*, 2010) and return to the cytosol by a specific sialin exporter SLC17A5 (Fig. 6, part 1). Indeed, SLC17A5 alterations are associated with pathological accumulation of lysosomal Sia storage called Sialic Acid Storage Disorders (SASD) (Strehle, 2003) including Salla disease which is an infantile sialic acid storage disease (Pertti Aula *et al.*, 1979).

Interestingly, acetylated analogs of ManNAc and Sia can be incorporated in cell through passive diffusion and are deacetylated by esterases to enter in the sialic acid biosynthesis pathway (Moons *et al.*, 2019) whereas in bacteria, a specific sialic acid transporter NanT is described embedded in the inner membrane of *Escherichia coli* to facilitate Sia incorporation (Martinez *et al.*, 1995). Furthermore, it has been shown that Neu5Gc could be incorporated into human cells through endocytosis and lysosomal sialic acid transporter (Bardor *et al.*, 2005) which raise many questions about its potential conversion to CMP-Neu5Gc in the nucleus, its transport, recognition and use by human sialyltransferases in the Golgi apparatus.

In the nucleus, Sias are activated with Cytidine Tri Phosphate (CTP) and its magnesium ion cofactor, a reaction catalyzed by CMP-sialic acid synthetase (CMAS, also known as CSS) leading to the formation of activated glycosyl nucleotide CMP-Neu5Ac. To be noted, deficiencies in Sia biosynthesis could lead to pathologies like sialuria with inhibition of GNE by CMP-Neu5Ac or sialidose which is an accumulation of sialoglycoconjugates with deficiencies of neuraminidases. During its transit to Golgi apparatus, CMP-Sia have an allosteric negative feedback to GNE (Chen *et al.*, 2016) to modulate cytosolic Sia quantities.

Afterwards, CMP-Sia are translocated across the membranes of the Golgi apparatus by SLC35A1, an antiporter that exchanges CMP-Sia for CMP. This mechanism was firstly mentioned studying UDP-Gal and CMP-Sia transport and wondering “how nucleotide sugars gain access to the transferases” (Fleischer, 1983). To study the role of SLC35A1, different systems were selected such as *Saccharomyces cerevisiae*, deficient for sialylation and CMP-Sialic acid transporter

(Berninsone *et al.*, 1997) and CHO-Lec2 and 6B2 mutant cells, deficient for CMP-Sia transporter (Eckhardt *et al.*, 1996). Complementation of these systems with the cloning of murine SLC35A1 led to effective and specific transport of CMP-Sia into yeast Golgi-enriched vesicles on one hand and to the detection of α 2,3-sialylated and polysialylated glycoproteins in CHO mutant cells on the other hand. Furthermore, it was shown that CMP-Neu5Gc could penetrate Golgi vesicles isolated from mouse liver and transfer of Neu5Gc onto endogenous acceptors, suggesting that CMP-Neu5Gc could be transported by mouse SLC35A1 (Lepers *et al.*, 1989). The identification of sialylation defects due to SLC35A1 inactivation in a young patient with high pathologies like severe hemorrhages, thrombocytopenia, respiratory distress syndrome and opportunistic infections led to SLC35A1 deficiency as type II CDG (Martinez-Duncker *et al.*, 2005).

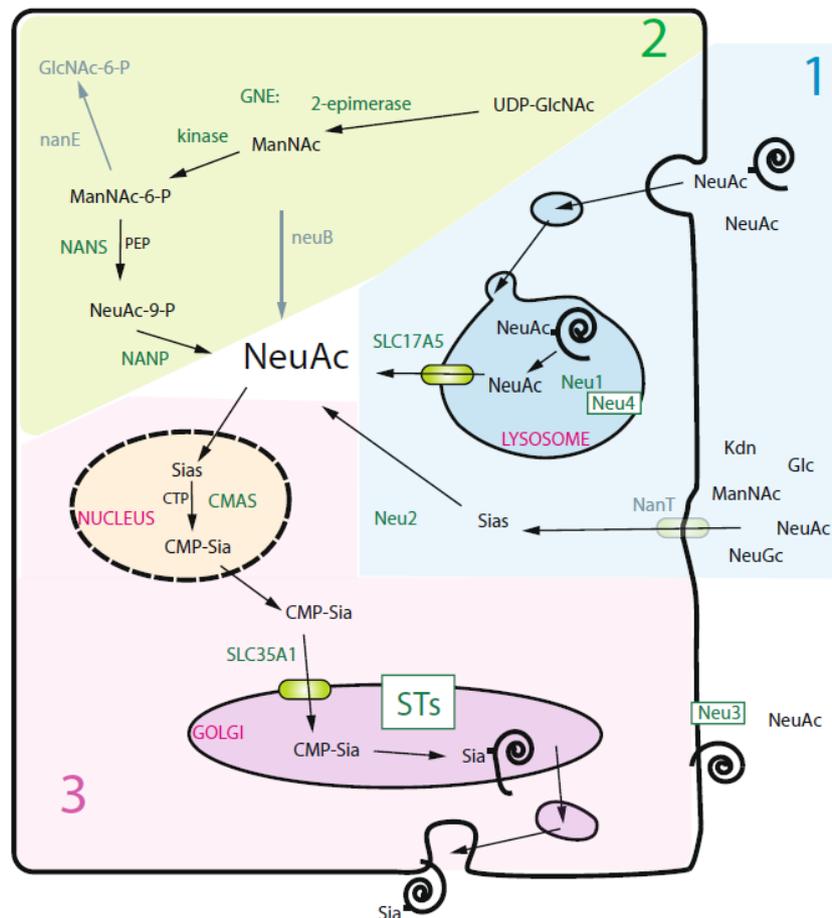


Figure 6: Sialic acid metabolism pathway in eukaryotic cells. Cytosolic Neu5Ac is coming from exogenous uptake (blue background) or from hexosamine pathway and the resulting UDP-GlcNAc molecule (green background). Then, Sia is activated in CMP-Sia in the nucleus by CMAS and transferred by various STs onto an acceptor in the Golgi apparatus. HUGO nomenclature is used for enzymes and transporters.

Illustration from Petit *et al.*, 2018.

In the Golgi lumen, sialyltransferases (developed in Chapter V) use CMP-Sia as donor substrate and catalyze their transfer onto acceptor glycoconjugate (Fig. 6, part 3). Finally, Sias associated to glycoconjugates are secreted, recycled or degraded by a cytosolic pyruvate lyase which cleaves Sias into pyruvate and ManNAc available for new Sia biosynthesis. Although complex, the Sia biosynthetic pathway has a common eukaryotic ancestral origin and was totally lost or partially maintained during eukaryotes evolution as proposed by phylogenetic studies (Petit *et al.*, 2018).

C. SIALYLATED GLYCOCONJUGATES IN PHYSIOLOGY

As illustrated in figure 5, sialylation is one of the terminal steps of glycosylation with fucosylation and N-acetylgalactosaminylation, which results in association of sialic acid residues to other monosaccharides such as galactose (Gal) or N-acetylgalactosamine (GalNAc) via α 2,3 or α 2,6 bonds or to other α 2,8 sialic acid residues (Fig. 5).

The terminal sialic acids of sialylated glycoconjugates carry negative charges and these particular physicochemical properties give sialoglycoconjugates a fundamental role in biological processes such as embryonic development and immunity (Abeln *et al.*, 2018). Particularly due to their hydrophilic and negative charges, sialic acids could be modulators of hydration status of high sialylated glycoproteins like mucins in aim to lubrication and protection of epithelia surfaces (Baos *et al.*, 2012). Interestingly, dynamical changes of sialic acids and sialoglycoproteins occur in gametes and during fertilization (Feng *et al.*, 2016). It has to be noted that sulfated sialoepitopes present at the sperm surface glycans are a necessary condition in the process of sperm cell maturation and the establishment of interaction between sperm and egg cells (Fliniaux *et al.*, 2022) and that sulfation of Sia is ubiquitous and essential for vertebrate development (Ertunc *et al.*, 2022).

The extraordinary biodiversity of sialoglycoconjugates described in vertebrates contributes to the functioning and maintenance of the cell by allowing it to interact with its environment. Especially in fish, a wide variety of sialoglycoconjugates composed of different sialic acids are found (Fig. 4) and on fertilized eggs, these sialoglycoconjugates form a particularly effective protective barrier against environmental factors and pathogens (Kitazume *et al.*, 1994).

D. SIALYLATED GLYCOCONJUGATES IN PATHOLOGICAL STATES

Sialoglycoconjugates also play a key role in many disease processes. In most cancers, sialylation levels increase and during tumor progression, cancer cells direct their metabolism towards increased synthesis of sialic acids and control sialylation of glycoproteins involved in metastatic processes (Almaraz *et al.*, 2012 ; Goni *et al.*, 2021). For example, sialyl-Lewis A (sLe^a) (Table 4) is a sialylated antigenic determinant expressed in the colorectal cancer that acquires tumor marker properties. Notably in the pancreas, CA19-9 is a diagnostic marker of sLe^a used for the monitoring and progression of pancreatic cancer. These tumour-associated carbohydrate antigens (TACA) such as sialyl-Lewis X (sLe^x) in colon cancer or sLe^a (Table 4) also help in early diagnosis, prognosis and monitoring in various cancers (Vajaria *et al.*, 2016 ; Martinez-Duncker *et al.*, 2011). Moreover, GD3 ganglioside is oftenly O-acetylated on its Sia residues when overexpressed in cancer (Cavdarli *et al.*, 2020).

Singularly in type 2 diabetes, patients display an increase of highly branched N-glycans with α 2,6-linked sialic acids (Rudman *et al.*, 2019). The fundamental role of sialoglycoconjugates appears clearly in a set of rare and hereditary pathologies with recessive autosomal transmission called CDG (Péanne *et al.*, 2018). CDG affect the proper development of the nervous system, as the example mentioned previously of SLC35A1-CDG where a deficiency of the Golgi transporter of Sia leads to a defect in sialylation of glycoconjugates. Sialuria and sialidose are also CDG with deficiencies of GNE inhibition and lysosomal neuraminidases respectively that lead to severe mental and motor retardation for patients (Strehle, 2003). Likewise, mutations of specific sialyltransferase ST3Gal V, named ST3GAL5-CDG, lead to high pathology at patient from failure to thrive, severe hearing, visual, motor, and cognitive impairment and respiratory chain dysfunction (Indellicato *et al.*, 2019).

Otherwise during microbial infections, many bacteria and viruses recognize sialic acids present on host cell glycoconjugates. For example, *Helicobacter pylori* expresses adhesins which recognize the sialic acids linked in α 2,3 of the oligosaccharide chains on the surface of gastric epithelial cells (Hirno *et al.*, 1996) to infect the gastric mucosa or the influenza virus Influenza possesses surface

haemagglutinins that bind specifically to α 2,6-linked sialic acid residues on the surface of host cells of the respiratory tract in humans (Couceiro *et al.*, 1993). It has been described that 9-O-Acetylation of Sia is directly used as primary receptor determinant by Influenza virus C for attachment to cell surface receptors (Rogers *et al.*, 1986). In host-pathogen interactions, sialic acids present at cell surface of pathogen such as bacteria (*E. coli* K1, *N. meningitidis* B and C) are remarkably described to confer a protective barrier and immune escape. Modulation of interaction can therefore have huge effects on zoonotic transmission and virulence of many viruses. Studies focusing on human coronavirus have demonstrated that sialylation is extremely important for interaction between spike (S) glycoprotein and its ACE receptor found at cell surface (Zhao *et al.*, 2020). S glycoprotein shows 9-O-acetylated Sia that allows binding mainly to sLe^x, confirmed by site-directed mutagenesis and *in vivo* assays (Tortorici *et al.*, 2019).

E. SIALYLATED EPITOPES : BIOMARKERS AND RECOGNITIONS

These last two decades, glycan-based biomarkers gain more and more importance due to the glycoconjugates essential role for cell functions and interactions. As mentioned for pancreatic cancer with CA19-9, sialylated markers could serve as diagnostic glycosignature and biomarkers. Aberrant sialylation is often correlated as a hallmark of cancer with progression, metastasis and linked with disease progression status but also more generally in human diseases. In inflammation and immunity, a great diversity of terminal sialoepitopes have important dynamic role to modulate the response. These sialylated motifs such as Lewis and sialyl Lewis antigens, Tn and sialyl-Tn antigens mediate the interactions between haematopoietic cells and their progenitors in the bone marrow and are involved in in regulating leukocyte trafficking (Reily *et al.*, 2019).

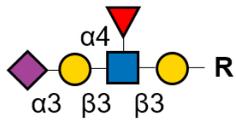
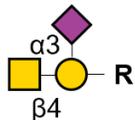
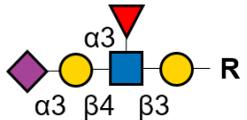
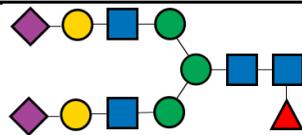
Siglecs and selectins are carbohydrate-binding proteins (CBP) known to regulate cell-cell adhesion or cell-protein interaction with a specific glycanic epitope recognition. In humans, Siglecs are a family of 16 members of sialic acid-binding lectins present on immune cell surface that regulate immune responses and cell homing through their interactions (Reily *et al.*, 2019). Sia quantities are modulated with siglecs which are considered as regulators of inflammation (Schnaar, 2016). Selectins have high affinity for sugar moieties and are markedly described in immune

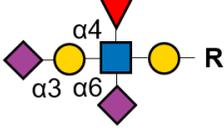
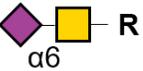
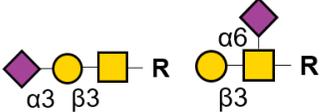
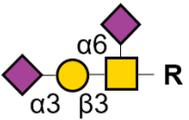
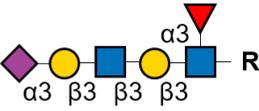
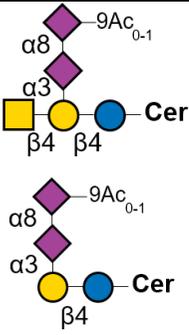
response. E- and P-selectins are known to initiate inflammation by cell-cell interaction with sLe^x, siglecs have regulatory effect against inflammation. Notably, Siglec-1 (sialoadhesin, CD169) expressed on activated monocytes and macrophages is involved in the clearance of α 2-3 sialylated human pathogens. Many siglecs are described to have an altered profile of expression in immune cells in HIV individuals. Siglec-7 and -9 are biomarkers for dysfunctional NK cells and Siglec-2 (CD22) and -6 for B cells interactions with their α 2-6 sialylated ligands (Colomb *et al.*, 2019).

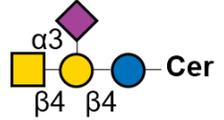
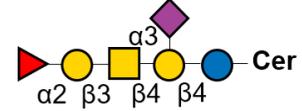
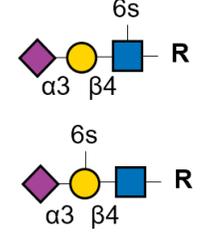
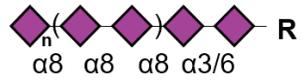
Modifications like 6-O-sulfation on sLe^x has direct impact for Siglec-8 recognition and will be useful to design Siglec-8-targeted agonists in allergic and inflammatory diseases (Pröpster *et al.*, 2016). Interestingly, Siglec-11 could selectively bind polySia and decrease pro-inflammatory activation on microglia cells and macrophages in spleen, intestine and liver tissues (Angata *et al.*, 2002b ; Wang and Neumann, 2010) and *in vivo* on mice (Karlstetter *et al.*, 2017). Siglecs bind to sialic acid through charges interaction between their conserved arginin residues in immunoglobulin domain with negative carboxyl group of Sia (Hartnell *et al.*, 2001) and their diversity could be explained by evolution mechanism against pathogen infection (Macauley *et al.*, 2014).

Sialoepitopes are signatures which serve as glycosylation-based biomarkers in lots of pathologies including cancer (Mereiter *et al.*, 2016) and neurodegenerative diseases like Alzheimer (Haukedal and Freude, 2021). Glycosylation changes and TACA are often observed in cancer cells and are considered as oncofoetal antigens similar to those observed in early development. These alterations are described as an increase of sialyl Lewis structures and sialylTn (sTn) antigen in O-glycans of MUC1 which promotes tumor growth and metastatic spread (Harduin-Lepers *et al.*, 2012). Sialyl-Lewis A (sLe^a) or sialyl-Lewis X (sLe^x), the E-selectin ligand glycans, are extensively expressed on numerous types of cancer cell surface, including colorectal, pancreatic, gastric, breast, prostate and lung cancer and is described to increase metastasis through extravasation (Saiki *et al.*, 1996). sT and sTn antigens are detected in cancer patient serum and serve as glycan-based biomarkers, notably sTn which alter adherence due to sialylation of O-glycans in cancer. All these sialoepitopes are listed in the table below.

Table 4: Representation of different sialylated epitopes structures reported as biomarkers in pathologies and diseases. The graphical representation is based on accepted SNFG conventions for the nomenclature of glycans and monosaccharides (Varki *et al.*, 2015a). SCLC and NSCLC : small and non-small cell lung carcinoma.

EPITOPE	SNFG REPRESENTATION OF GLYCAN STRUCTURE	GLYCAN CLASS	PATHOLOGIES	LITTERATURE
Sialyl Lewis A (CA19-9, sLe^a)		O-glycans	Gastrointestinal tumors Often elevated in cancer patient serum (pancreatic, gastric, colorectal)	Tang <i>et al.</i> , 2016 Ferreira <i>et al.</i> , 2019
Sd^a/Cad		Sd ^a on N-glycans, major Cad on core-1 and core-3 O-glycans and glycolipids (GM2)	Downregulated in colon cancer (in balance with sLe ^x)	Dall'Olio <i>et al.</i> , 2014 Groux-Degroote <i>et al.</i> , 2021
Sialyl Lewis X (sLe^x)		sLe ^x present on N-glycans, GSL, O-glycans, on MUC7, MUC16 in ovary	Carcinoembryonic antigen (CEA) Breast, gastric and colorectal cancer CA125 antigen in MUC16 tandem repeat Autoimmune Sjögren's syndrome	Mereiter <i>et al.</i> , 2016 Trinchera <i>et al.</i> , 2017 Ferreira <i>et al.</i> , 2019 Karlsson and Thomsson, 2009 Felder <i>et al.</i> , 2014
AFP-L3		Aberrant core-fucosylation on α -foetal protein (AFP) N-glycans	Hepatocellular carcinoma	Li <i>et al.</i> , 2001 Rojas <i>et al.</i> , 2018 Zhou <i>et al.</i> , 2021

Disialyl Lewis A (Di-sLe^a)		O-glycans	Colon cancer	Tsuchida <i>et al.</i> , 2003
Sialyl Tn (CD175s, sTn)		Sialylation on Tn antigen on O-glycans, on MUC1	Bladder cancer ; nephropathy Breast cancer ; pancreatic adenocarcinoma	Sewell <i>et al.</i> , 2006 Munkley, 2016
Sialyl-T Thomsen-Friedenreich (sT)		Sialylation on TF antigen on O-glycans, on MUC1	Prostate, ovarian, lung, breast (serum) cancer ; pancreatic adenocarcinoma	Taylor-Papadimitriou <i>et al.</i> , 1999 Patil <i>et al.</i> , 2015
Disialyl T		O-glycans	Human pancreatic cells	Tsang <i>et al.</i> , 2022
VIM-2 (CD65s)		O-glycans	Human leukemia cells	Robinson <i>et al.</i> , 1994
GD3, GD2 ± O-acetylation		GSL	Melanoma, neuroblastoma, acute lymphoblastic leukemia, breast cancer, glioma, SCLC, osteosarcoma	Cavdarli <i>et al.</i> , 2019a Cavdarli <i>et al.</i> , 2019b Cavdarli <i>et al.</i> , 2020 Daniotti <i>et al.</i> , 2013

GM2		GSL	Overexpressed in Tay-Sachs disease Melanoma	Daniotti <i>et al.</i> , 2016
Fucosyl GM1		GSL	SCLC	Drivsholm <i>et al.</i> , 1994
6-/6'-sulfosialyl LacNAc		N-glycans on PSGL-1 (P-Selectin glycoprotein ligand 1)	Inflammatory dendritic cells	Schäkel <i>et al.</i> , 2002 Hofer <i>et al.</i> , 2019
Neu5Gc		N-/O-glycans	Oncofoetal human colonic, melanoma, retinoblastoma and breast cancer cells	Samraj <i>et al.</i> , 2014
Polysialic acid chains (PolySia)		N-/O-glycoproteins and some glycolipids	Mental disorders and brain diseases NSCLC metastasis and lung cancer aggressiveness, gliomas Nephroblastoma and neuroblastoma Autoimmune disease scleroderma	Schnaar <i>et al.</i> , 2014 Thiesler <i>et al.</i> , 2022 Elkashef <i>et al.</i> , 2016 Villanueva-Cabello <i>et al.</i> , 2022

IV. POLYSIALIC ACID (POLYSIA): STRUCTURE AND BIOLOGICAL ROLES

A. STRUCTURE OF POLYSIA

Firstly isolated and described as colominic acid on *E. coli* K-235 bacteria capsules by Goebel and Barry in 1957, polySia refers to association of monomers of sialic acids in polymers ranging from 8 to 400 Sia residues (Sato and Kitajima, 2013). PolySia are considered as a modification of glycoproteins that occurs mainly on N-glycans and sometimes on O-glycans to form polymeric anionic chains. Their biosynthesis and regulation are ensured by enzymes named polysialyltransferases and will be developed further in Chapter VI. These polymers adopt a dynamic conformation which could be helicoidal and/or linear depending on their type of linkage: in α 2,8-linkage in vertebrates and even α 2,9 in bacteria (Sato and Kitajima, 2013). Moreover, conformational studies by NMR demonstrated that 8 Sia residues are required for polySia helicoidal structures recognized by an anti-polySia antibody (Michon *et al.*, 1987; Nagae *et al.*, 2013). The structural diversity of polySia illustrated in figure 7 is based on their degree of polymerization (DP) that can reach 400 residues, the type of glycosidic bonds established, the nature of the sialic acids (Neu5Ac, Neu5Gc, Kdn) and their possible modifications like methylation, acetylation, sulfation or lactonization (Sato and Kitajima, 2021).

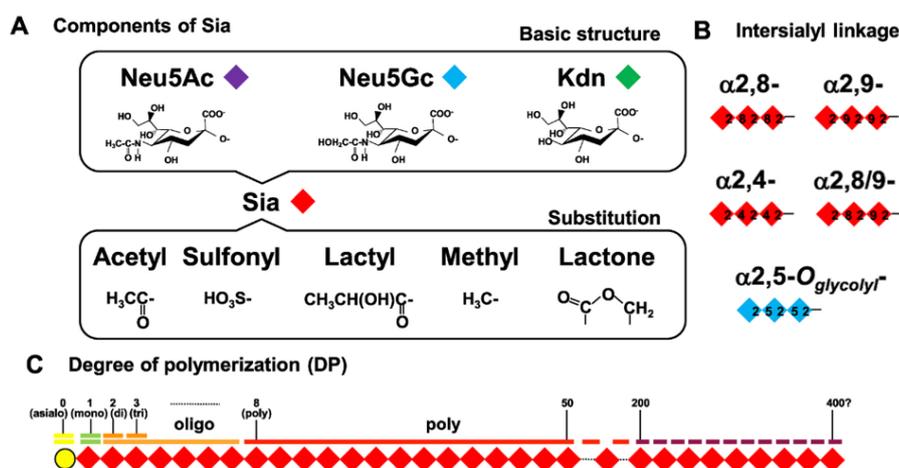


Figure 7: Diversity in oligo-/polySia chains at different levels. According to the nature of Sia (A), the intersialyl-linkage established (B), the size (DP) (C) and potential modification (acetylation, methylation, sulfation, lactylation or lactonization), oligoSia and polySia are extremely diverse in free polysaccharides chains or associated on glycolipids and glycoproteins.

Inspired from Sato and Kitajima, 2021

Table 5: List of glycoproteins described to be polysialylated in the nature. Adapted from Teinturier-Lelèvre thesis, 2006 ; Sato, 2013 ; Boyancé thesis, 2021.

Name	Expression profile	Glycan	Glycosylation sites	Enzymes	References
Neural Cell Adhesion Molecule (NCAM)	Nervous system, and immune cells (NK, NKT, LT CD4 ⁺ , neutrophils, monocytes cells, progenitor hematopoietic cells)	<i>N</i> -glycans	Asn ⁴⁴⁹ and Asn ⁴⁷⁸ in Ig5 domain	ST8Sia II and ST8Sia IV	Curreli <i>et al.</i> , 2007 Galuska <i>et al.</i> , 2010a Colley <i>et al.</i> , 2014 Schnaar <i>et al.</i> , 2014 Villanueva-Cabello <i>et al.</i> , 2022
Voltage-gated sodium channel α-subunit	Electric eel electroplax membranes and rat brain	?	?	?	James and Agnew, 1989 Zuber <i>et al.</i> , 1992
CD-36 scavenger receptor	Human, and murine milk	<i>O</i> -glycans	?	?	Yabe <i>et al.</i> , 2003
Neuropilin-2 (NRP-2)	Mature dendritic cells, microglia, murine macrophages	<i>O</i> -glycans	Specific 17 aa sequences between MAM and FV/VIII domains	ST8Sia IV	Curreli <i>et al.</i> , 2007 Rollenhagen <i>et al.</i> , 2013 Stamatos <i>et al.</i> , 2014 Werneburg <i>et al.</i> , 2016
Synaptic Cell Adhesion Molecule 1 (SynCAM 1)	NG-2 glial cells, oligodendrocyte precursor cells	<i>N</i> -glycans	Asn ¹¹⁶ in Ig1	ST8Sia II	Galuska <i>et al.</i> , 2010a Rollenhagen <i>et al.</i> , 2012
C-C chemokine receptor-7 (CCR-7)	Mature dendritic cells	<i>N</i> - / <i>O</i> -glycans	?	ST8Sia IV	Kiermaier <i>et al.</i> , 2016
E-selectin ligand-1 (ESL-1)	Microglia and human THP-1 macrophages	?	?	ST8Sia IV	Werneburg <i>et al.</i> , 2016

ST8Sia II	Almost ubiquitous in different cell types during development and adult (described next paragraph)	<i>N</i> -glycans	Asn ⁸⁹ , Asn ²¹⁹ (Asn ²³⁴)	ST8Sia II	Mühlenhoff <i>et al.</i> , 1996b Close and Colley, 1998 Close <i>et al.</i> , 2001
ST8Sia IV	Almost ubiquitous in different cell types during development and adult (described next paragraph)	<i>N</i> -glycans	Asn ⁷⁴	ST8Sia IV	Mühlenhoff <i>et al.</i> , 1996a Close and Colley, 1998 Close <i>et al.</i> , 2000
PSGP and hyosophorins	Oviduct mucin of various salmonid Hyosophorins on cortical alveoli of salmonid eggs Ovarian fluids of salmonids Sea urchins	<i>O</i> -glycans	?	?	Inoue and Iwasaki, 1978 Iwasaki and Inoue, 1985 Kitajima <i>et al.</i> , 1986 Inoue <i>et al.</i> , 1988 Kitajima <i>et al.</i> , 1988 Sato <i>et al.</i> , 1993 Sato, 2012 Kitazume <i>et al.</i> , 1994 Miyata <i>et al.</i> , 2006

B. POLYSIALYLATION : ACCEPTORS AND BIOLOGICAL ROLES

In the biosphere, polySia is distributed in many organisms like bacteria, vertebrates, echinoderms such as sea urchins (Kitazume *et al.*, 1994 ; Miyata *et al.*, 2004 ; Miyata *et al.*, 2006).

i. *PolySia in bacteria*

Among their membrane polysaccharides, bacteria as *E. coli* and *N. meningitidis* display long polySia chains, named colominic acid, with DP = 100-200 of Neu5Ac residues α 2,8- and/or α 2,9-linked between them (Sato and Kitajima, 2013). These polyNeu5Ac chains are essential for pathogenesis and virulence, allowing bacteria to evade the immune response of the mammalian host which recognizes it as “self” (Colley *et al.*, 2014 ; Villanueva-Cabello *et al.*, 2022).

ii. *PolySia in human and vertebrates*

Firstly identified in rat brain (Finne *et al.*, 1983), polyNeu5Ac found on the neural Cell Adhesion Molecule (NCAM, also named CD56) has been intensively studied in the mammalian central nervous system (CNS) where it plays an essential role during brain development with neuronal plasticity due to its repulsive properties (Schnaar *et al.*, 2014). Indeed, the three isoforms of NCAM of 120, 140 and 180 kDa could carry polySia on Asn⁴⁴⁹ and Asn⁴⁷⁸ in their Ig5 domain (Kleene and Schachner, 2004) (Table 5).

Due to its length and negative charges, polySia is surrounded by a large hydration shell which increases the hydrodynamic volume of its protein carriers. This leads to steric hindrance and modify protein function (Colley *et al.*, 2014). In contrast, PolySia-NCAM displays DP about 26 to 35 Sia residues (Galuska *et al.*, 2006 ; Mori *et al.*, 2017 ; Mori *et al.*, 2020) whereas higher DP are reported on bacteria and eukaryotic cells surface (Nakata and Troy, 2005) highlighting a large polySia heterogeneity. Polysialylation of NCAM limits homophilic *cis*- and *trans*- interactions on the same cell or between two cells, which prevent NCAM dimerization (Storms and Rutishauser 1998 ; Mori *et al.*, 2017). Through anti-adhesive properties due to steric encombrement (Rutishauser, 2008), polySia-NCAM promotes neurons migration and neuronal plasticity (Angata *et al.*, 2007) (Fig. 8A). Indeed, NCAM^{-/-} transgenic mice showed a decrease of 96.5% of brain

polySia expression in comparison to WT mice, confirming that NCAM is the major polySia acceptor in brain (Galuska *et al.*, 2010a).

In the immune system, polySias associated with neuropilin-2 (NRP-2) (Curreli *et al.*, 2007) and the chemokine receptor CCR-7 (Kiermaier *et al.*, 2016) (Table 5) are believed to play a role in the migration of dendritic cells. NCAM is also expressed on immune cells as monocytes, neutrophils and NK cells (Drake *et al.*, 2008 ; Moebius *et al.*, 2007; Stamatou *et al.*, 2014). Recently, polySias have drawn attention due to their attractive properties since these negatively charged polymers can bind and modulate the function of neurotransmitters and factors such as brain-derived neurotrophic factor (BDNF), dopamine and fibroblast growth factor (FGF2) in the CNS (Ono *et al.*, 2012 ; Hane *et al.*, 2015 ; Sato and Hane, 2018) and to protect these factors from proteolytic cleavages (Sato and Kitajima, 2021). Moreover, this attractive effect of polySia was also described during inflammation where neutrophils can generate Neutrophil Extracellular Traps (NETs) to capture pathogens (Galuska *et al.*, 2017c ; Zlatina and Galuska, 2019). These NETs are composed of antimicrobial peptides (AMPs) and histones toxic for pathogenic agents (Papayannopoulos, 2018) or which could enhance infectivity in case of SARS-CoV-2 (Hong *et al.*, 2022). NETs produced in excess become cytotoxic for the cells (Chen *et al.*, 2014). To be noted, nanoparticles with polySia inhibit the release of NETs (Bornhöfft *et al.*, 2019) and colominic acid could inhibit histones cytotoxicity *in vitro* interacting mainly with H2A and H2B (Galuska *et al.*, 2017c), suggesting importance of DP and charge of polySia for interactions.

PolySia glycoconjugates could be conveyed on cell surface through an unknown Ca^{2+} -dependent mechanism (Werneburg *et al.*, 2016 ; Thiesler *et al.*, 2021) and could be released in the extracellular medium through proteolysis and/or neuraminidase cleavages (Sumida *et al.*, 2015). PolyNeu5Ac chains released in the blood stream have a protective role of endogenous cells *in vivo* against histone-mediated cytotoxicity (Zlatina *et al.*, 2018) (Fig. 8B). In lung, soluble polysialylated NCAM has cytoprotective function binding nucleosomes and reducing NET cytotoxicity (Ulm *et al.*, 2013). PolySia could also bind lactoferrin and prevent excessive NETosis (Kühnle *et al.*, 2019a) through its lactoferrin domain, without loss of lactoferrin antimicrobial activity (Kühnle *et al.*,

2019b). Endoneuraminidase-N (endoN) treatment to remove polySia from cells lead to a drastic decrease of protection against histone cytotoxicity, confirming polySia role in the NETosis process (Ulm *et al.*, 2013).

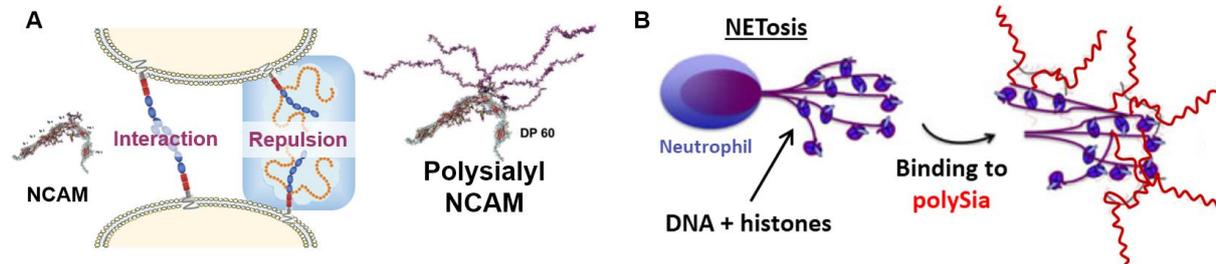


Figure 8: Repulsive and attractive properties of polySia. (A) In the nervous system, polysialylated NCAM promotes neuronal migration and plasticity through its repulsive mechanism. Inspired by Sato and Kitajima, 2021. (B) In the immune system, polySia could interact with charged biomolecules to modify their function, as the example of excessive histones cytotoxicity during NETosis process. Inspired by Galuska *et al.*, 2017b.

PolySia are mainly reported in organogenesis, cell growth, brain development and neuronal plasticity in the CNS and notably decrease around 70% 17 days after birth (Schnaar *et al.*, 2014). Interestingly, removal of polySia is correlated with a decrease of 25% in the distance between cells (Mindler *et al.*, 2021) while there is a huge heterogeneity of polySia chains length on cell surface (Drake *et al.*, 2008 ; Nakata and Troy, 2005).

Polysialic acid chains are also described in different organs such as in heart where they are involved in myocardium formation and neural innervation area (Delgado *et al.*, 2021). They are also described in testicles playing a role in cytodifferentiation and spermatogenesis regulation at adulthood and in placenta for cytotrophoblaste migration and adhesive interactions. They could be involved in the nutrient intake to the fetus during pregnancy (Hachem *et al.*, 2021 ; Galuska *et al.*, 2017a). NCAM positive cells like heart, dermal and smooth muscular cells, lung, liver, kidney, digestive and testis tissues present polysialylation (Galuska *et al.*, 2017a).

A few other polySia-career glycoproteins are reported in the literature (see Table 5). For example, E-selectin ligand-1 (ESL-1) and NRP-2 are polysialylated on their O-glycans on macrophages (Werneburg *et al.*, 2016) and Synaptic Cell Adhesion Molecule 1 (SynCAM-1) is polysialylated on its N-glycans on polydendrocytes (Rollenhagen, 2012). Essential for migration of dendritic cells, CCL-21 binds polySia with its C-terminal tail and could then interact with CCR-7

(Rey-Gallardo *et al.*, 2010 ; Kiermaier *et al.*, 2016). CD36 scavenger receptor from milk glycoprotein (Yabe *et al.*, 2003), α -subunit of the voltage Na⁺-sensitive channel in rat brain (Zuber *et al.*, 1992) and enzymes themselves, the two polysialyltransferases ST8Sia II and ST8Sia IV, can be polysialylated (Close and Colley, 1998). Indeed, there are not a lot of natural carriers identified for polySia (Table 5) and until now, only polyNeu5Ac is described in human (Janas and Janas, 2011 ; Davies and Varki, 2015), whereas more complex polySia are found in other vertebrate species, such as polyKdn in rat (Ziak *et al.*, 1996 ; Ziak *et al.*, 1999) and other polySia in fishes.

iii. PolySia in fishes

The nature, distribution and quantity of these polySia are very diverse and play extremely important roles depending on the fish species considered. Interestingly, the quantity and nature of polySia chains found in fish serum are more important and varied than in human serum (Zlatina *et al.*, 2018). As reported in rat brain, polySia on α -subunit of the voltage Na⁺-sensitive channel is also described in fish teleost like eel *Electrophorus electricus* (Fig. 9) (James and Agnew, 1989).

Inoue and Iwasaki have discovered 200 kDa glycoproteins, named hyosphorins, located in cortical alveoli of rainbow trout eggs (Iwasaki and Inoue, 1985). Salmonid fish eggs hyosphorins, also reported as polysialylated glycoproteins (PSGPs), are O-glycans with more than fifty percent of Sia by weight found on tandem repeats of a tridecapeptide. PSGPs display polySia chains on their O-glycans implicated in sperm acrosomal reaction and preventing polyspermy during fertilization (Taguchi *et al.*, 1994 ; Fliniaux *et al.*, 2022).

Four species of salmonids: *Oncorhynchus mykiss* (rainbow trout) (Fig. 9), *Oncorhynchus nerka adonis* (Jokanee), *Oncorhynchus masou ishikawai* (yamane) and *Salmo gairdneri* have diverse polySia and PSGP structures described (Kitajima *et al.*, 1986 ; Inoue *et al.*, 1988 ; Kitajima *et al.*, 1988 ; Sato *et al.*, 1993). In the rainbow trout *Oncorhynchus mykiss*, polySia chains are described with Neu5Gc, whereas in other salmonid species, a mixture of polyNeu5Gc, polyNeu5Ac and Kdn-rich glycoproteins have been described (Sato *et al.*, 1993). Kdn residues are mainly described as capped

at the non-reducing end of oligo-, polySia chains (Nadano *et al.*, 1986) although polySia containing many Kdn residues have been found in ovarian fluid (Sato *et al.*, 1993).

Other salmonids like *Salvelinus leucomaenis pluvius* (Japanese common char) and *Salvelinus namaycush* (lake trout) are described with specific polySia structures containing exclusively (Neu5Gc- α 2,8-Neu5Gc)_n repetitions or with alternations of (Neu5Gc- α 2,8-Neu5Gc)_n and (Neu5Ac- α 2,8-Neu5Ac)_n with an average of 6 Sia residues but variability between 2 and 24 residues by chains. Modifications are also reported on PSGP of unfertilized eggs of *Oncorhynchus nerka adonis* with O-acetylations on C-4, -7 and -9 positions of Neu5Gc and -9 position of Kdn (Sato, 2013) suggesting an important role during oogenesis (Iwasaki *et al.*, 1990). These polySias linked in α 2,6 to the inner GalNAc residue of O-glycans on PSGP provide a protective effect against pathogens in the reproductive system (Kitazume *et al.*, 1994) and are described to be resistant to bacterial exosialidases required for infection (Iwasaki *et al.*, 1984). Indeed, these polySia, associated on egg cells membranes could prevent from aquaculture pathogens infections such as oomycetes, saprolegnia, parasitica, fungi. In echinoderms, sea urchins display also complex polySia chains with α 2,5, α 2,8 and α 2,9-linked Sia and reported modifications like sulfation and hydroxylation (Miyata *et al.*, 2006).

This incredible polySia diversity observed in the various bony fish orders raises the question of the biosynthetic enzyme diversity and specificity. Previous phylogenetic studies conducted in the team indicated that functional diversity could result from distinct evolutionary genetic events: in the case of *Salmonidae*, the gain or loss of polysialyltransferase genes (Venuto *et al.*, 2020) and in the case of sea urchins, several duplication in tandem leading to a large multicopy family of polysialyltransferases with many variations associated (Harduin-Lepers *et al.*, 2008).

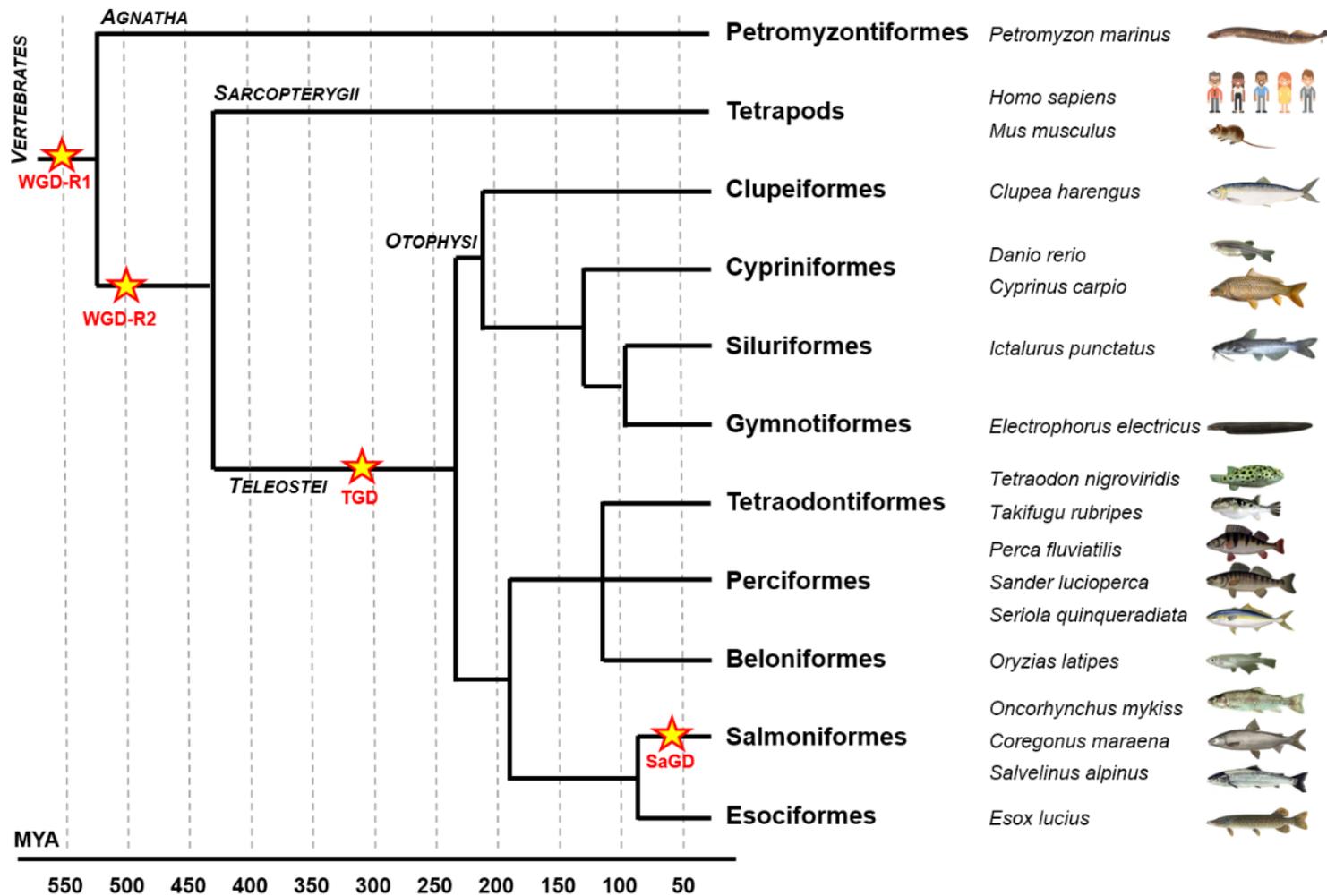


Figure 9: Phylogenetic tree highlighting the different teleost fish orders from vertebrates evolution. Whole genome duplication rounds R1 and R2 (WGD-R1 and WGD-R2), teleost-specific and salmonid-specific genome duplications rounds (TGD and SaGD) are represented to scale of millions of years ago (MYA). Agnatha and sarcopterygii are reported to anchor the phylogenetic tree after WGDR-1 and WGDR-2.

Inspired by Volff, 2014 and Venuto thesis, 2021.

C. POLYSIA EXPRESSION DYSREGULATED IN PATHOLOGY

PolySia expression are altered in mental disorders such as schizophrenia, depression, bipolar and epilepsy in parallel to variations and single nucleotide polymorphisms (SNPs) on the polySia-carrier or on the enzyme genes (Schnaar *et al.*, 2014 ; Thiesler *et al.*, 2022). PolySia are often described as oncodevelopmental antigens reexpressed in cancer with a role in tumour metastasis and immunosurveillance escape (Jarahian *et al.*, 2021 ; Thiesler *et al.*, 2022). As described in table 5, polySia is detected in different cancers such as neuroblastomas with a DP of 55 Sia residues (Livingston *et al.*, 1988 ; Hildebrandt *et al.*, 1998), colorectal cancers, small and non-small cell lung carcinoma (SCLC and NSCLC) (Komminoth *et al.*, 1991 ; Michalides *et al.*, 1994), Wilm's tumor (Zuber et Roth, 1990), glioma (Suzuki *et al.*, 2005), breast cancer (Soukhtehzari *et al.*, 2022) and are correlated with increase of metastasis (Elkashef *et al.*, 2016), decrease of immune responses and bad prognosis (Tanaka *et al.*, 2001 ; Seidenfaden *et al.*, 2003 ; Elkashef *et al.*, 2016). Interestingly, polySia sera levels are different between male and female and are elevated in the serum of systemic sclerosis patient (Tajik *et al.*, 2020 ; Villanueva-Cabello *et al.*, 2022), suggesting a sex-dependant role in the balance for immune response.

D. BIOTECHNOLOGICAL AND BIOTHERAPEUTICAL USE OF POLYSIA

PolySia have physico-chemical properties like water solubility, hydrophilicity, low viscosity, good biocompatibility and degradation *in vivo* which are really interesting for the development of biotechnological and biotherapeutical applications.

Many companies and groups of research are interested in polySia properties. First examples reported chemical polysialylation using colominic acid leading to successful drug delivery system with low immunogenicity (Gregoriadis *et al.*, 1993; Fernandes and Gregoriadis, 2001; Gregoriadis *et al.*, 2005). For invasive meningitis, candidate polysialylated vaccine was developed, optimized and tested in humans with modifications as propionylation and phenylacetylation to increase immune response against *N. meningitidis* group B (Krug *et al.*, 2012). PolySia derivatives on micelles were also developed for improvement of drug delivery systems (Zhang *et al.*, 2014). Another study

showed efficiency in the production of antibodies against *N. meningitidis* group C coupling α 2-9-oligoSia to glycolipids (Liao *et al.*, 2016).

Classical ways for glycoproteins production in bacteria and eukaryotic cells lead to facilitated productions, relatively stable, biodegradable and non-immunogenic glycoproteins. However, these glycoproteins have low hydrophilicity and low pharmacokinetics characteristics. Thus, polySia-based strategies were also investigated to modify characteristic and modulate pharmacokinetics of glycoconjugates. Chemical modification with Poly Ethylene Glycol (PEG) coupling (i.e. PEGylation), lead to better solubility, hydrophilicity, stability and half-life of proteins extended (Constantinou *et al.*, 2010). Nevertheless, PEGylation has limited uses as it may alter function of protein, increases the risk of side effects and toxicity in case of high dosages and long-term therapies, can be immunogenic and is expensive to produce (Bukowski *et al.*, 2002 ; Caliceti and Veronese, 2003 ; Jevsevar *et al.* 2010).

In comparison, enhanced strategy with chemical coupling by reductive amination or enzymatic addition of polySia improves glycoconjugates properties without alterations of functions (Gregoriadis *et al.*, 2005) and have to be developed for personalized therapeutics (Chen, 2014). Notably biotechnological companies as Xenetic Biosciences developed polyXen project, human preclinical and clinical studies to change the apparent hydrodynamic radius, improve stability, half-life and reducing clearance *in vivo* for pharmacological action of biomolecules such as Erythropoietin (EPO) for anemia, DNase I for cystic fibrosis and Factor VIII (FVIII) for haemophilia treatments (Zhang *et al.*, 2010 ; Meng *et al.*, 2019 ; Vorobiev *et al.*, 2013). Unlike PEG, polySia has the advantage to be biodegradable and metabolized in Sia units before its rapid excretion in the urine (Bader and Wardwell, 2014).

Neurodegenerative diseases such as Alzheimer, multiple sclerosis or age-related macular degeneration are often associated with an inflammatory origin, notably with proinflammatory cytokines, complement activation and exacerbated immune response. These last years, project named "PolySia avDP20" is based on the use of polySia with an average degree of polymerization of 20 Sia residues to prevent *in vitro* complement activation through anti-inflammatory effect

mediated by Siglec-11 interaction (Shahraz *et al.*, 2015 ; Shahraz *et al.*, 2022) and *in vivo* reducing disease symptoms of multiple sclerosis mouse model (Karlstetter *et al.*, 2017 ; Langmann and Fauser, 2017 ; Liao *et al.*, 2021). In CNS damage, polySia biomaterials developed and tested on mice shown a better axonal growth on damaged area, suggesting polySia efficiency in this context (Masand *et al.*, 2012). In other regenerative process after corticospinal injury, ST8Sia IV expression in mouse model allowed the recruitment of progenitor cells for axonal growth on the damaged area (El Maarouf *et al.*, 2006). PolySia could be use also to counteract NET associated cytotoxicity during inflammation (Fig. 10). Using colominic acid, *in vitro* and *in cellulo* experiments shown an inhibition of NETs cytotoxicity (Zlatina *et al.*, 2018).

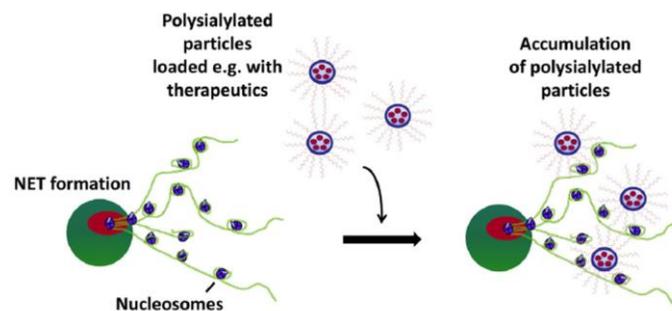


Figure 10: PolySia nanoparticles accumulated have a protective role against histone cytotoxicity released during NETosis. PolySia beads are able to bind and accumulate at histones/DNA-fibers via their polySia chains. Protective role of histones and polySia chains which form a complex.

Illustration from Galuska *et al.*, 2017c

Bacteria membrane, which are mainly composed of acidic phospholipids (PL) on their two sheets, are more sensitive to a targeted-release of histones coupled to polySia. Targeting bacteria after infection could cause less side effects and damages to human cells, which display neutral PL in extra / acidic PL in intra, and could be also investigated.

PolySia is described to be reexpressed or overexpressed in tissues during cancer progression and could serve as biomarkers for bad prognosis progression of cancer. In cancer therapy development, Creative Biolabs company has designed polysialylated derivatives glycoconjugate vaccines to overcome immune tolerance to polySia and facilitate the immune targeting of tumor cells. The chemically modified Sia precursor Man-NPr was used for biochemical engineering of cell surface with N-propyl polysialylation. Profacgen company has developed a highly selective

platform with nanomolar affinity to target polySia structures and positive cancerous cells and test drugs candidate for cancer therapy. In cancer therapies, two polySia-based strategies were developed recently using endoN coupled to oncolytic viruses to target and kill *in vitro* and *in vivo* polySia⁺ cancer cells (Martin *et al.*, 2018) or using an anti-polySia antibody coupled to a cytotoxic molecule which show an efficient targeting of polySia⁺ cells and decrease of their survival (Cox *et al.*, 2019). Through their StimuXen project, Lipoxen company developed a technology to protect leukocytes in chemotherapy patients. They demonstrated in pre-clinical trials that polysialylation on Granulocyte colony-stimulating factor (G-CSF) improved pharmacokinetics and pharmacodynamics of the factor in the bone marrow to produce more leukocytes and preserve the whole-body ability to fight infection. In summary, PolySia have many biotechnological and biotherapeutical interests and tend to be more and more explored, understood and controlled for development of new personalized strategies.

The major challenge that remains to be investigated is on an enzymatic point of view : how the machinery of glycosylation and particularly of sialylation is able to produce such a diversity of sialylated glycoconjugates and how these enzymes could regulate protein-carbohydrate interactions, initiation of sialylation and even elongation for polysialylation chains biosynthesis ?

V. SIALYLTRANSFERASES AND THEIR FUNCTIONAL BIODIVERSITY

The incredible biodiversity of glycoconjugates requires an equivalent diversity of biosynthetic enzymes known as glycosyltransferases. Among them, sialyltransferases (STs) (EC.2.4.3.x) are the last sequentially acting and inverting enzymes (Harduin-Lepers, 2010 ; Harduin-Lepers, 2023 ; Harduin-Lepers *et al.*, 2001 ; Harduin-Lepers *et al.*, 2005) involved in sialylation reaction, biosynthesis and bioregulation of sialylated glycoconjugates. STs catalyze the transfer of Sia onto matured glycans in the *trans*-Golgi compartment and *trans*-Golgi Network of cells. More precisely, STs are two substrates Leloir-type metal-independent enzymes that catalyze the stereo- and regio-specific transfer of sialic acid from the activated sugar donor Cytidine-MonoPhosphate- β -Sialic acid (CMP- β -Sia) to a large variety of glycolipid or glycoprotein acceptor substrates. The functional

organization of these enzymes in the Golgi apparatus is not completely understood yet and represents a current limitation to the study of one specific ST substrate specificity.

Indeed, there are 20 human STs described and grouped according to their activity and their modular nature in the family of glycosyltransferases GT29 in the Carbohydrate-Active enZymes (CAZy) database proposed by Dr B. Henrissat and collaborators which lists all the enzymes involved in carbohydrate metabolism (Coutinho and Henrissat, 1999 ; Cantarel *et al.* 2009, Lombard *et al.*, 2013 ; Drula *et al.*, 2022).

A. GENE ORGANIZATION AND REGULATION OF ST GENES EXPRESSION

Reported in the table 6, the twenty STs genes identified in the human genome (Harduin-Lepers, 2010 ; Harduin-Lepers, 2013 ; Harduin-Lepers *et al.*, 2001 ; Harduin-Lepers *et al.*, 2005) are grouped in four families: ST3GAL, ST6GAL, ST6GALNAC and ST8SIA (Harduin-Lepers, 2010). STs genes are polyexonic and widely dispersed in the genome on several chromosomes (Table 6). Distribution of exon/intron boundaries and exons sizes of ST genes is similar delineating groups of genes belonging to the same subfamily and family likely originating from a common ancestor (Harduin-Lepers, 2010 ; Harduin-Lepers *et al.*, 2001 ; Harduin-Lepers, 2023). Indeed, this gene architecture specific of each family is conserved in all vertebrates STs and is a parameter for identification of STs in vertebrate genomes (Harduin-Lepers *et al.*, 2005 ; Chang *et al.*, 2019). STs are mainly regulated at the transcriptional level by transcription factors such as Sp1, USF, NF-kB for *ST3GAL1* and *ST3GAL5* genes (Taniguchi *et al.*, 2001 ; Kim *et al.*, 2001 ; Higai *et al.*, 2006), or HNF-1, DBP, Sp1, Oct-1 for *ST6GAL1* gene (Svensson *et al.*, 1992 ; Xu *et al.*, 2003 ; Aasheim *et al.*, 1993 ; Wang *et al.*, 1993 ; Aas-Eng *et al.*, 1995 ; Lo and Lau, 1996), NF-kB for *ST8SIA1* gene (Bobowski *et al.*, 2013) or Sp1 for *ST8SIA2* and *ST8SIA4* (Yoshida *et al.*, 1996), although numerous other transcription factors could interfere with STs expression as described by GeneHancer Regulatory Element database (<https://www.genecards.org/Guide/GeneCard#enhancers>).

The expression of STs gene is highly regulated in a spatio-temporal manner (Kitagawa and Paulson, 1994a) and regulation takes place at different transcriptional and post-transcriptional levels, according to cell and tissue-type. They can be ubiquitously expressed or expressed

specifically during organogenesis or at specific developmental stages, under physiological and/or pathological states. Several isoforms of these enzymes may be expressed as a result of alternative exonic splicing events, use of multiple promoters and transcription factors. The most studied example is *ST6GAL1* gene which drives the expression of in several mRNA isoforms observed in different tissues (Taniguchi *et al.*, 2000). At least four main promoters are known for human *ST6GAL1* gene expression designated as P1, P2, P3 and P4 (Dall'Olio *et al.*, 2004 ; Dorsett *et al.*, 2021) where all are tissue-specific and physiopathological-dependent.

Epigenetic mechanisms can also regulate STs expression and is altered in pathologies. Particular, DNA methylation of the *ST6GAL1* gene (Dorsett *et al.*, 2021) is reported hypomethylated in lung cancer (Vojta *et al.*, 2016) or hypermethylated in glioblastoma (Kroes and Moskal, 2016) and bladder cancer (Antony *et al.*, 2014). MiRNAs, non-coding nucleic acids transcript also called microRNAs, regulate gene and mRNA expression of STs. In cancers, overall strategies are developed such as high-throughput fluorescence assay for miRNA interaction (miRFluR) with *ST6GAL1* and *ST6GAL2* genes and subdivided miRNAs in downregulators and upregulators (Jame-Chenarboo *et al.*, 2022).

Table 6: The 20 human sialyltransferase genes. HGNC : Hugo Gene Community Nomenclature. Chr : Chromosome, region and position of genes ; Exons : number of exons.

Inspired from Szabo thesis, 2017 and Vallejo-Ruiz thesis, 2001

Gene name	Other names	Chr (Gene ID)	HGNC	Exons
ST3GAL family				
ST3GAL1	SIAT4A, ST3O, SIATFL	8q24.22 (6482)	10862	13
ST3GAL2	SIAT4B	16q22.1 (6483)	10863	7
ST3GAL3	SIAT6, ST3N	1p34.1 (6487)	10866	26
ST3GAL4	SIAT4, SIAT4C, CGS23, FLJ11867, NANTA3, STZ	11q24.2 (6484)	10864	19
ST3GAL5	SIAT9, GM3 synthase	2p11.2 (8869)	10872	17
ST3GAL6	SIAT10	3q12.1 (10402)	18080	16
ST6GAL family				
ST6GAL1	SIAT1, CDw75, ST6N	3q27.3 (6480)	10860	9

ST6GAL2	SIAT2	2q12.3 (84620)	10861	14
ST6GALNAC family				
ST6GALNAC1	SIAT7A	17q25.1 (55808)	23614	12
ST6GALNAC2	SIAT7, SIAT7B, STHM	17q25.1 (10610)	10867	11
ST6GALNAC3	SIAT7C	1p31.1 (256435)	19343	11
ST6GALNAC4	SIAT7D, SIAT3C	9q34.11 (27090)	17846	6
ST6GALNAC5	SIAT7E, MGC3184	1p31.1 (81849)	19342	5
ST6GALNAC6	SIAT7F	9q34.11 (30815)	23364	10
ST8SIA family				
ST8SIA1	SIAT8, SIAT8A, GD3 synthase	12p12.1 (6489)	10869	5
ST8SIA2	SIAT8B, HsT19690, STX	15q26.1 (8128)	10870	7
ST8SIA3	SIAT8C	18q21.31 (51046)	14269	4
ST8SIA4	SIAT8D, PST, PST1	5q21.1 (7903)	10871	6
ST8SIA5	SIAT8E	18q21.1 (29906)	17827	9
ST8SIA6	SIAT8F	10p12.33 (338596)	23317	12

B. *ST* GENES EXPRESSION IN PHYSIOPATHOLOGY

i. *In physiology*

In humans, sialylation occurs in every tissue although sialyltransferases expression aren't ubiquitous for all. From data collected in all databases and normalized pipeline of The Human Protein Atlas, STs expression level was quantified in all tissues. The most studied *ST3GAL1* and *ST6GAL1* genes are ubiquitous and mainly expressed in heart, kidney thyroid and liver for *ST3GAL1* and highly expressed in liver for *ST6GAL1* whereas other ubiquitous enzymes like *ST3GAL2* or *ST8SIA1* are expressed at lower level. Some specific tissues are the site for higher STs expression such as in the brain for *ST3GAL5*, *ST6GALNAC6*, *ST8SIA2*, *ST8SIA3*, *ST8SIA4*, *ST8SIA5* and *ST8SIA6* in the reproductive system for *ST3GAL4*, *ST3GAL5*, *ST6GALNAC2* or in the digestive tissue mainly for *ST3GAL3* and *ST6GALNAC1* (Zhang *et al.*, 2020).

ii. *In cancers and diabetes*

Sialylation and STs protein expression level are often altered in pathological state as in cancers where hypersialylation is now recognized as a hallmark (Rodrigues and Macauley, 2018 ;

Hugonnet *et al.*, 2021 ; Munkley, 2022). Aberrant STs expression is closely correlated with cancer progression, angiogenesis, metastasis and drug resistance. To be noted, most STs reported in tumor invasiveness of bladder, breast, ovarian, pancreatic cancers and leukemia are ST6Gal I, ST3Gal I, ST3Gal II, ST3Gal IV, ST6GalNAc I and ST6GalNAc V (Zhang *et al.*, 2020 ; Hugonnet *et al.*, 2021). Sialylation of the two carbohydrate antigens sLe^a and sLe^x results from overexpression of three enzymes involved in their biosynthesis : ST3Gal III, ST3Gal IV and ST3Gal VI. So far, ST6GAL I is the most overexpressed gene found in cancers and pathologies like colorectal cancer and metastasis, pediatric leukemia, gastric, breast, prostate, ovarian cancers and metastasis and type II diabetes (Park and Lee, 2013 ; Swindall *et al.*, 2013 ; Schultz *et al.*, 2013 ; Rudman *et al.*, 2019 ; Schultz *et al.*, 2016 ; Wei *et al.*, 2016 ; Dorsett *et al.*, 2021 ; Garnham *et al.*, 2019). The two polysialyltransferases (polySTs) ST8SIA II and ST8SIA IV are overexpressed in cancers like neuroblastoma, SCLC and associated with metastasis (Falconer *et al.*, 2012 ; Guo *et al.*, 2020).

iii. In genetic and mental diseases

Deficiencies of STs expression could lead to CDG with severe pathologies, particularly two enzymes involved in gangliosides biosynthesis, *ST3GAL3* and *ST3GAL5*. Variants of *ST3GAL5* genes are described in patients suffering from pathologies like respiratory chain dysfunction and cognitive impairment (Indellicato *et al.*, 2019), in salt & pepper syndrome (Boccutto *et al.*, 2014) as well as Rett syndrome-like phenotype (Lee *et al.*, 2016) associated with mental disorders, epilepsy, visual deficiencies, intellectual disability, choreoathetosis and deafness (Gordon-Lipkin *et al.*, 2018 ; Inamori and Inokuchi, 2022). *ST3GAL3* deficiency is correlated with West syndrome characterized by disorders in neural development and epilepsy (Edvardson *et al.*, 2013 ; van Diepen *et al.*, 2018 ; Khamirani *et al.*, 2021). Furthermore, SNPs are reported in polySTs *ST8SIA4* and *ST8SIA2* associated to mental disorders like schizophrenia and bipolar disorders (Schnaar *et al.*, 2014). Independent studies from Japan, China and Australian cohort has shown that variations of *ST8SIA2* gene are reported in schizophrenic patients (Arai *et al.*, 2006 ; Tao *et al.*, 2007 ; McAuley *et al.*, 2012). Interestingly in *ST8Sia II* sequence, mutation reported in a schizophrenic patient leads to reduced *ST8Sia II* activity and the production of shorter polySia chains (Isomura *et al.*, 2011).

All these data converge towards the fact that STs and aberrant or deficient sialylation represent great biomarkers (Munkley and Scott, 2019 ; Zhang *et al.*, 2020) and targets for personalized therapeutics and treatment.

C. PROTEIN ORGANIZATION OF STs

STs adopt a type II topology (Paulson and Colley, 1989) in the Golgi apparatus, i.e. a short cytoplasmic tail on the N-terminal side about a dozen of amino acids, a single transmembrane domain (TMD) of about twenty amino acids with specific Leu and Phe hydrophobic residues for embedded in the Golgi membrane (Patel and Bajali, 2006), followed by a stem region of extremely variable length then a large catalytic domain of 250 to 300 aa exposed in the lumen of the Golgi apparatus (Harduin-Lepers, 2013). Some STs like ST6Gal I also exist as soluble forms and have been reported in colostrum, milk and plasma (Hudgin and Schachter, 1971 ; Paulson *et al.*, 1977 ; Paulson et Colley, 1989 ; Kaplan *et al.*, 1983) and BACE1 protease could be involved in STs cleavages as reported for ST6Gal I (Kitazume *et al.*, 2003 ; Kitazume *et al.*, 2005). Interestingly, these cleaved extracellular STs remain active for sialylation of exogenous glycoproteins (Sugimoto *et al.*, 2007) and in the systemic circulation (Lau, 2013). A study using cleaved ST6Gal I shows that this enzyme remains active to sialylate efficiently targeted cells (Lee *et al.*, 2014).

Although the functional organization of sialyltransferases in intracellular membranes is not well known, recent studies suggest that these enzymes function in complex with other glycosyltransferases as described for ST6GAL1 and B4GALT1 cooperation in the Golgi membranes (Hassinen *et al.*, 2010 ; Khoder-Agha *et al.*, 2019). In addition, Golgi environment such as pH, ions and redox homeostasis (Rivinoja *et al.*, 2009 ; Rivinoja *et al.*, 2012) influence the correct final sialylation status of glycoconjugates (Kellokumpu, 2019).

In addition, STs undergo a series of post-translational modifications (PTM) along their biosynthesis such as *O*- and *N*-glycosylation and disulfide-bond formation. These modifications influence their subcellular localization (Ma *et al.*, 1999) and trafficking in the cell (Harduin-Lepers *et al.*, 2001 ; Jeanneau *et al.*, 2004 ; Martina *et al.*, 1998 ; Bhide *et al.*, 2017a). Indeed, different *in vitro* and *in cellulo* studies demonstrated that these PTM modulate STs folding, stability, dimer

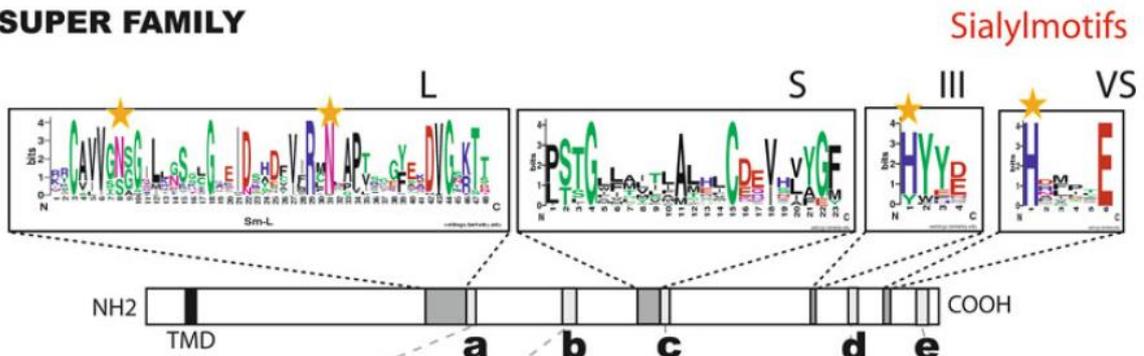
assembly and could have a significant impact on their enzymatic activity (Trombetta, 2003 ; Mühlenhoff *et al.*, 2001 ; Mikolajczyk *et al.*, 2020 ; Harduin-Lepers, 2023). To summarize, all these studies suggest a direct link between the structure, PTM and function of STs.

Despite relatively low sequence identities between the 20 human STs (Datta and Paulson, 1997 ; Paulson and Rademacher, 2009), all STs sequences are characterized by four conserved peptide motifs known as sialylmotifs (SMs) (Harduin-Lepers, 2010 ; Harduin-Lepers, 2013). Multiple sequence alignments (MSA) of vertebrate sialyltransferase sequences helped defining these conserved motifs which are described in all the GT29 sequences (Petit *et al.*, 2018).

These conserved motifs, namely the sialylmotifs Large (L), Short (S), III, Very Short (VS) are involved in binding of the ST substrates and catalytic function (Datta, 2009) (Fig. 11a):

- **SM L** is involved in interaction with donor substrate as proved by mutagenesis (Datta and Paulson, 1995). Composed by 48-49 aa, this motif has 5 aa strictly conserved during evolution (Livingston and Paulson, 1993) and conserved cystein involved in intramolecular disulfide bonds for optimal conformation of STs (Datta *et al.*, 2001 ; Qian *et al.*, 2001 ; Rao *et al.*, 2009).
- **SM S** is involved in interaction with acceptor substrate as proved by mutagenesis (Datta *et al.*, 1998). SM S is composed by 23 aa with Gly and Cys extremely conserved (Datta and Paulson, 1997) and the second cystein involved in intramolecular disulfide bonds with those in SM L.
- **SM III** is involved in catalytic activity for STs with 4 aa, notably a Tyr and a His conserved and involved in interaction with phosphate group of donor substrate (Jeanneau *et al.*, 2004).
- **SM VS** is involved in catalytic activity for STs with 4 aa, notably a Glu and a His well conserved (Geremia *et al.*, 1997 ; Jeanneau *et al.*, 2004) with the His considered as the catalytic base (Rao *et al.*, 2009).

a SUPER FAMILY



b FAMILY

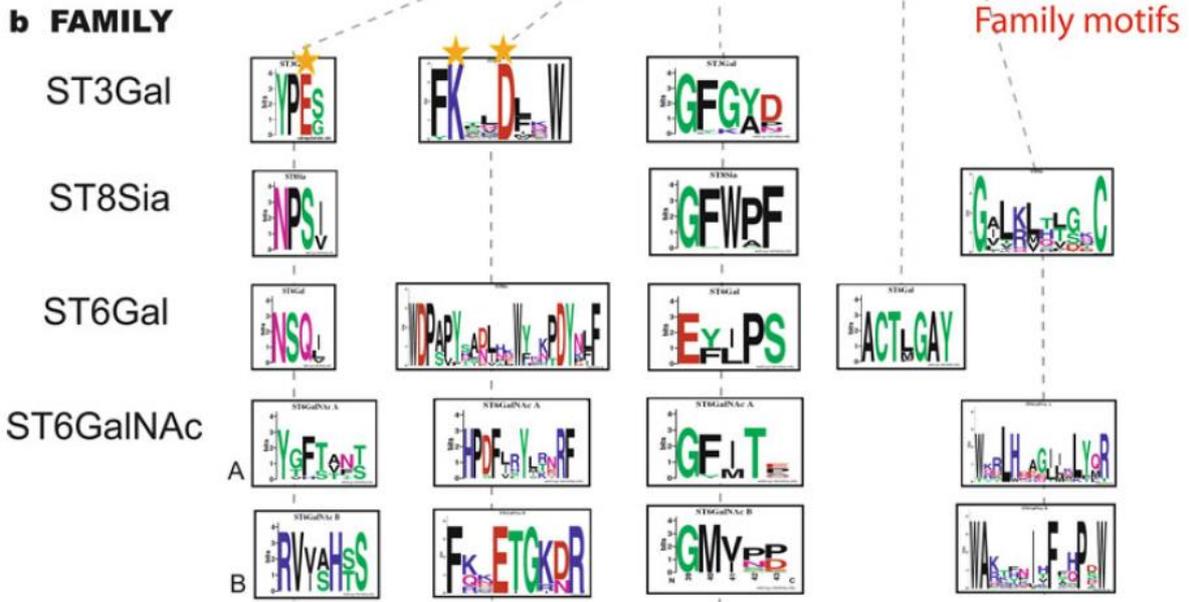


Figure 11: The two levels of amino acids conservation in STs. The 4 sialylmotifs L, S, III and VS are represented in the catalytic domain and characterized superfamily of STs. Specific motifs a, b, c, d and e are conserved in each STs family named family motifs. Yellow stars notify aa involved in interaction with substrates. Results from MSA with 42 ST3GAL, 13 ST6GAL, 23 ST6GALNAC and 37 ST8SIA vertebrate sequences.

Illustration from Harduin-Lepers *et al.*, 2013

Moreover, another level of conserved peptide motifs defines family-motifs of STs and 5 family motifs a, b, c, d, e were described recently (Harduin-Lepers, 2010 ; Patel and Bajali., 2006) (Fig. 11b). These family motifs are likely involved in enzymatic specificity defining each family ST3GAL, ST6GALNAC, ST6GAL or ST8SIA. In particular, YPE and NPS motifs for ST3GAL and ST8SIA families respectively are involved in acceptor substrate recognition nearby SM L (Okajima *et al.*, 1999a ; Fukumoto *et al.*, 1999; Angata et Fukuda, 2003). Other motifs are characteristics of families such as (Y/A)GF(K/G)(Y/A) for ST3GAL, (AN)YG(F/M) for ST6GAL or (IIL)(F/Y)GFWPF for ST8SIA (Fig. 12). Specificity Determining Positions (SDPs)

are critical amino acid positions likely involved in functional specificity in each subfamilies as described for ST3GAL and ST6GAL family (Petit *et al.*, 2015 ; Teppa *et al.*, 2016).

D. MECHANISM ACTION OF SIALYLATION REACTION

Firstly proposed by Horenstein, the mechanism for sialylation reaction is initiated by recognition of CMP-Neu5Ac by the catalytic domain of the STs (Horenstein and Bruner, 1996). These inverting enzymes reverse Sia anomery β from CMP-Sia to anomery α during transfer onto the acceptor monosaccharide. Interaction between enzymes, donor and acceptor substrates allow release of CMP by nucleophile substitution S_N2 to form an oxocarbenium transition-state molecule (Fig. 12) (Audry *et al.*, 2011 ; Montgomery *et al.*, 2017a). Indeed, sialylmotif III is located in the binding pocket of STs and particularly its His is involved in different steps with : the deprotonation of hydroxyl group of the acceptor, the interaction with electrophilic part of charged Sia for transition-state and the release of CMP from the donor substrate after Sia transfer, as describe for porcine ST3Gal I in interaction with CMP-Neu5Ac (Rakic *et al.*, 2013).

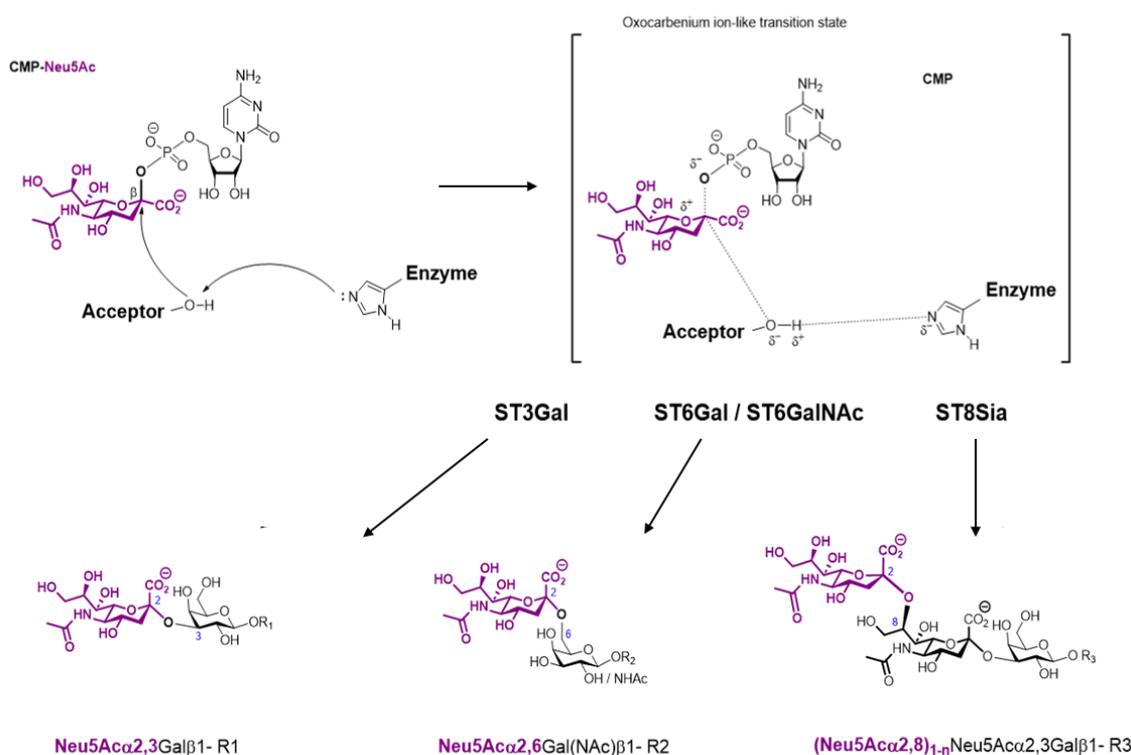


Figure 12: Sialylation reaction mechanism. Sialylation is divided in different steps with interactions between acceptor, donor and catalytic residue of STs for a S_N2 mechanism and inverting transfer of Sia.

Illustration from Harduin-Lepers, 2023

E. STs ENZYMATIC SPECIFICITY

STs are often and mistakenly considered as « one enzyme-one linkage » enzymes as hypothesized by Roseman with the concept that each glycosidic linkage is the product of a single enzyme (Rini *et al.*, 2009). Nevertheless, STs have to be considered as much more complex and broader in their enzymatic specificity. Biochemically, the 20 mouse and human STs are the most studied (Harduin-Lepers, 2008) and little is known about the activities of fish or amphibian enzymes. On one hand, human enzymes are very specific depending on the linkage they formed and the nature of the substrate they are sialylating. On the other hand, these enzymes also present functional redundancy *in vitro* and display an enzymatic promiscuity revealed by the study of their evolutionary history (Rohfritsch *et al.*, 2006). Indeed, they are classified according to their enzymatic specificities towards acceptor substrates, the monosaccharide acceptor and the type of glycosidic linkage formed:

- **ST3GAL** enzymes catalyzed the transfer of Sia forming α 2-3 linkage onto a Gal.
- **ST6GAL** enzymes catalyzed the transfer of Sia forming α 2-6 linkage onto a Gal (also on GalNAc despite their name).
- **ST6GALNAC** enzymes catalyzed the transfer of Sia forming α 2-6 linkage onto a GalNAc (also on GlcNAc despite their name).
- **ST8SIA** enzymes catalyzed the transfer of Sia forming α 2-8 linkage onto a Sia.

The glycosidic linkage with anomery α (Kolter and Sandhoff, 1997) is established between hydroxyl group, carried by anomeric carbon in C-2 of Sia and those in C-3/6/8/9 of the monosaccharide acceptor. The second level of enzymatic specificity concerns the acceptor structure itself where STs could recognize: aglycone moiety of glycoconjugate as for glycolipids, the underlying glycan and other linked-monosaccharides, or even polypeptide chains as shown for ST6GalNAc (Samyn-Petit *et al.*, 2000 ; Harduin-Lepers *et al.*, 2000) and for polysialyltransferases ST8Sia II and ST8Sia IV (Nakata *et al.*, 2006 ; Liao *et al.*, 2020) .

Some STs are highly specific towards their acceptor substrates like ST3Gal V which only generate G_{M3} ganglioside (Table 7). Otherwise, ST3Gal I and ST3Gal II catalyze the transfer of Sia

residue with the same α 2-3 linkage onto the same acceptor terminal Gal residue on Gal β 1-3GalNAc (type 3 disaccharide) and contribute to O-glycans biosynthesis (Kitagawa and Paulson, 1994b ; Kim *et al.*, 1996, Giordanengo *et al.*, 1997). Both enzymes have low *in vitro* activity on Gal β 1-3GlcNAc (type 1 disaccharide) and no *in vitro* activity on Gal β 1-4GlcNAc (type 2 disaccharide) (Kojima *et al.*, 1994). Indeed, ST3Gal I and ST3Gal II displays redundant enzymatic activities although their gene expression in tissues is different. *ST3GAL1* gene is mainly expressed in spleen and weakly in brain whereas these ratios are opposite for *ST3GAL2* gene as described in mouse (Kono *et al.*, 1997).

ST6GALNAC family members show differences in terms of specificity (Table 7): although ST6GalNAc I have an enzymatic specificity towards Tn antigen to synthesize sTn antigen *in vitro* and *in vivo* (Ikehara *et al.*, 1999a ; Kurosawa *et al.*, 2000 ; Marcos *et al.*, 2004), ST6GalNAc II could also synthesize sTn antigen *in vitro* but mainly sT antigen *in vivo* (Kurosawa *et al.*, 1996 ; Samyn-Petit *et al.*, 2000 ; Kono *et al.*, 2000). ST6GalNAc III and ST6GalNAc IV have an even more restricted specificity towards a few O-GP and glycolipids to generate $G_{D1\alpha}$ (Sjoberg *et al.*, 1996; Lee *et al.*, 1999 ; Harduin-Lepers *et al.*, 2000). In contrast, ST6GalNAc V and ST6GalNAc VI synthesize di-sLe^a on GlcNAc residues from gangliosides (Tsuchida *et al.*, 2003).

ST6Gal (ST6Gal I and ST6Gal II) catalyze the transfer of a Sia in α 2,6 bond on a Gal residue of the LacNAc structure (Gal β 1-4GlcNAc) (Fig. 13) but also in α 2,6 on a GalNAc residue of LacDiNAc (GalNAc β 1-4GlcNAc), in inverse activity ratios (Rohfritsch *et al.*, 2006).

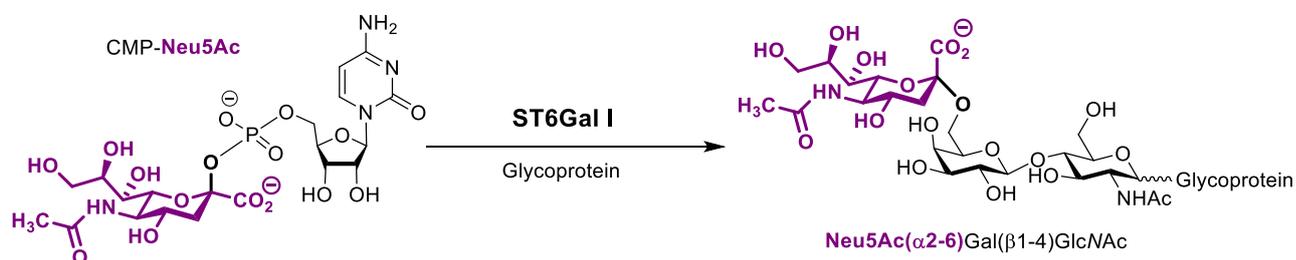


Figure 13: Representation of sialic acid transfer catalysis Neu5Ac (purple) from the activated donor substrate CMP-Neu5Ac by the human ST6Gal I enzyme onto a LacNAc (Gal β 1-4GlcNAc) motif of a glycoprotein.

Adapted from Noel *et al.*, 2018

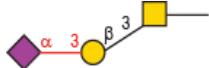
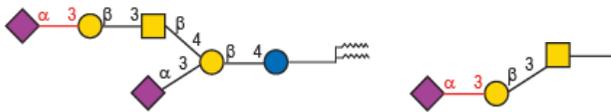
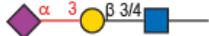
These enzymatic specificities can be explained by their molecular evolution from a common ancestor *st6gal1/2* which have been duplicated during vertebrate's evolution in two copies for acquisition of new functions and specificities (Teppa *et al.*, 2016).

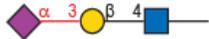
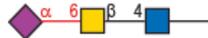
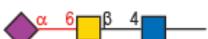
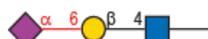
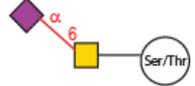
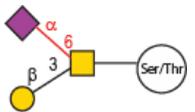
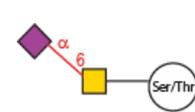
In contrast, some STs have more relaxed donor substrate specificity (Datta, 2009 ; Harduin-Lepers, 2010), although their donor substrate specificities are less explored (Harduin-Lepers, 2023). CMP-Neu5Gc, CMP-Neu5,9Ac₂ (Higa and Paulson, 1985), and CMP-Kdn (Angata *et al.*, 1998b) could be acceptable donor substrates by some mammalian sialyltransferases (Li and Chen, 2012).

For ST enzymatic activity, a minimal catalytic domain essential for STs activity is reported in the stem region and the catalytic domain between 70 and 90 aa before SM L for all STs (Vallejo-Ruiz *et al.*, 2001 ; Chen and Colley, 2000 ; Donadio *et al.*, 2003). Truncature in the stem region impacts STs enzymatic activity. Various deleted forms of the first 28, 35, 42 and 60 aa from N^{ter} ST6Gal I sequence (i.e. Δ28, Δ35, Δ42 and Δ60ST6Gal I) shows higher efficiency of Sia transfer towards transferrin, desialylated fetuin and orosomuroid compared to the full length ST6Gal I (Legaigneur *et al.*, 2001). In parallel, distinct kinetic parameters were found for two truncated ST6Gal I with a K_m for LacNAc 10-fold increased for Δ108ST6Gal-I (84 mM) as compared to Δ89ST6Gal-I (8.3 mM) (Luley-Goedl *et al.*, 2016). About ST3Gal I, different constructs like Δ25ST3Gal I and Δ56ST3Gal I led to efficiently active enzyme whereas Δ76ST3Gal I and Δ105ST3Gal I were not active (Vallejo-Ruiz *et al.*, 2001). These observations highlights that the stem region, extremely variable in terms of size, could influence enzyme stability, flexibility and may modulate acceptor recognition and specificity as described for ST6Gal I (Legaigneur *et al.*, 2001 ; Donadio *et al.*, 2003). Furthermore, mutagenesis in SMs and family motifs demonstrate various effects in ST6Gal I activity using surface plasmon resonance (SPR) (Laroy *et al.*, 2001) and for ST3Gal I activity (Jeanneau *et al.*, 2004).

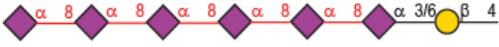
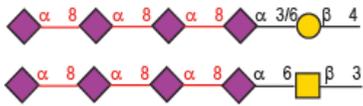
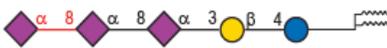
Finally, ST6Gal I could form dimers through its disulfide bound (Ma and Colley, 1996 ; Qian *et al.*, 2001) which impact its structure stability and its Golgi localization (Fenteany and Colley, 2005). This structural complexity, which may impact donor substrate recognition and enzymatic specificity, will be discussed in the next part about « 3D structure of STs ».

Table 7: Human protein sialyltransferases and their enzymatic specificities. Accession number for database UniProt ; AA quantities. GP : glycosylproteins / GL : glycolipids. Profile expression from The Human Protein Atlas. Inspired from Teintener-Lelièvre thesis, 2006 and Harduin-Lepers, 2023.

Enzyme	UniProt	EC nb	AA	Substrates specificity and structure formed	Main expression profile in tissue	References
ST3GAL family						
ST3Gal I	Q11201	2.4.3.2	340	<p>Neu5Acα2-3Galβ1-3GalNAc-Ser/Thr / O-GP & GL</p> 	Liver, kidney, skeletal muscle, spleen, pancreas, liver, placenta	Kitagawa and Paulson, 1994a
ST3Gal II	Q16842	2.4.3.4	350	<p>Neu5Acα2-3Galβ1-3GalNAc- / O-GP & GL</p> 	Skeletal muscle, spleen, heart	Kim <i>et al.</i> , 1996
ST3Gal III	Q11203	2.4.3.6	375	<p>Neu5Acα2-3Gal β -1,3/4-GlcNAc- / N-/O-GP & GL</p> 	Skeletal muscle, spleen	Kitagawa et Paulson, 1993
ST3Gal IV	Q11206	2.4.3.2	333	<p>Neu5Acα2-3Gal- β -1,4-GlcNAc- Neu5Acα2-3Gal- β -1,3-GalNAc- / N-/O-GP & GL</p> 	Placenta, ovaries, adrenal gland and testicles	Kitagawa and Paulson, 1994b
ST3Gal V	Q9UNP4	2.4.3.9	418	<p>Neu5Acα2-3Gal- β -1,4-Glc-Cer / GL G_{M3}</p> 	Ubiquitary	Ishii <i>et al.</i> , 1998

ST3Gal VI	Q9Y274	2.4.3.9	331	Neu5Ac α 2-3Gal- β 1,4-GlcNAc- / N-/O-GP & GL 	Liver, heart muscle, brain	Okajima <i>et al.</i> , 1999a
ST6GAL family						
ST6Gal I	P15907	2.4.3.1	406	Neu5Ac α 2-6Gal β 1-4GlcNAc- / N-GP  	Ubiquitary, liver mainly	Grundmann <i>et al.</i> , 1990
ST6Gal II	Q96JF0	2.4.3.1	529	Neu5Ac α 2-6Gal β 1-4GlcNAc- / N-GP  	Thyroid gland and Brain	Takashima <i>et al.</i> , 2002 Krzewinski-Recchi <i>et al.</i> , 2003
ST6GALNAC family						
ST6GalNAc I	Q9NSC7	2.4.3.3	600	(Neu5Ac α 2-3) ₀₋₁ (Gal β 1-3) ₀₋₁ Neu5Ac α 2-6GalNAc- Ser/Thr O-GP 	Digestive tissues : colon, intestine	Ikehara <i>et al.</i> , 1999a
ST6GalNAc II	Q9UJ37	2.4.3.3	374	Gal β 1-3[Neu5Ac α 2-6]GalNAc-Ser/Thr / O-GP  	Ubiquitary, mainly retina and skin	Samyn-Petit <i>et al.</i> , 2000
ST6GalNAc III	Q8NDV1	2.4.3.7	305	Neu5Ac α 2-3Gal β 1-3[Neu5Ac α 2-6]GalNAc- / O-GP & GL (G _{D1α})	Brain, thyroid gland and kidney	Tsuchida <i>et al.</i> , 2005

				Neu5Ac α 2-3Gal β 1-3[Neu5Ac α 2-6]GalNAc-		
				O-GP & GL (GD $_{1\alpha}$)		
ST6GalNAc IV	Q9H4F1	2.4.3.7	302		Ubiquitary, mainly pancreas	Harduin-Lepers <i>et al.</i> , 2000
				GL (GD $_{1\alpha}$) & N-/O-GP		
ST6GalNAc V	Q9BVH7	2.4.3.7	336		Brain, lung, smooth muscle, retina	Okajima <i>et al.</i> , 1999b ; Ikehara <i>et al.</i> , 1999b
				GL (GD $_{1\alpha}$ & All α -series gangliosides) & N-/O-GP		
ST6GalNAc VI	Q969X2	2.4.3.7	333		Ubiquitary, mainly brain, colon, ovary	Tsuchida <i>et al.</i> , 2003 Senda <i>et al.</i> , 2007
				ST8SIA family		
				Neu5Ac α 2-8Neu5Ac α 2-3Gal β 1-4Glc-Cer : GL (GD $_{3}$)		
ST8Sia I	Q92185	2.4.3.8	356		Brain, parathyroid gland, lymph node	Nara <i>et al.</i> , 1994 Sasaki <i>et al.</i> , 1994
ST8Sia II	Q92186	2.4.3.8	375	(Neu5Ac α 2-8) $_n$ Neu5Ac α 2-3/6-Gal β 1- / N-/O-GP	Brain, foetal kidney and heart	Scheidegger <i>et al.</i> , 1995

				 : polySia		
ST8Sia III	O43173	2.4.3.8	380	<p style="text-align: center;">Neu5Acα2-8Neu5Acα2-3Galβ1-4-GlcNAc- N-/O-GP & GL</p> 	Brain and foetal liver	Lee <i>et al.</i> , 1998
ST8Sia IV	Q92187	2.4.3.8	359	<p style="text-align: center;">(Neu5Acα2-8)_nNeu5Acα2-3/6Gal-β1- / N-/O-GP</p>  : polySia	Ubiquitary, mainly placenta, spleen, thymus, brain	Nakayama <i>et al.</i> , 1995
ST8Sia V	O15466	2.4.3.8	376	<p style="text-align: center;">GL (G_{T3}, G_{D1c}, G_{T1a}, G_{Q1b}, G_{T3})</p> 	Brain, adrenal gland	Kim <i>et al.</i> , 1997
ST8Sia VI	P61647	2.4.3.8	398	<p style="text-align: center;">Neu5Acα2-8Neu5Acα2-3(6)-Gal / O-GP & GL</p> 	Breast, brain and lung	Teinturier-Lelièvre <i>et al.</i> , 2005

F. 3D STRUCTURES OF SIALYLTRANSFERASES

i. Bacterial STs

Tertiary structure is essential to understand the molecular basis for donor and acceptor substrate specificities of STs as well as their catalysis. The first reported STs crystallization was for bacteria Cst-II from *Campylobacter jejuni* belonging to the GT-42 family (Chiu *et al.*, 2004). Cst-II is a multifunctional α 2-3/8-STs, different from their animal counterparts (GT-29) and able to synthesize various glycosidic linkages. Cst-II adopt a Rossman type for GT-A like conformation, or variant 2 GT-A fold with an alternated structure in sandwich $\alpha/\beta/\alpha$ folds (Mestrom *et al.*, 2019) (Fig. 14). Crystal structure of Cst-II provides key insights into the sialyl-transfer mechanism with identification of His₁₈₈ as the catalytic base. Mutation into Ala (i.e. Cst-II-H188A) impairs enzymatic activity where Cst-II-H188A is not able to catalyze Neu5Ac transfer onto natural acceptor substrate (Chan *et al.*, 2009).

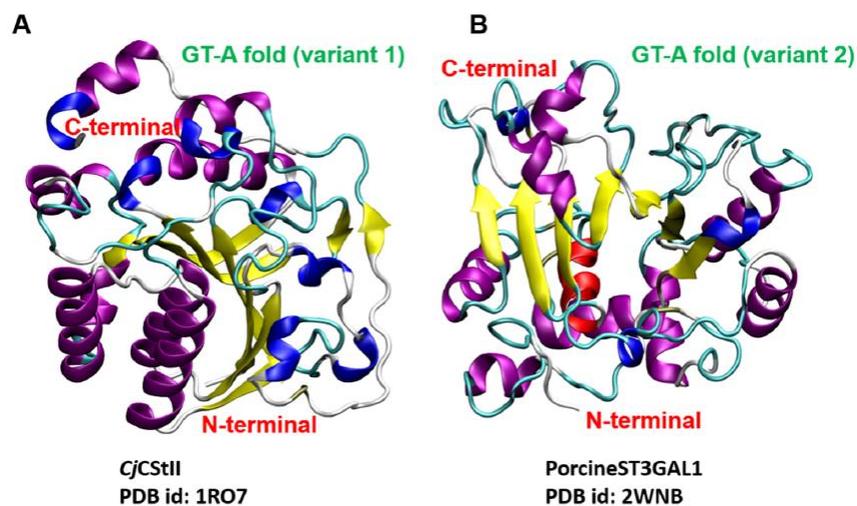


Figure 14: GT folds with (A) GT-A fold for Cst-II enzyme and (B) GT-A-like fold for porcine ST3Gal I.

Illustration from Grewal *et al.*, 2021

ii. Mammalian STs

In parallel, eukaryotic GT-29 STs adopt also a GT-A like fold structure (Fig. 14), which differs from other GT-A fold glycosyltransferases since they do not have a DxD motif and do not require a divalent cation for their activity (Breton *et al.*, 2005 ; Harrus *et al.*, 2018). Computational and system biology approaches helped to predict donor specificity of GTs enzymes from diverse species

and enabled their deep evolutionary classification into 53 families according their structural features and fold (Taujale *et al.*, 2021). However, the divergent and atypical GT-A families such as GT-29 and GT-42 sialyltransferase families do not possess the canonical GT-A motifs and don't align well with other GT-A families (Taujale *et al.*, 2020).

The first 3D structure of mammalian STs was solved by X-ray crystallography in 2009 for the porcine ST3Gal I (Rao *et al.*, 2009) with a catalytic domain organized in alternation of α/β folds with 12 α -helices (ribbons) and 7 twisted β -strands (arrows) and a topology on 7612345 (Fig. 15). As illustrated on the cartoon figure 17, a correlation between α/β folds and sialylmotifs has been established. Particularly, SM L (in green, Fig.15B) encompasses 4 α -helices and 4 β -strands of ST3Gal I to forms part of the donor binding site. SM S (in orange, Fig.15B) comprises helix α 11 and strand β 6 close to C3 disulfide bond (Fig. 15A) and is reported to be involved in both acceptor and donor substrates recognition as shown by site-directed mutagenesis (Datta *et al.*, 1998). SM III (in red, Fig. 15B) is located near to the phosphate-binding site and contains His³⁰² which stabilizes ribose ring of CMP from the donor substrate. Finally, SM VS (in purple, Fig. 15B) is composed by helix α 12 and a loop region containing the catalytic base His³¹⁹, essential for enzymatic activity. Between SM III and SM VS, a segment of 12 residues (305-316) corresponds to the family motif “d” (Fig. 11) and is described as a disordered “lid” domain (magenta dashed line in Fig. 15) and a flexible loop that could be involved in the donor binding substrate mode.

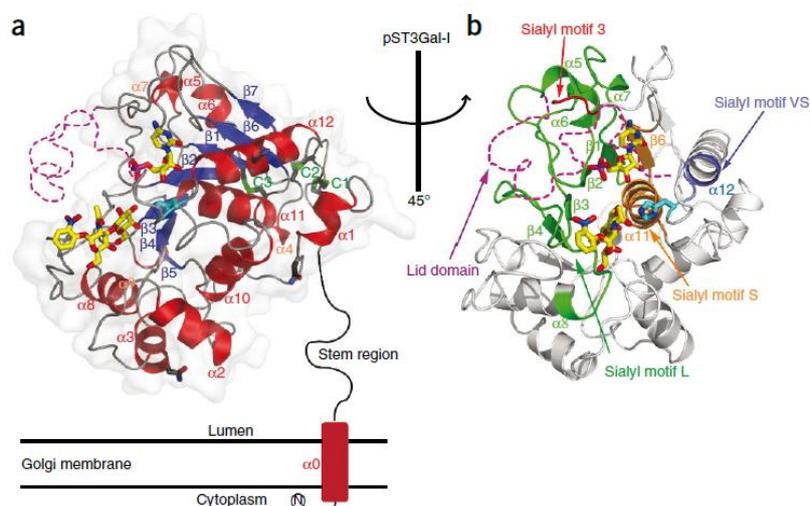


Figure 15: 3D structure of the catalytic domain of porcine ST3Gal I. (a) N^{ter} side is located in the cytoplasm, transmembrane domain is embed in the Golgi membrane with the helix α 0. The stem region

could be the site of proteolytic cleavage. On C^{ter} side, catalytic domain is represented with alternation of 12 α -helices (ribbons) and 7 twisted β -strands (arrows). (b) Representation with the 4 SMs L, S, III and VS.

Illustration from Rao *et al.*, 2009

Although donor binding site with this flexible loop seems highly conserved in GT-29 as shown for the docking of CMP-Neu5Ac with ST3Gal I and ST6Gal I (Rao *et al.*, 2009 ; Kuhn *et al.*, 2013), major variations occur in the acceptor site according the nature of enzymes, nature of its acceptor substrate (Sia, Gal, GalNAc) and linkage established. For GT-A fold variant 2 structure, catalysis in STs occurs in a close and active conformation that results from the movement of loops and/or domains induced from the initial binding of the CMP-Sia (Audry *et al.*, 2011). The flexibility and positioning of these loops confer substrate specificity and correctly orient the nucleophilic group (Harduin-Lepers, 2010 ; Harduin-Lepers, 2013 ; Moremen and Haltiwanger, 2019).

Furthemore, crystal structure of rat ST6Gal I analyzed with docking and molecular dynamics simulations highlights that helix α 12 with catalytic His acts as base to deprotonate hydroxyl group of the terminal Gal residue acceptor for nucleophilic attack on C2 of Neu5Ac donor (Meng *et al.*, 2013). Co-crystallization of human ST6Gal I with CMP (Kuhn *et al.*, 2013) confirms the S_N2 sialylation mechanisms with inversion of configuration proposed by Horenstein and Bruner.

iii. Homo- and heterodimerization of mammalian STs

To be noted, human ST6Gal I comprises 6 important Cys involved in intra- and intermolecular disulfide bonds to stabilize its structure and form homo- and heterodimers (Qian *et al.*, 2001 ; Hassinen *et al.*, 2010 ; Khoder-Agha *et al.*, 2019). In the TMD, Cys²⁴ is required for homodimerization whereas Cys¹⁸¹ and Cys³³² are respectively located in SM L and SM S and are important for localization in the *trans*-Golgi. Cys³⁵⁰ and Cys³⁶¹ are located in SM III and family motif “d” and are part of the active site, essential for the catalytic activity. Mutant of Cys¹³⁹ into Ser (i.e. C139S) is still active and more stably localized in Golgi as homodimer. Nevertheless, homodimerization of ST6Gal I reduces its ability to bind CMP-Neu5Ac (Qian *et al.*, 2001).

In parallel for ST6Gal I, relevant heterodimers were reported in complex with B4GALT1. Co-transfection with plasmid constructs of B4GALT1 and ST6Gal I and bimolecular fluorescence

complementation indicates heterodimerization occurring *in vivo* (Hassinen *et al.*, 2010). Indeed, ST6Gal I/B4GALT1 heterodimers are reported through lateral interactions via highly charged surface domains. ST6Gal I uses its non-catalytic surface for interaction with B4GALT1, leaving its active site exposed and accessible for donor and acceptor substrate binding. The two enzymes interact laterally via *cis*-configuration on the same side of the Golgi (Khoder-Agha *et al.*, 2019).

Similarly, crystallization of human ST8Sia III revealed a GT-A fold variant 2 structure with 7 parallel β -strands flanked by multiple helices. The general spatial arrangement for the 4 SMs are similar to those described for porcine ST3Gal I although the active site clef of ST8Sia III lies at the junction of conserved motifs SM S, SM VS and SM III (Volkers *et al.*, 2015). Moreover, dimerization of ST8Sia III has been reported in the crystals and reflects solution behavior (Fig. 16). Each monomer of ST8Sia III has its active site oriented facing the membrane, suggesting a role in the stabilization of the enzymes. In contrast, no dimerization was described for the two human polysialyltransferase ST8Sia II and ST8Sia IV (Volkers *et al.*, 2015).

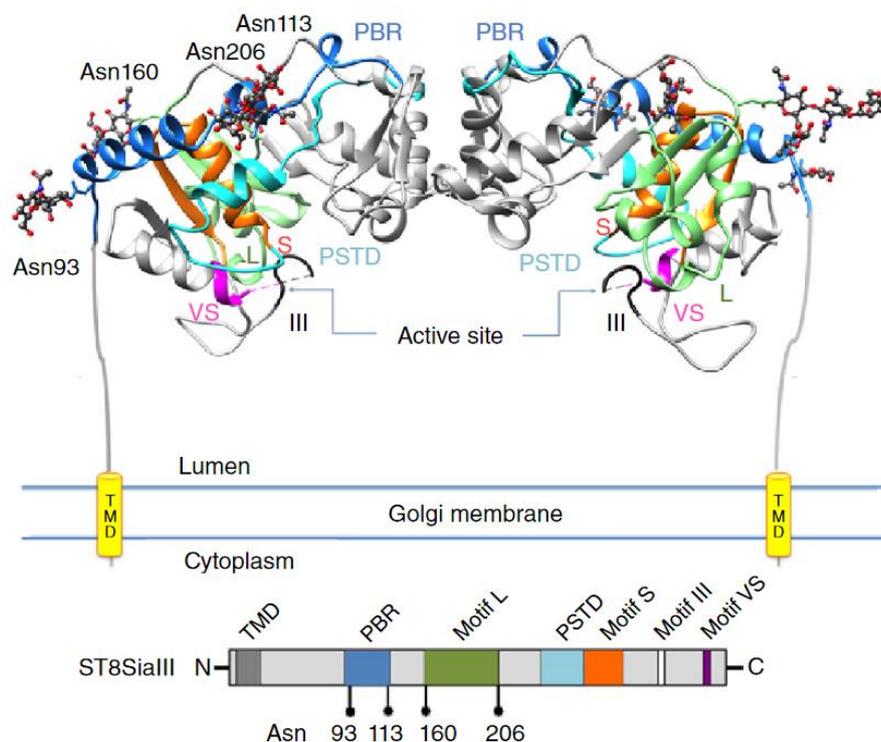


Figure 16: Overall structure and dimerization of human ST8Sia III. Each monomer is represented with its 4 sialylmotifs (L, S, III and VS), the 4 N-glycosylation sites oriented in surface and the two polysialyltransferase-like motifs (PBR-like and PSTD-like).

Illustration from Volkers *et al.*, 2015

ST8Sia III shares around 35% sequence identity with ST8Sia IV and has “PBR-like” and “PSTD-like” motifs (Fig. 16) partially aligned with PBR and PSTD motifs of human polysialyltransferases ST8Sia II and ST8Sia IV (Volkers *et al.*, 2015). Nevertheless, no effective polysialyltransferase activity were reported up to now for ST8Sia III, which is not able to polysialylate NCAM and itself (Angata *et al.*, 2000).

To summarize, crystal structures of GT-29 enzyme-substrate complexes (reported in Table 8) offer many advantages to understand how the sialylation mechanism takes place and how interaction and binding of substrates with specific domains of enzymes could influence their stereoselectivity and regioselectivity to generate such a wide diversity of sialylated glycoconjugates.

Table 8: Crystal structures of sialyltransferases in the Protein Data Bank (PDB). Ligands abbreviations are : 4-amino-1- β -D-ribofuranosyl-2(1H)-pyrimidinone (CTN) ; 2-(acetylamino)-2-deoxy-6-O-sulfo- β -D-glucopyranose (NGS) ; N-acetyl-2-deoxy-2-amino-Gal (A2G). Inspired by Harrus *et al.*, 2020.

Enzyme	Organism	AA range	3D-structure	Ligand substrates	References
ST3GAL family					
ST3Gal I	<i>Sus scrofa</i>	46 – 343	2WML	Unliganded	Rao <i>et al.</i> , 2009
			2WNB	A2G, CMP, Gal	
			2WNF	A2G, Gal	
ST6GAL family					
ST6Gal I	<i>Rattus norvegicus</i>	95 – 403	4MPS	GlcNAc	Meng <i>et al.</i> , 2013
			<i>Homo sapiens</i>	89 – 406	4JS1
	4JS2	β -Man, CMP, β -L-Fuc, Gal, Man, GlcNAc			Harrus <i>et al.</i> , 2020
	132 – 406	6QVS	Unliganded		
6QVT	CMP-Neu5Ac				
ST6GALNAC family					
ST6GalNAc II	<i>Homo sapiens</i>	66 – 374	6APJ	GlcNAc	Moremen <i>et al.</i> , 2018
			6APL	CMP, GlcNAc	
ST8SIA family					
ST8Sia III	<i>Homo sapiens</i>	81 – 380	5BO6	β -Man, CDP, α Fuc, GlcNAc	Volkers <i>et al.</i> , 2015
			5BO7	α Fuc, CTP, GlcNAc	
			5BO8	Man, α Fuc, β -Man, GlcNAc	
			5BO9	GS, α Fuc, Gal, CDP-3-F-Neu5Ac, GlcNAc, Sia	Not published
			5CXY	GlcNAc, β -Man, α -Fuc	

G. MOLECULAR PHYLOGENY AND EVOLUTION OF STs

Evolutionary biology is a science focused on the evolutionary processes that produced the diversity of life on Earth. Combining molecular phylogeny and evolution strategies enabled over the past 25 years to gain insights the biodiversity of sialylation machinery and molecular evolution of STs correlated with their enzymatic specificities in Metazoa.

Notably, recent breakthrough in this field led to the reconstruction of the sialylation pathway including GT-29 STs in the ancestor of eukaryotes (Petit *et al.*, 2018). Briefly, Last Eukaryota Common Ancestor (LECA) already possessed two ancestors of GT-29 sequences (i.e. ST6Gal/ST6GalNAc III-VI and ST8Sia/ST6GalNAc I-II/ST3Gal) probably inherited from an endosymbiosis event of α -proteobacteria (Petit *et al.*, 2018 ; Harduin-Lepers, 2023). A recent review summarizes glycomics profile in different taxa which could be a consequence of their direct evolution from LECA (West *et al.*, 2021). Interestingly, GT29 ST-related homolog sequence were identified in the genome of Eukaryota classes (SAR, Opisthokonta, Archeplastida, Chromalveolata), α -proteobacteria and in protist organisms (Petit *et al.*, 2018). During eukaryotes evolution, this sialylation pathway was partially maintained, with modification and molecular evolution, or totally lost like in the Streptophyta (Petit *et al.*, 2018). In plants, putative sialyltransferase and CMP-Sialic acid transporter genes were identified as well as traces of sialylation in *Arabidopsis thaliana* (Shah *et al.*, 2003). Until now, no direct evidences of enzymatic activity of sialylation were detected in plants (Séveno *et al.*, 2004). This GT-29 patchy distribution in Eukaryota raises the question of the evolutionary origin of the sialylation machinery and sialic acids as proposed below (Fig. 17) (Harduin-Lepers, 2023).

GT-29 STs are widespread in Deuterostoma and few sequences were described in invertebrate Deuterostomia like sea urchin and Protostoma. These GT-29 sequences share common conserved SM L, S, III and VS suggesting a common ancestor (Harduin-Lepers *et al.*, 2005 ; Teppa *et al.*, 2006 ; Petit *et al.*, 2018). For instance, arthropods like the insect *Drosophila melanogaster* only have one ST6GAL-like sequence, DSIAT (Koles *et al.*, 2004 ; Repnikova *et al.*, 2010) which represents an orthologue of the ancestor of the vertebrate's ST6Gal I and ST6Gal II (Petit *et al.*, 2010). Figure

17 illustrates the evolutionary scenario of three families ST3GAL, ST6GAL and ST8SIA studied these past years.

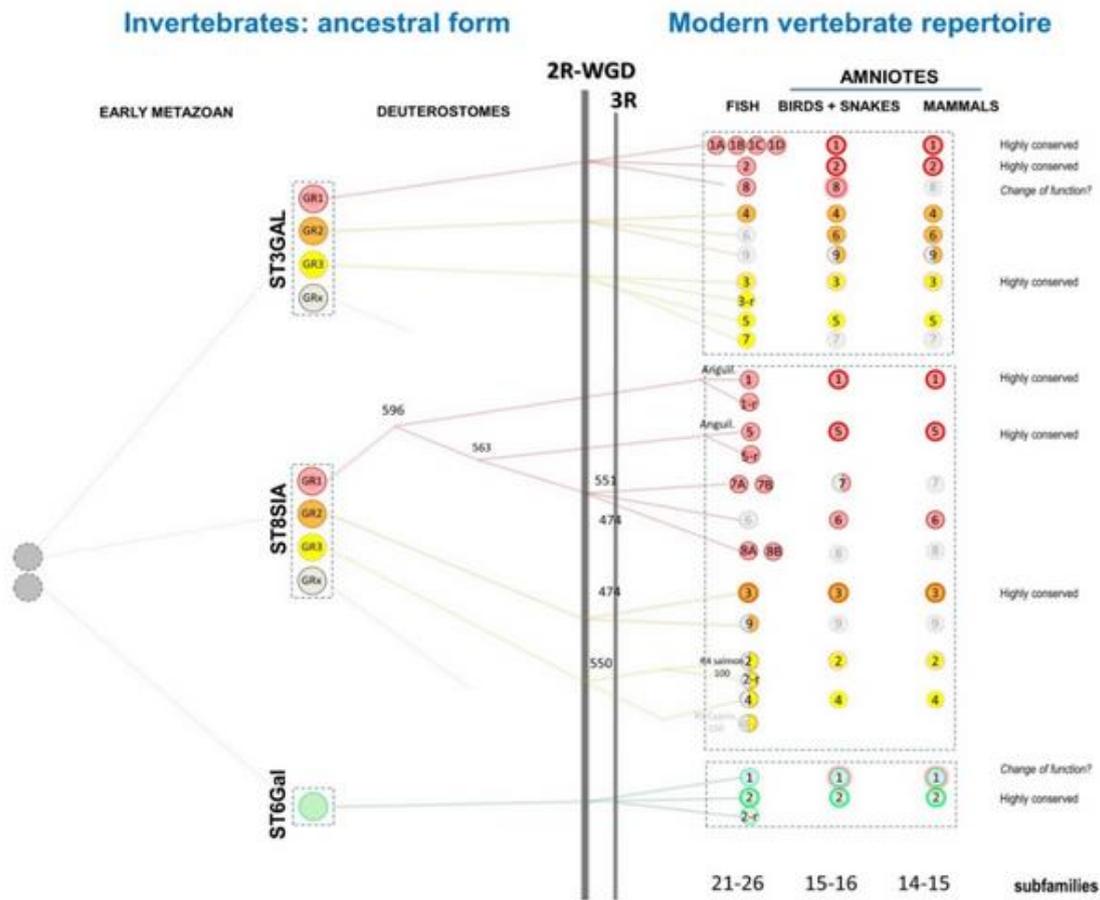


Figure 17: Illustration of the evolutionary scenario of GT-29 sialyltransferases in Metazoa. The two original groups of GT-29 sialyltransferases present in LECA gave rise to the four families ST3GAL, ST6GAL, ST6GALNAC and ST8SIA subdivided into groups of enzymes (GR) after gene duplications in Metazoa, which were maintained or not across evolution of Deuterostoma. Up to now, only ST8SIA, ST6GAL and ST3GAL evolutionary history could be reconstructed in Deuterostoma, therefore ST6GALNAC are not represented here. A burst of innovations occurred at the base of vertebrates after two rounds of whole genome duplication events (WGD-R1 and WGD-R2), after the Teleost-specific whole genome duplication event TGD and the species-specific genome duplication events (salmonid-specific 4R (Ss4R) and carp-specific 4R (Cs4R)) which were maintained or not in the various vertebrate branches.

Illustration from Harduin-Lepers, 2023

The functional diversification of sialyltransferases in vertebrates is explained by several rounds of whole genome duplication (WGD-R1 and R2) that took place around 555 and 500 million years ago (MYA) at the root of early vertebrates. A third WGD event took place around 320 MYA in Teleosts (bony fishes) and is considered as a teleost-specific WGD (TGD) (Fig. 17) (Petit *et al.*,

2010 ; Teppa *et al.*, 2016 ; Chang *et al.*, 2019 ; Venuto *et al.*, 2020). Another WGD event took place in Salmonidae (salmonid-specific WGD called SaGD), which occurred around 25-100 MYA. After these WGD events, duplicated genes could have diverse fates: they could be lost after accumulation of mutations. Notably, our team reported that *st3gal6*, *st3gal9* and *st8sia6* were lost in teleosts or *st3gal7* and *st8sia8* were lost in tetrapods (Fig. 18) and *st3gal8* lost in mammals likely through pseudogenization (Petit *et al.*, 2010 ; Teppa *et al.*, 2016 ; Chang *et al.*, 2019). Nevertheless, duplicated genes could be maintained in genomes and led to non-functional and/or functional enzymes with specific activity. Indeed, these enzymes could acquire new activity (neofunctionalization) or conserved ancestral gene function (subfunctionalization) (Fig. 18) (Teppa *et al.*, 2016). The mammalian ST6Gal I is an example of neofunctionalization and an accumulation of changes both in the catalytic domain and in the N-terminus part of the protein was observed during amniotes evolution (Petit *et al.*, 2010). Biochemical evidences showed that human ST6Gal family members (ST6Gal I and ST6Gal II) catalyze the transfer of a Sia in α 2,6 bond on a Gal residue of the LacNAc structure (Gal β 1-4GlcNAc) and also in α 2,6 on a residue of N-acetylgalactosamine of LacDiNAc (GalNAc β 1-4GlcNAc), in inverse activity ratios (Rohfritsch *et al.*, 2006). DSIAT, like the human ST6Gal II, has a preferential sialylation activity on LacDiNAc (Koles *et al.*, 2004) rather than on LacNAc. These distinct specificities could be explained by the molecular evolution of STs in vertebrates (Teppa *et al.* 2016) (Fig. 18) and it has been hypothesized the zebrafish ST6Gal I enzyme would show a different enzymatic specificity from that described for human ST6Gal I, although this remains to be demonstrated.

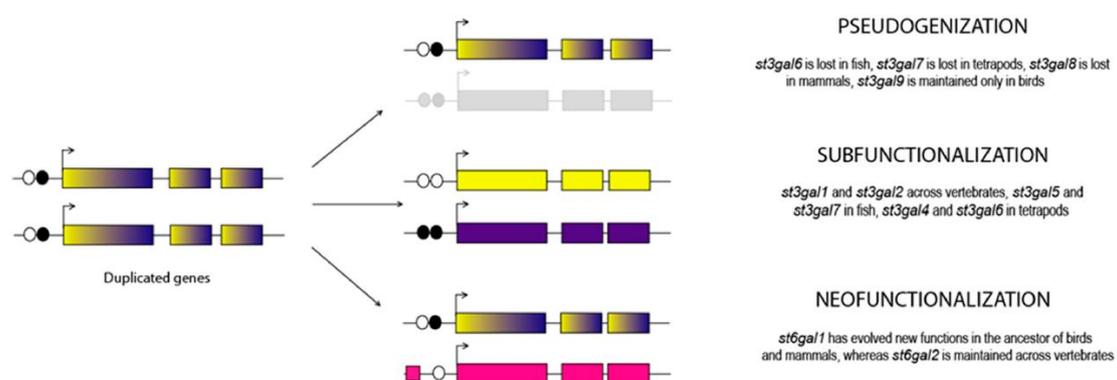


Figure 18: Evolutionary fate of ST gene duplicated after WGD events. Ancestral polyexonic ST genes duplicates could undergo diverse fates as loss of genes : **pseudogenization**, conservation of part of ancestral

gene function : **subfunctionalization**, or even acquire new function after accumulation of mutations or different regulation : **neofunctionalization**.

Illustration from Teppa *et al.*, 2016

Furthermore, structural studies have shown that Tyr₁₂₂ present in the human ST6Gal I is involved in LacNAc recognition motif (Kuhn *et al.*, 2013) whereas all vertebrates ST6Gal II display an His in this position (Teppa *et al.*, 2016). It was shown using specificity determining position (SDP) prediction, *in silico* mutagenesis from Tyr to His (Y122H) and docking studies that this change would allow the recognition of LacDiNAc motif (Teppa *et al.*, 2016) (Fig. 19). Therefore, the relationships between structure and function with the change of specific amino acids could modulate enzymatic specificity, although it needs to be enzymatically proven for vertebrates ST6Gal I as for zebrafish ST6Gal I.

Altogether, these studies have shown that these evolutionary events are likely responsible for the presence of 20 human and over 30 STs genes in the zebrafish *Danio rerio*.

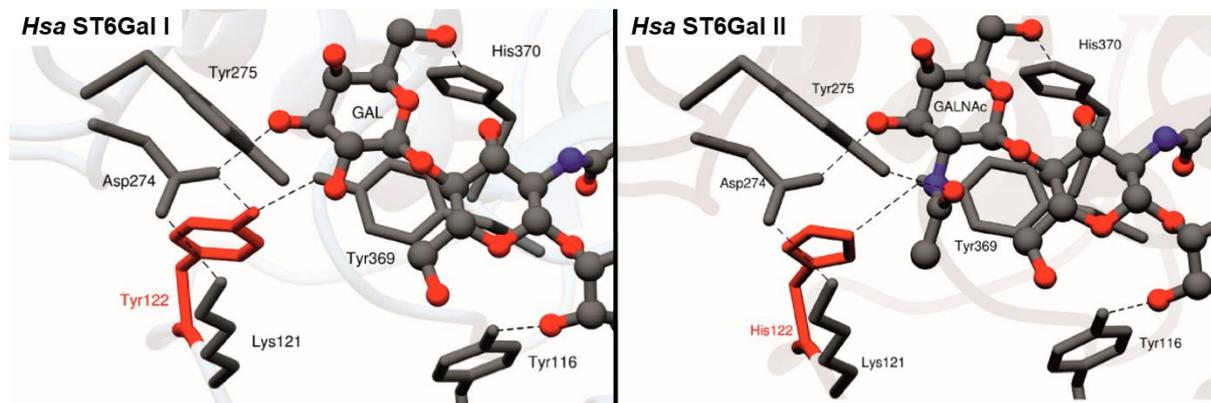


Figure 19: Crystallographic data of Tyr₁₂₂ and His₁₂₂ residues from the human ST6Gal I and ST6Gal II in interaction with Gal and GalNAc residues respectively.

Illustration from Teppa *et al.*, 2016

VI. POLYSIALYLTRANSFERASES

A. DESCRIPTION OF ST8SIA FAMILY

The ST8SIA family is comprised of 6 members in mammals known to be enzymatically active. In mammals, the ST8SIA enzymes catalyze the transfer of Sia onto other Sia residues with α 2-8 linkages onto glycoproteins or glycolipids. Phylogenetic analyses of Metazoan ST8SIA indicated that they can be classified into three subgroups according their enzymatic activity:

- **Mono- α 2-8-sialyltransferases** : ST8Sia I, ST8Sia V are responsible for gangliosides G_{D3} and G_{T3} biosynthesis respectively (Sasaki *et al.*, 1994; Nara *et al.*, 1994; Nakayama *et al.*, 1995 ; Kono *et al.*, 1996 ; Teinturier-Lelièvre *et al.*, 2005 ; Harduin-Lepers, 2010) although ST8Sia V could be active onto other glycolipids such as G_{M1b}, G_{D1a}, G_{T1b} and G_{Q1c}. Cloned and characterized in mouse and human, ST8Sia VI is involved on diSia formation onto O-glycosylproteins (Takashima *et al.*, 2002 ; Teinturier-Lelièvre *et al.*, 2005).
- **Oligo- α 2-8-sialyltransferase** : ST8Sia III catalyzes the transfer of one or many Sia onto glycoproteins and glycolipids to form oligoSia chains with DP = 3 to 7 (Lee *et al.*, 1998 ; Angata *et al.*, 2000 ; Harduin-Lepers, 2010) and could be responsible for polySia initiation (Sato *et al.*, 2001).
- **Poly- α 2-8-sialyltransferases (polySTs)** : ST8Sia II and ST8Sia IV are responsible for polySia chains formation on glycoproteins (Kojima *et al.*, 1996 ; Nakayama *et al.*, 1995 ; Angata *et al.*, 1998a) with notable differences in terms of tissues expression (Table 7), enzymatic specificities and DP sizes, with longer polySia generated for ST8Sia IV than ST8Sia II (Angata *et al.* 2002a).

B. EVOLUTION OF ST8SIA FAMILY

i. All ST8Sia enzymes

Phylogenetic distribution of α 2,8-sialyltransferases (ST8Sia) was assessed in Metazoa. In invertebrates, five ancestral *st8sia* genes are present (*st8sia2/4*, *st8sia3/9*, *st8sia1*, *st8sia5* and *st8sia6/7/8*) and nine *st8sia* gene subfamilies have arisen after the two rounds of WGD events, at the base of vertebrates (Fig. 20) (Harduin-Lepers, 2008 ; Harduin-Lepers, 2010 ; Venuto *et al.* 2020).

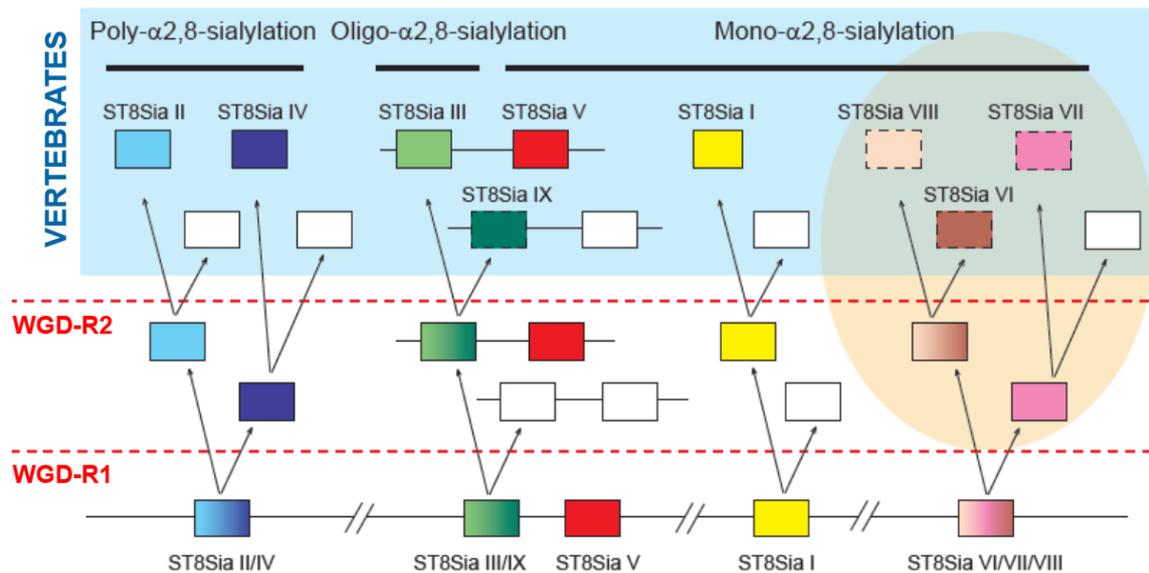


Figure 20: Distribution of ST8Sia genes during vertebrates evolution.

Graphical abstract of Venuto *et al.*, 2020b.

The evolutionary history of ST8Sia in vertebrates indicates multiple gene losses and/or gains during vertebrate evolution (Harduin-Lepers *et al.*, 2008). In fact, it was shown very recently that this was explained by a loss of the *st8sia6* gene in fish and the existence of a paralogous gene renamed *st8sia8* (or *st8sia6*-like) which disappeared in mammals (Chang *et al.*, 2019) and whose enzymatic activity needs to be characterized. Likewise, synteny analysis revealed the presence of *st8sia7* in *Danio rerio* (Fig. 20). Altogether these studies suggest the existence of new sialyltransferase activities in fish, explaining the extraordinary diversity of sialoglycans in these species.

ii. Polysialyltransferases

Molecular phylogeny of eukaryotic ST8Sia highlights the early emergence of polysialyltransferases genes in Deuterostomia around 750 MYA since ancestral sequences of polySTs were reported in invertebrates (the amphioxus *Branchiostoma floridae*) (Harduin-Lepers *et al.*, 2008). Particularly in vertebrates, WGD events played a crucial role in the evolution and functional diversity of ST8Sia. WGD-R1, WGD-R2, TGD and SaGD are followed by a rapid evolution of sequences and/or the loss of certain duplicated genes which catalyze the evolution of new adaptations (Braasch *et al.*, 2016), explaining divergence of vertebrate ST8Sia enzymes.

Present in Deuterostomia, ancestral *st8sia2/4* gene was segregated in two *st8sia2* and *st8sia4* genes after WGD-R1 (550 MYA) during the emergence of vertebrates (Harduin-Lepers *et al.*, 2008).

To be noted, the presence of one *stδsia4* gene in lamprey (*Petromyzon marinus*) (Fig. 21) could be explained by the loss of *stδsia2* gene after these duplications (Venuto *et al.*, 2020b). Then, tetrapods like *Homo sapiens* and *Mus musculus* present two polysialyltransferases *stδsia2* and *stδsia4* genes coding for ST8Sia II and ST8Sia IV active enzymes. Interestingly, there is a particular distribution of these polyST genes in fish teleost (Venuto *et al.*, 2020b) with generation of new paralogues resulting from specific events of gene and genome duplications and gene loss in the various orders of fish (Macquoen and Johnston, 2014).

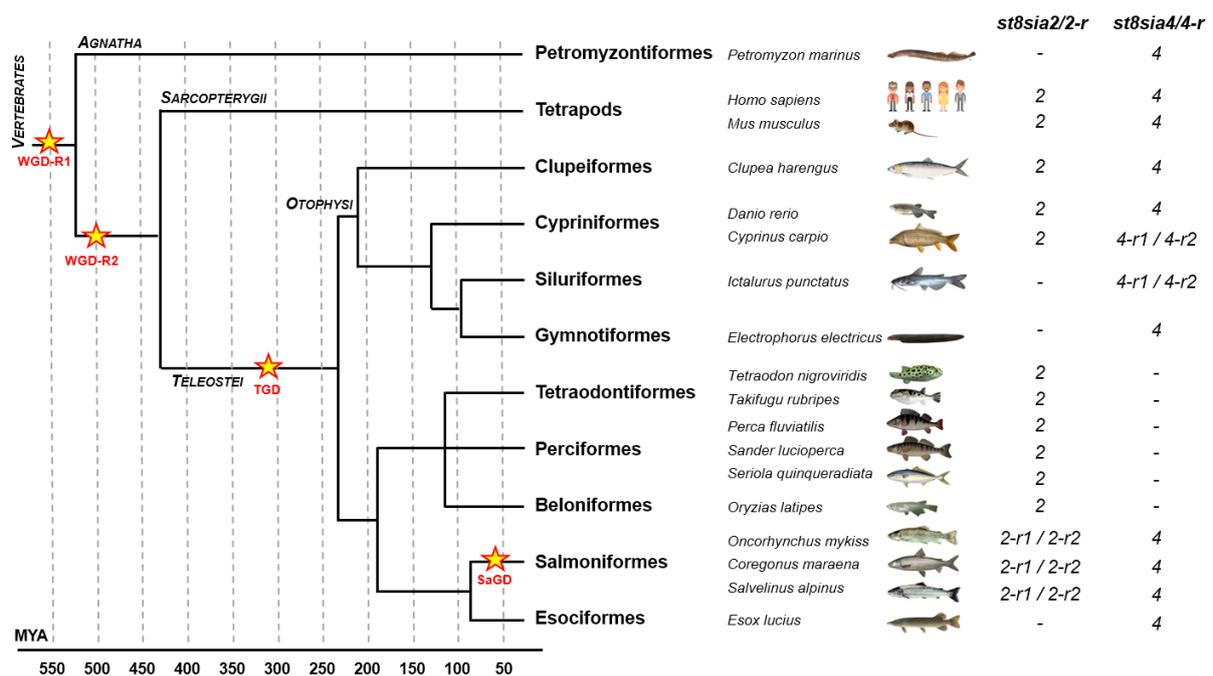


Figure 21: Distribution of polyST genes in vertebrates after WGD events. WGD-R1, WGD-R2, TGD and SaGD are specific duplication events which could explain the polyST gene distribution, particularly in teleost and salmonid fishes. Represented to scale of MYA. Inspired by Venuto *et al.*, 2020b.

Especially, the orthologue of the *stδsia4* gene was lost in perciformes, tetraodontiformes and beloniformes like the medaka *Oryzias latipes* and the orthologue of the *stδsia2* gene was lost in siluriformes like *Ictalurus punctatus* and esociformes like *Esox lucius* (Harduin-Lepers *et al.*, 2008; Venuto *et al.*, 2020b). Furthermore, a species-specific genome duplication event led to two copies of *stδsia4* gene in *Cyprinus carpio* and in *Ictalurus punctatus* whereas two copies of *stδsia2* gene are reported in all salmonids after SaGD like in *Oncorhynchus mykiss* and *Coregonus maraena* (Fig. 21). Synteny analysis of polySTs, i.e. the study of adjacent genes framing *stδsia2* and/or *stδsia4* genes onto a chromosome, showed a highly conserved *stδsia2* and *stδsia4* relationships for genes loci in

tetrapods and ancient fish. Particularly after SaGD, *st8sia2-r1* and *st8sia2-r2* genes share the same synteny in salmonids confirming a common origin (Venuto *et al.*, 2020b).

This distribution further suggested molecular evolution of these enzymes with potentially new functions for the remaining paralogue. These fish genes likely define new subfamilies with novel and as yet unknown enzymatic specificities (Venuto *et al.*, 2020a; Venuto *et al.*, 2020b).

C. ROLE OF VERTEBRATES POLYSTS

i. Biosynthesis of polySia

PolySTs are responsible for the biosynthesis of polySia as they catalyze the transfer of several Sia residues on sialylated glycoconjugates. In mammals, ST8Sia II and ST8Sia IV are the two polySTs responsible for the elongation of α 2,8-linked polyNeu5Ac chains (Close and Colley, 1998) (Table 7). These two enzymes are *trans*-Golgi type II enzymes and share around 59 % of sequence identity (Nakayama *et al.*, 1995 ; Livingston and Paulson, 1993 ; Harduin-Lepers *et al.*, 2008) and are notably very different to the bacterial polyST synthase like in *E. coli* K1 (Troy, 1992) or Cst-II from *C. jejuni* (Chiu *et al.*, 2004) which belong to the GT-42 CAZy family.

PolySia chains are different in size and properties depending on the enzymes which catalyze their synthesis. It was shown recently that human ST8Sia II and ST8Sia IV synthesize polySia with a DP = 32 and 35 Sias residues respectively (Mori *et al.*, 2017). Likewise, ST8Sia II catalyzes formation of approximately 40 Sias residues transferred on NCAM *in vitro*, lower than the 60 Sias transferred by ST8Sia IV *in vitro* (Angata *et al.*, 2000). Moreover, polySia-NCAM synthesized by ST8Sia II displayed a repulsive property toward other polySia-NCAM and an attractive field toward factors like BDNF and FGF2 whereas polySia-NCAM generated by ST8Sia IV exhibited only attractive properties toward polySia-NCAM, BDNF and FGF2 (Mori *et al.*, 2017). To be noted, ST8Sia II and ST8Sia IV could act in cooperation to generate longer DP polySia than independently *in vitro* (Angata and Fukuda, 2003). Consequently, these enzymes have a direct impact to modulate the size and function of polysialylated glycoconjugates.

ii. Expression in mammals

In mammals, polySTs have been highly studied in the brain (Schnaar *et al.*, 2014). The Human Protein Atlas reported an ubiquitous expression of *st8sia4* mRNA, mainly in spleen, bone marrow, lymphoid tissues and brain whereas *st8sia2* mRNA is more restricted to brain (Table 7). In mice embryos, the two polyST transcripts were barely detected at embryonic day E8 (Ong *et al.*, 1998). 21 days after birth, their mRNA expression decrease of 95% for *st8sia2* and 60% for *st8sia4* (Oltmann-Norden *et al.*, 2008) and only *st8sia4* transcript persist in the adult brain (Ong *et al.*, 1998 ; Eckhardt *et al.*, 2000). Double knockout (KO) *st8sia2*^{-/-} *st8sia4*^{-/-} in mice is lethal after 8 weeks, showing that polySTs and polySia are crucial for development (Weinhold *et al.*, 2005). Individual knockout (KO) mice on either *st8sia4* or *st8sia2* genes highlights polyST involvements in cognition, plasticity, learning memories (Markram *et al.*, 2007 ; Schnaar *et al.*, 2014). *St8sia2* KO led to 60% less polySia in comparison to WT whereas no significant change of polySia level was detected in *st8sia4* KO mice, suggesting a compensation by *st8sia2* *in vivo* (Oltmann-Norden *et al.*, 2008). To be noted, the presence of both polySTs *in vivo* is correlated with higher polySia chains (DP = 90) suggesting their cooperation during polySia formation (Galuska *et al.*, 2008).

As described in Chapter V-B-ii, *ST8SIA2* and *ST8SIA4* genes overexpression in humans are correlated with resistance to anti-cancerous drugs in myeloma (Ma *et al.*, 2015) but could be targeted to modulate tumor cell migration (Al-Saraireh *et al.*, 2013) and in metastatic cancer (Falconer *et al.*, 2012). Likewise in genetic disorders, SNPs of *ST8SIA2* gene are in linked with schizophrenia and neuropsychiatric disorders (Arai *et al.*, 2006).

iii. Expression in zebrafish

Zebrafish also exhibit *st8sia2* and *st8sia4* genes responsible for polysialylation in brain. Knockdown of *st8sia2* or *st8sia4* expression using morpholinos directly affect the axonal growth during CNS development in zebrafish (Langhauser *et al.*, 2011). To be noted, zebrafish ST8Sia II is catalytically active on NCAM *in vitro* and *in vivo* (Marx *et al.*, 2007). Interestingly, ST8Sia II is the principal responsible for polySia on NCAM throughout life as shown by *in situ* hybridization and *st8sia2* transcript is expressed in the CNS more ubiquitously than mammals. In contrast, *st8sia4*

transcript is weakly detectable and ST8Sia IV is not responsible for polySia on NCAM *in vivo* although this enzyme is catalytically able to produce polySia *in vitro* (Marx *et al.*, 2007).

While in mammals major polySia is driven by a strong expression of ST8Sia IV, polySia in zebrafish is associated with ST8Sia II and the role of ST8Sia IV remains to be elucidated.

iv. Specificity towards acceptor substrates

In contrast to most other GTs, polySTs seems to be highly protein-specific as only few polysialylated proteins have been identified in mammals and in vertebrates (Table 5).

First of all, NCAM is the most abundant cell adhesion molecules at cell surfaces and the most studied in the nervous system to be polysialylated (Mühlenhoff *et al.*, 1996a). There are three isoforms of NCAM (180, 140 and 120 kDa) which have 5 immunoglobulins group (Ig1 – Ig5) and 2 fibronectin domains (FN1-2). PolySia are attached on Ig5 NCAM to form PolySia-NCAM (Mühlenhoff *et al.*, 1998; Inoue et Inoue, 2001 ; Nakata et Troy, 2005). To be noted, Ig5-FN1 domains contain the docking site for polySia biosynthesis and elongation on NCAM (Mendiratta *et al.*, 2005 ; Close *et al.*, 2003).

Whereas NCAM could be polysialylated by both human polyST (Galuska *et al.*, 2006), only ST8Sia II could polysialylate SynCAM-1 *in vivo* through its N-glycan of Ig1 domain (Rollenhagen *et al.*, 2013) and only ST8Sia IV could polysialylate NRP-2 through its core-1 O-glycans (Rollenhagen *et al.*, 2012), suggesting differences of specificities for acceptor substrates recognition between these two polySTs (Mühlenhoff *et al.*, 2013). To be noted *in vitro*, SynCAM-1 could be polysialylated by ST8Sia II and ST8Sia IV (Galuska *et al.*, 2010a).

Interestingly about donor substrates, no enzymatic evidences were reported yet for the use of Neu5Gc and Neu5Ac directly by human ST8Sia II and ST8Sia IV to generate polyNeu5Ac and polyNeu5Gc or only indirect demonstrations as shown in SH-SY5Y human cells (Davies and Varki, 2015). The high specificity of vertebrate polyST towards acceptor substrates and the lack of studies with various donor substrates led biotechnological companies to use bacterial polyST for enzymatic addition of polyNeu5Ac on a large number of molecules for therapeutic use as reported in Chapter IV-D.

D. SEQUENCES OF MAMMALIAN POLYST AND CONSERVED MOTIFS

Sequence alignments of ST8Sia II and ST8Sia IV show a high conservation of the SMs L, S, III and VS and of conserved motifs that are unique to polySTs (Fig. 22): a polybasic domain (PBR, 35 aa) (Foley *et al.*, 2009 ; Bhide *et al.*, 2017b) and a polysialylation domain (PSTD, 32 aa) required respectively for the recognition of the protein acceptor and the catalytic activity (Bhide *et al.*, 2017b ; Nakata *et al.*, 2006). Catalytic His (i.e. His³⁴⁶ for ST8Sia II and His³³¹ for ST8Sia IV) are essential for polyST activity as shown by mutagenesis studies activity (Close *et al.*, 2003). Furthermore, basic aa located near to the PBR are involved in the substrate recognition: Arg⁸² of ST8Sia IV is important for NCAM and NRP-2 recognition whereas on ST8Sia II, Lys¹¹⁴ is important for SynCAM-1 recognition and Lys¹¹⁴ and Lys¹¹⁸ are involved in NCAM recognition (Mühlenhoff *et al.*, 2013).

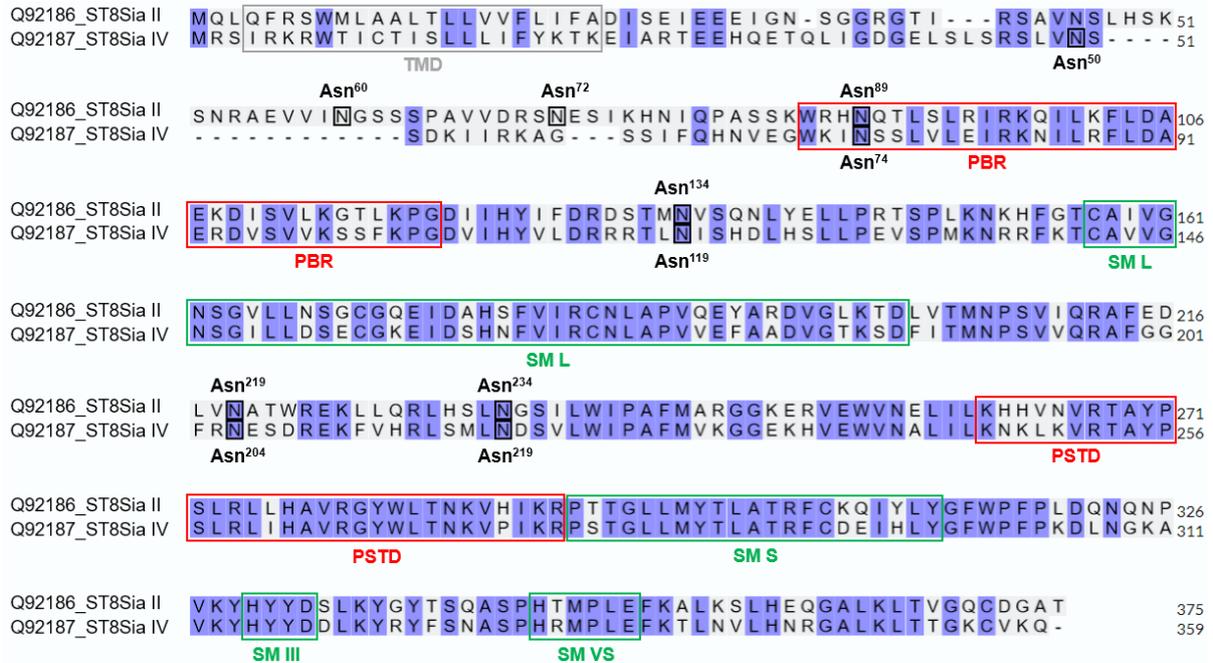


Figure 22: MSA of human polyST ST8Sia II and ST8Sia IV using UniProt ClustalW. Protein sequence identities are colored in blue. TMD is framed in blue, SMs L, S, III and VS are framed in green and PBR and PSTD motifs are framed in red. The 6 and 5 N-glycosylation sites for ST8Sia II and ST8Sia IV respectively are framed and reported in black. Accession number for human ST8Sia II: Q92186 and for human ST8Sia IV: Q92187.

i. Cysteines involved in polyST stability

Conserved Cys residues have been reported in polyST sequences: three Cys located in SM L, one Cys in SM S and one Cys located at C^{ter} position (Fig. 22) (Angata and Fukuda, 2003). Using mass spectrometry, Angata and collaborators show for ST8Sia IV that these Cys are involved in

intramolecular bonds : the Cys¹⁴² in SM L and Cys²⁹² in SM S form a disulfide bound and that the second Cys in SM L (Cys¹⁵⁶) and Cys³⁵⁶ at the C^{ter} form a second disulfide bridge (Angata *et al.*, 2001). To be noted, no homodimer was detected through ST8Sia IV Cys residues (Angata and Fukuda, 2003). Moreover, they demonstrated using site-directed mutagenesis that mutation on these Cys completely abolished enzymatic activity of ST8Sia IV, highlighting that these aa are crucial for the 3D structure and for the catalytic activity of ST8Sia IV (Angata *et al.*, 2001).

ii. PBR motif

PBR motif consists of 35 aa located near to the stem region in position 71–105 for ST8Sia IV and between position 80–120 for ST8Sia II (Nakata *et al.*, 2006 ; Foley *et al.*, 2009). Some aa located in PBR have a crucial role in acceptor-substrate recognition as they make specific contacts with NCAM. As reported in Table 9, replacement of aa residues like Arg⁹³ in ST8Sia IV or Arg⁹⁷ in ST8Sia II by Ala residues have higher impact on polySia on NCAM rather than on the enzyme themselves (i.e. autopolySia) (Bhide *et al.*, 2018).

Table 9: Impact of replacing ST8Sia IV and ST8Sia II PBR residues on substrate polysialylation. Higher significant impact of aa substitutions on enzymatic activity are reported in red.

Data retrieved from Bhide *et al.*, 2018

<i>PolySia with ST8Sia IV</i>	R72A	R82A	K83A	R87A	R93A	K99A	K103A	
NCAM	89	49	73	67	17	97	70	
NRP-2	78	9	89	72	95	12	107	
ST8Sia IV	88	47	88	109	98	24	98	
<i>PolySia with ST8Sia II</i>	R87A	R95A	R97A	K98A	K102A	K108A	K114A	K118A
NCAM	83	87	33	87	66	23	42	31
SynCAM-1	77	79	51	97	21	59	3	2
ST8Sia II	100	83	75	102	95	102	32	15

NMR studies show direct electrostatic interactions between FN1-NCAM and PBR-ST8Sia IV through an acidic patch composed by Arg⁸², Lys⁸³ and Arg⁹³ (Bhide *et al.*, 2017). Substitutions by Ala residues of Glu⁹² and Asp⁹⁴, aa surrounding Arg⁹³, largely inactivated ST8Sia IV (Foley *et al.*, 2009). Indeed, Glu⁹², Arg⁹³ and Asp⁹⁴ are located in an essential loop involved in acceptor substrate recognition (Zhou *et al.*, 2015). This loop begins with Phe⁸⁸ and Leu⁸⁹ essential for intramolecular

interaction with PSTD (Zhou and Huang, 2022). Lys¹⁰², Lys¹¹⁴, Lys¹¹⁸ in PBR are also required for interaction between polysialylated-ongoing acceptor and polySTs (Bhide *et al.*, 2018).

iii. PSTD motif

PSTD motif is composed of 32 aa behind SM S (Sevigny *et al.*, 1998 ; Nakata *et al.*, 2006) located in positions 246-277 for ST8Sia IV and 261-292 for ST8Sia II that are essential for elongation and polysialylation capacity of the enzymes (Nakata *et al.*, 2006 ; Huang *et al.*, 2015). Notably, three positively charged aa Arg²⁶⁵, Lys²⁷⁶ and Arg²⁷⁷ in C^{ter} of PSTD (Fig. 22) are crucial to initiate polySia process as shown by site-directed mutagenesis (Foley *et al.*, 2009).

Different PSTD mutants in ST8Sia IV such as R252L, K272I, I275K or R277I led to a reduction of at least 75% of polyST activity (Nakata *et al.*, 2006), indicating that Arg²⁵², Lys²⁷², Ile²⁷⁵ and Arg²⁷⁷ are required for catalytic activity onto NCAM. Likewise, mutants with switched basic residues like KR2RK (i.e. R276K and K277R switch) led to a drop of polyST activity suggested that not only number of basic aa residues in the PSTD important for polySia but also their position (Nakata *et al.*, 2006).

In parallel, NMR were used to study interaction of PSTD with CMP-Sia and polySia (Liao *et al.*, 2020). Three aa segments were reported in the PSTD: a short α -helix H1 from Val²⁵¹ to Ala²⁵⁴, a longer α -helix H2 from Ser²⁵⁷ to Asn²⁷¹ and finally a shorter C^{ter} loop of Ile²⁷⁵, Lys²⁷⁶ and Arg²⁷⁷. The largest chemical shift perturbation (CSP) was detected for H1-PSTD with polySia, CMP-Sia and Sia whereas H2 shows also high CSP for PSTD-CMP-Sia and PSTD-polySia, consistent with a conformational change in H2 during interactions. Interestingly, some non-basic residues between Val²⁵¹ and Val²⁷³ are required for polySia on NCAM, suggesting that enzymatic activity is not only dependent of basicity in PSTD (Liao *et al.*, 2020).

NMR and Wenxiang Diagram approaches were also developed for PSTD graphical study with distribution of its hydrophobic and hydrophilic aa residues on α -helix structure. These approaches helped to determine the key aa involved in intra- and intermolecular interactions, trying to understand how polysialylation reaction is processed. Notably in ST8Sia IV, Val²⁶⁴, Tyr²⁶⁷, Trp²⁶⁸

are the three key aa for PSTD-polySia interaction whereas Ala²⁶³ and Tyr²⁶⁷ are crucial residues for the PBR-PSTD intramolecular interaction. (Zhou and Huang, 2022).

Finally, based on NCAM-polySTs interaction through PBR and PSTD motifs, Troy and Colley teams made hypothesis of a first protein-protein interaction and after initiation of polySia, a progressive switch leading to a protein-glycan interaction required for polySia elongation observations (Colley *et al.*, 2014 ; Bhide *et al.*, 2018). The molecular mechanism for PolySia reaction isn't completely understood. A concept for polySTs activity was summarized recently (Fig. 23) (Galuska *et al.*, 2017) gathering different hypotheses and evidences previously collected.

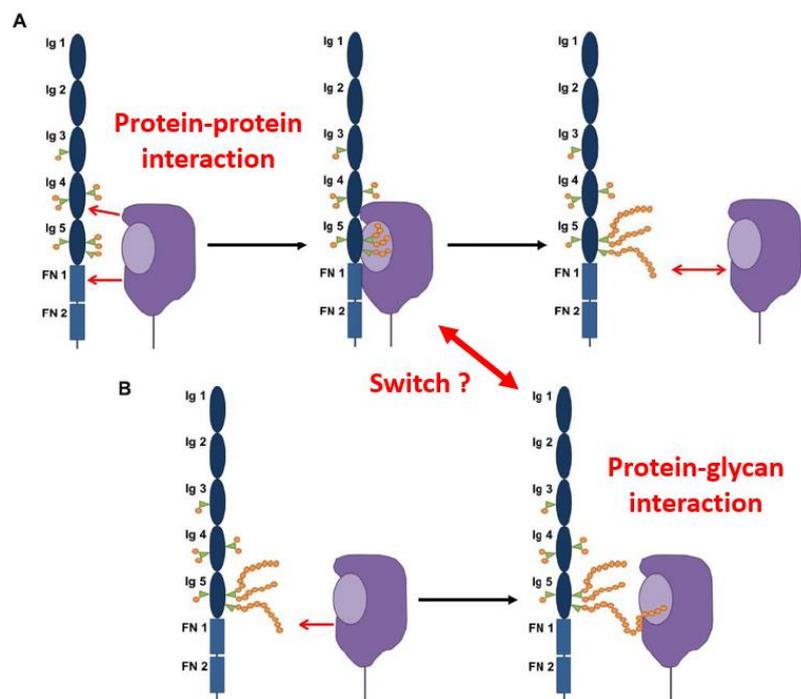


Figure 23: Concept of polysialylation: from protein-protein interaction switching to protein-glycan interaction. The initial interaction between polyST through its basic PBR region and NCAM substrate through its acidic FN1 domain lead to initiation of polysialylation (A) before a switch for polySia recognition to elongate the chains (B).

Illustration adapted from Galuska *et al.*, 2017a

iv. N-glycosylation and autopolySia of polySTs

PolySTs have bifunctional role as they exhibit auto- and NCAM polysialylation activity. Furthermore, ST8Sia II and ST8Sia IV display respectively 6 and 5 N-glycosylation sites as confirmed by PNGase-F treatment (Mühlenhoff *et al.*, 2001 ; Close *et al.*, 2001). Transient

transfection of recombinant V5-tagged ST8Sia II and ST8Sia IV in mammalian CHO cells led to their main expression in the Golgi membrane and at low level in secreted forms in the extracellular space. These forms have been reported highly glycosylated with polySia, particularly for ST8Sia IV (Close and Colley, 1998). Moreover, the expression of human ST8Sia IV in mammalian cells initially polySia-deficient like CHO2A10 cells led to autopolySia on its N-glycans (Mühlenhoff *et al.*, 1996b) whereas its production in insect cells, without sialylation system, led to no autopolySia (Angata *et al.*, 2001). Interestingly in each case, ST8Sia IV is catalytically active suggesting that autopolySia is not essential for its enzymatic activity.

Table 10: Impact of N-glycosylation on enzymatic activity of ST8Sia II and ST8Sia IV. PolySia levels are described *in vivo* or *in vitro* on NCAM or on the enzyme itself. Data retrieved from Mühlenhoff *et al.*, 2001.

Mutants	<i>In vivo</i> polySia on NCAM + enzyme	<i>In vitro</i> autopolySia on enzyme	<i>In vitro</i> polySia on NCAM
ST8Sia II			
WT	+	+	+
N60Q	+/-	+	+
N72Q	+	+	+
N89Q	+/-	-	-
N134Q	+	+	+
N219Q	+/-	-	+/-
N234Q	+/-	+/-	+/-
ST8Sia IV			
WT	+	+	+
N50Q	+	+	+
N74Q	-	-	-
N119Q	+	+	+
N204Q	+	+	+
N219Q	+	+	+

In parallel, N-glycosylation mutants of polySTs were generated in CHO2A10 cells to decipher importance of N-glycosylation sites in enzymatic activity (Mühlenhoff *et al.*, 2001). Substitution of Asn⁷⁴ decreased *in vitro* and *in vivo* enzymatic activity of ST8Sia IV whereas no individual mutant significantly affected *in vivo* polysialylation activity of ST8Sia II. Interestingly, double mutant of Asn⁸⁹/Asn²¹⁹ completely inactive enzymatic activity of ST8Sia II *in vivo* (Table 10). Consequently, these enzymes autopolysialylate their own N-glycans, Asn⁷⁴ for ST8Sia IV and Asn⁸⁹ and Asn²¹⁹ for ST8Sia II (Table 5), to modulate their enzymatic activity onto acceptor substrates like NCAM.

On the other hand, mutagenesis in PBR motif of ST8Sia IV showed a decrease of autopolySia activity whereas reversion of Glu⁹² and Asp⁹⁴ residues enhanced autopolySia but significantly reduced NCAM polySia (Foley *et al.*, 2009). Furthermore, autopolySia of ST8Sia II and ST8Sia IV seems required *in vivo* for SynCAM-1 or NRP-2 to promote longer polySia chains (Mühlenhoff *et al.*, 2013) but not for NCAM (Bhide *et al.*, 2018 ; Close *et al.*, 2001).

To summarize, autopolySia has a fine role to modulate enzymatic activity of polySTs and polySia level and size generated on acceptor substrates.

v. ST8Sia III : polyST or oligoST ?

PolyST studies oftenly mentioned ST8Sia III has an actor involved in the polysialylation process (Angata *et al.*, 2000 ; Sato *et al.*, 2000). As reported in chapter V-F-ii, ST8Sia III sequence is composed by the four conserved SM, shares around 33-35% of sequence similarity with each polyST and has two “PBR-like” and “PSTD-like” motifs (Angata et Fukuda, 2003). ST8Sia II, ST8Sia III and ST8Sia IV could used mono- and disialyl LacNAc residues for activity however polySia were only detectable with ST8Sia II and ST8Sia IV *in vitro* and *in vivo* (Angata *et al.*, 2000). Few studies are described about ST8Sia III and particularly in zebrafish where ST8Sia III is weakly expressed during CNS development and is involved in myotome formation (Bentrop *et al.*, 2008). To summarize, ST8Sia III couldn't be considered as a polyST as it catalyzes the transfer of maximum 2-3 Sia residues, suggesting an enzymatic activity of oligoST.

E. TELEOST FISH POLYST

Only few studies focused on the particular distribution of polyST genes in teleost (Fig. 21) and their enzymatic activity potentially responsible for such complex polySia chains as those described in salmonids (Sato *et al.*, 1993). Indeed, Kitajima and collaborators investigated enzymatic activity of polySTs in rainbow trout (*O. mykiss*) (Asahina *et al.*, 2006). They identified and cloned the three polySTs: one ST8Sia IV, named rtPST which shares 75% of identity with the human ST8Sia IV, and two ST8Sia II, named rtSTX-ov and rtSTX-em, which share 92% of sequence identity between them and around 70% with the human ST8Sia II. Indeed, rtPST is the most abundant polyST whereas rtSTX-ov and rtSTX-em have low or no expression in ovary during oogenesis (Asahina *et*

al., 2006). RtPST displayed *in vivo* activity whereas only low activity was detected *in vitro*. In contrast, each rtSTX showed no *in vivo* and very low *in vitro* activities. To be noted, rtPST produced in COS-1 cells exhibits autopolySia whereas no autopolySia were detectable for rtSTX-ov and rtSTX-em (Asahina *et al.*, 2006).

The authors developed enzymatic assays using the three rainbow trout polySTs and the two human polySTs with radiolabeled CMP-[¹⁴C]Neu5Ac on NCAM and PSGP. Human ST8Sia IV showed 7.8 and 4.4 times higher activity toward PSGP and NCAM than the recombinant rtPST. Likewise, human ST8Sia II displayed about 16 and 48 times higher activity toward PSGP and NCAM than rtSTXs (Asahina *et al.*, 2006). Since rtSTX-ov and rtPST co-exist in ovary, their co-incubation enhanced polyST activity on PSGP, suggesting a cooperative role of these enzymes.

In this study, these teleost polySTs exhibit low enzymatic activity *in vitro* with CMP-Neu5Ac and it raises the question of their use of donor substrates like CMP-Neu5Gc or CMP-Kdn. Indeed, these potential specificities towards other donor substrates should be investigated to explain notably the presence of diverse polySia on O-glycoproteins in salmonid ovary.

In parallel, we explored if the molecular evolution and particular distribution of polySTs in teleosts could be responsible for distinct enzymatic specificities than those reported in mammals. From the three polyST genes identified in *Coregonus maraena*, PCR analysis demonstrated distinct expression profiles for *st8sia2r-1* and *st8sia2r-2* genes in various salmonid tissues. Their highest mRNA expression levels were detected in gonads and telencephalon and in gonads and spleen respectively. Moreover, *st8sia4* gene showed low mRNA levels expressed in the brain, head kidney, gills, gonads and spleen (Venuto *et al.*, 2020a ; Venuto *et al.*, 2020b).

Until now, no entire 3D structure is described for the two human ST8Sia II and ST8Sia IV. Based on the closest described 3D model of the human ST8Sia III (PDB: C5BO6B) (Volkers *et al.*, 2015), other models were generated using Phyre2 server (Huang *et al.*, 2017). In Venuto *et al.*, 2020b, we used YASARA software and took advantage of the PSTD NMR structure of human ST8Sia IV (PDB : 6AHZ) (Peng *et al.*, 2019) to compare PSTD motif between each ST8Sia IV and each ST8Sia II, ST8Sia II-r1 and ST8Sia II-r2 from human and salmonid species (Fig. 24).

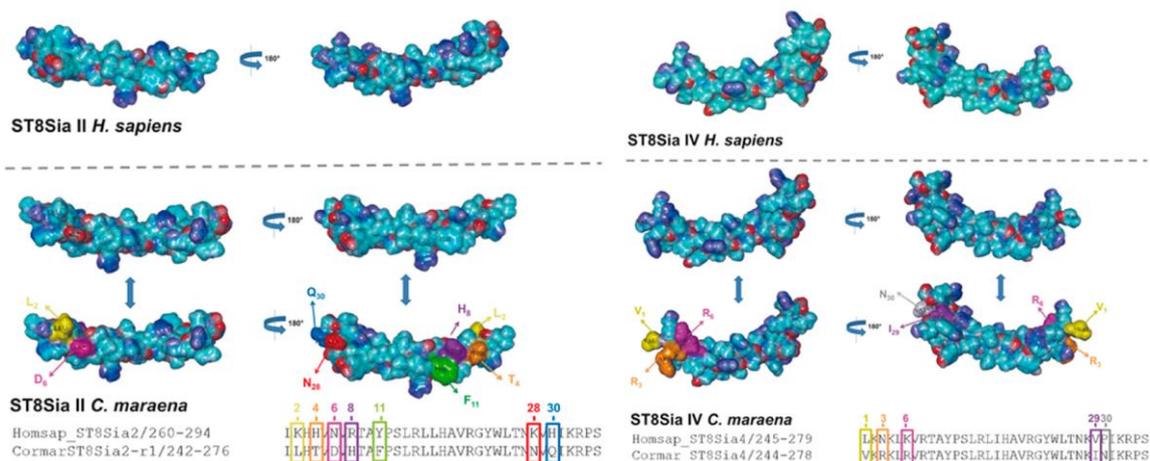


Figure 24: PSTD-electrostatic potential surfaces representations of human and coregone ST8Sia II/II-r and ST8Sia IV. Substitutions between human and coregone PSTD of ST8Sia II/II-r and ST8Sia IV are reported under the YASARA PSTD models. Illustration from Venuto *et al.*, 2020b

Among aa changes noticed between human and coregone polySTs, some of them could have direct impact on enzymatic specificity (Venuto *et al.*, 2020b). In ST8Sia II/II-r, Asn²⁶⁵ to Asp²⁶⁸ (reported 6 in pink, Fig. 24) could influence interaction between PSTD and polySia. Lys²⁸⁶ to Asn²⁸⁹ (reported 28 in red, Fig. 24) is correlated to lower enzymatic activity *in vitro* in rtSTXs (Angata *et al.*, 2006). In each ST8Sia IV, no change is observed for Lys^{246/245}, Lys^{248/249}, Arg^{252/253} (Fig. 24), the basic residues reported of highest impact on enzymatic activity (Nakata *et al.*, 2006).

Nevertheless, other aa changes like Asn²⁴⁷ to Arg²⁴⁶ (reported 3 in orange, Fig. 24) and Lys²⁵⁰ to Arg²⁴⁹ (reported 6 in purple, Fig. 24) may influence interaction with donor substrates, especially with CMP-Neu5Ac and polySia (Nakata *et al.*, 2006 ; Zhou *et al.*, 2015 ; Liao *et al.*, 2020).

Consequently, 3D-modeling of the PSTD unveiled potential changes of function for salmonid polySTs (Venuto *et al.*, 2020b) and raise the question of the impact on the donor substrate specificities of these enzymes.

VII. METHODS TO STUDY SIALYLATION AND SIALYLTRANSFERASES

A. PRODUCTION OF GLYCOENGINEERED ENZYMES AND ACCEPTORS

i. In bacteria system

Prokaryotes are the easier systems used as cell factories for protein production as they have a high yield and low fermentation costs and times (Pratama *et al.*, 2021). In contrast, bacterial hosts can't produce glycoproteins with mammalian-type glycosylation and display immunogenic and different *N*-glycan structures than in mammals. They also encounter difficulties for PTMs, solubility and are often met with non-functional protein aggregates (Table 11). Although their biosynthetic pathways differ from eukaryotes, the identification of O- and N-glycosylation pathways in pathogenic *C. jejuni* led to subsequently transfer of various enzymes by engineering of *E. coli* (Jaffé *et al.*, 2014 ; Pratama *et al.*, 2021). These genetic modifications have increased the efficiency of glycoprotein production in *E. coli* and allowed control of glycosylation with less heterogeneity.

Bacterial systems were oftenly chosen for further structural analysis of mammalian and bacterial STs. For instance, rat ST6Gal I, human ST6Gal I and porcine ST3Gal I were produced in *E. coli* to obtain high level of production allowing purification and crystallization (Meng *et al.*, 2013 ; Harrus *et al.*, 2020 ; Rao *et al.*, 2009). Belonging to GT-38 family, bacterial polyST Cst-II from *C. jejuni* were notably produced in *E. coli* and purified for structural analysis (Chiu *et al.*, 2004). PolyST from *N. meningitidis* was produced in *E. coli* allowing its characterization (Willis *et al.*, 2008) and its development for polySia applications like increasing half-life of factor IX in the serum (Lindhout *et al.*, 2011). Although a good catalyst, this bacterial polyST was not the most stable enzyme, likely due to absence of critical disulfide bonds and PTMs as found for human polySTs.

Structure-based rational design, mutagenesis and biochemical studies (Benkoulouche *et al.*, 2019 ; Cui *et al.*, 2022) could help for the refinement of bacterial polySTs kinetic properties (efficiency, regio- and stereo-selectivity) required to prepare polySia for a large-scale production in applied biocatalysis system (Flitsch *et al.*, 2012). Based on directed evolution, mutants of polySTs were notably developed to improve physicochemical and biochemical properties like solubility and thermal stability (Janesch *et al.*, 2019a).

Table 11: Advantages and disadvantages of cell system for glycoprotein production.Data retrieved from Swiech *et al.*, 2014

Cell system	Bacteria	Plant	Insect	Yeast	Mammalian
Speed and Yield	High	Low	Medium	High	Medium
Cost	Low	Medium	Medium	Low	Medium-High
Glycosylation and PTM	Low	Medium	Medium	Low-Medium	High
Folding	Low	Medium	Medium	Low	High

ii. In eukaryotic system

Eukaryotic cell lines are now widely used as factories and remain the preferred host to produce recombinant glycoproteins as reported in a recent global approach for production of all GTs (Moremen *et al.*, 2018). Glycoprotein expression depends of the nature of cell types (insect, mammals, human, yeast) because glycosylation profile is different for each of them and could lead to various scales and quality of production (Table 11).

Plant and insect cell systems could be used for production as they can synthesize complex glycoproteins. Since they produce xylose, α 1-3 fucose or paucimannosidic N-glycans, these platforms have been glycoengineered by genome edition to produce complex N-glycans with Gal and Sia residues (Gomord *et al.*, 2010). To be noted, insect-cell expression was used for the production and crystalization of human ST8Sia III (Volkers *et al.*, 2015). Likewise in yeast which produce mainly high-mannose type, mutants were developed to customize glycosylation and could be scalable in fermentation for high level of production (Betenbaugh *et al.*, 2004).

As listed in Table 11, mammalian cell systems also have advantages and disadvantages. Interestingly, most of recombinant biotechnological and therapeutic glycoproteins are produced in Chinese Hamster Ovary cells (CHO) allowing a large-scale expression (Jaffé *et al.*, 2014) with human-like glycosylation, PTMs and chaperones required for proper folding and function of glycoproteins (Betenbaugh *et al.*, 2004 ; Mestrom *et al.*, 2019). Particularly, CHO cells were transiently transfected to study enzymatic activity of human polySTs ST8Sia II and ST8Sia IV

(Mühlenhoff *et al.*, 1996b ; Close and Colley, 1998). Nevertheless, culture conditions in mammalian cells are more challenging than bacteria, notably with expensive growth media, slower growth rates and susceptibility to viral and bacterial infection (Swiech *et al.*, 2014).

Human cell systems have been also optimized for glycoprotein production. For instance, Human Eukaryotic Kidney (HEK293) cells have the advantage to display human glycosylation and to be highly transfectable as described with transient transfection for all GTs (Moremen *et al.*, 2018 ; Noel *et al.*, 2018). Indeed, CHO cells produce only α 2-3-sialylated N-glycans whereas HEK293 cells display an heterogeneous mixture of α 2-3- and α 2-6-sialylated N-glycans (Fujitani *et al.*, 2013) (Table 12). Finally, new complex cell-free strategies like SIMPLEx tend to be developed for soluble expression of integral membrane proteins as reported for human ST6Gal I (Jaroentomeechai *et al.*, 2022) and should be more investigated in the future particularly to control glycosylation patterns of production.

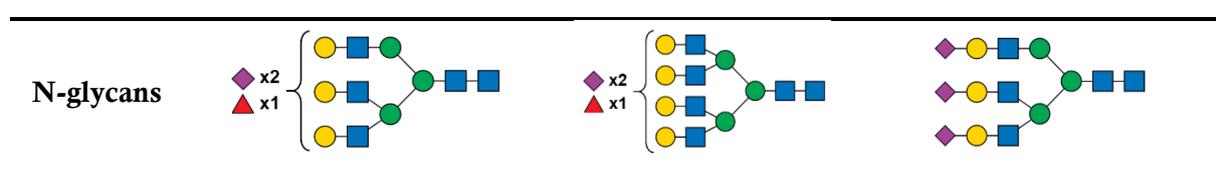
Table 12: Glycomic profiles and sialylated glycoconjugates in (A) HEK293 and (B) CHO-K1 cells.

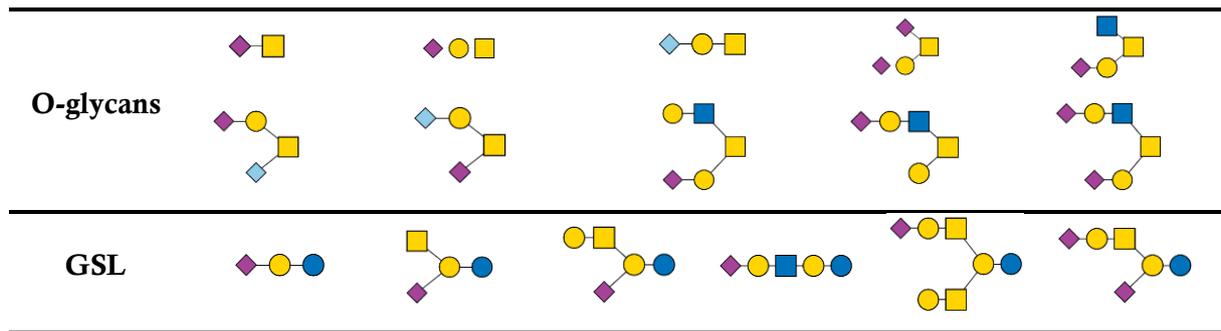
Data retrieved from suppdata Fujitani *et al.*, 2013

(A) HEK293 cells

Structure	Proportion per glycosylation (%)	Profile	Proportion per structure (%)
N-glycans type	45	High-mannose	80
		Small oligoman.	15
		Complex/hybrid	5
O-glycans series	45	Core 1	90
		Core 2	5
		Core 3	5
GSL	10	Ganglio-	75
		Lacto-	15
		Globo-	10

Sialylated glycoconjugates in HEK293 cells

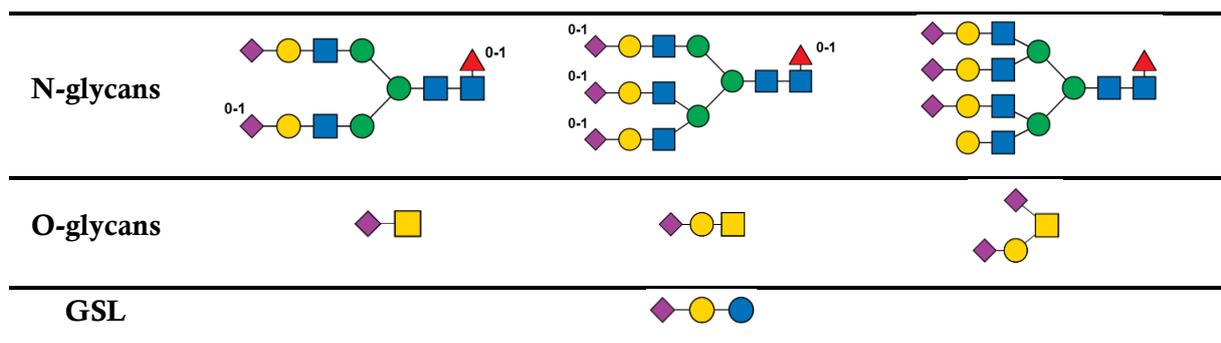




(B) CHO-K1 cells

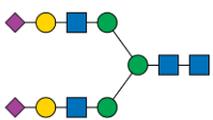
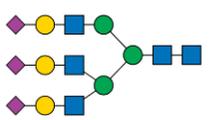
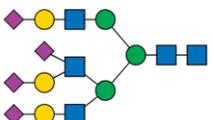
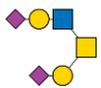
Structure	Proportion per glycosylation (%)	Profile	Proportion per structure (%)
N-glycans type	70	High-mannose	70
		Small oligoman.	5
		Complex/hybrid	25
O-glycans series	15	Core 1	100
		Core 2	0
		Core 3	0
GSL	15	Ganglio-	95
		Lacto-	5
		Globo-	0

Sialylated glycoconjugates in CHO-K1 cells



In parallel, HEK293 cells system allowed the production of other glycoproteins which could serve for polySTs as sialylated acceptor substrates used during my thesis, such as bovine fetuin, which is a glycoprotein carrying three complex triantennary-type N-glycans presenting the sialylated LacNAc (Gal β 1-4GlcNAc) epitope (Table 13) as well as three mucin-type O-glycans (disialylated core 1: Neu5Ac α 2-3Gal β 1-3[Neu5Ac]GalNAc-Ser/Thr) (Baenziger & Fiete, 1979) or can be directly used as platform for exogenous sialylation on cell surface (Mbua *et al.*, 2013).

Table 13: Description of sialylated glycanic structures of acceptor substrates used to study sialyltransferase activity. Representation were done with GlycoGlyph (Mehta and Cummings, 2020).

Acceptor substrate	N-/O-glycans sialylated	Symbolic representation	Sialylated glycanic structure	References
Fetuin	N- : 3 tri-antennary sialylated N-glycans		Neu5Ac α 2-6(3)Gal β 1-4GlcNAc β 1-2Man α 1-3[Neu5Ac α 2-6(3)Gal β 1-4GlcNAc β 1-2Man α 1-6]Man β 1-4GlcNAc β 1-4GlcNAc-Asn	Baenziger and Fietze, 1979
			Neu5Ac α 2-6(3)Gal β 1-4GlcNAc β 1-2[Neu5Ac α 2-6(3)Gal β 1-4GlcNAc β 1-4]Man α 1-3[Neu5Ac α 2-6(3)Gal β 1-4GlcNAc β 1-2Man α 1-6]Man β 1-4GlcNAc β 1-4GlcNAc-Asn	
			Neu5Ac α 2-6(3)Gal β 1-4GlcNAc β 1-2[Neu5Ac α 2-6(3)Gal β 1-4[Neu5Ac α 2-6]GlcNAc β 1-4]Man α 1-3[Neu5Ac α 2-6(3)Gal β 1-4GlcNAc β 1-2Man α 1-6]Man β 1-4GlcNAc β 1-4GlcNAc-Asn	
	O- : core-1 mucin-type		Neu5Ac α 2-6GalNAc-Ser/Thr	
		Neu5Ac α 2-3Gal β 1-3GalNAc-Ser/Thr		
		Neu5Ac α 2-3Gal β 1-3[Neu5Ac α 2-3Gal β 1-4GlcNAc β 1-6]GalNAc-Ser/Thr		



Galβ1-3[Neu5Acα2-6]GalNAc-Ser/Thr



Neu5Acα2-3Galβ1-3[Neu5Acα2-6]GalNAc-Ser/Thr

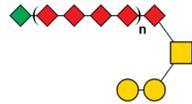
PSGP-L
PSGP-H

O- : core-1 and core-2 mucin-type disialylated (-L) or oligo- / polysialylated from 2 to 20 Sia (-H)

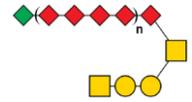


Kdnα2-8(Neu5Acylα2-8)_{n~2-20}Neu5Acylα2-6[Galβ1-3]GalNAc-Ser/Thr

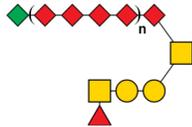
Inoue and Inoue, 1997



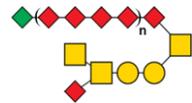
Kdnα2-8(Neu5Acylα2-8)_{n~2-20}Neu5Acylα2-6[Galβ1-4Galβ1-3]GalNAc-Ser/Thr



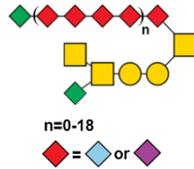
Kdnα2-8(Neu5Acylα2-8)_{n~2-20}Neu5Acylα2-6[GalNAcβ1-3Galβ1-4Galβ1-3]GalNAc-Ser/Thr



Kdnα2-8(Neu5Acylα2-8)_{n~2-20}Neu5Acylα2-6[Fuca1-3GalNAcβ1-3Galβ1-4Galβ1-3]GalNAc-Ser/Thr

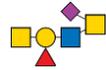


Kdnα2-8(Neu5Acylα2-8)_{n~2-20}Neu5Acylα2-6[GalNAcβ1-4[Neu5Acylα2-3]GalNAcβ1-3Galβ1-4Galβ1-3]GalNAc-Ser/Thr

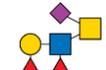


Kdn α 2-8(Neu5Acyl α 2-8)_{n \approx 2-20}Neu5Acyl α 2-6[GalNAc β 1-4[Kdn α 2-3]GalNAc β 1-3Gal β 1-4Gal β 1-3]GalNAc-Ser/Thr

Bovine Submaxillary Mucin (BSM)	O- : Core -1, Core -2, Core -3, Core -5		Neu5Acyl α 2-6GalNAc-Ser/Thr	Kim <i>et al.</i> , 2020
			Neu5Ac α 2-6[GlcNAc β 1-3]GalNAc-Ser/Thr	
			Neu5Ac α 2-3Gal β 1-3GalNAc-Ser/Thr	
			Neu5Gc α 2-6[GalNAc α 1-3]GalNAc-Ser/Thr	
			Neu5Gc α 2-6[Gal β 1-4GlcNAc β 1-3]GalNAc-Ser/Thr	
			Neu5Ac α 2-6[Fuc α 1-3(4)Gal β 1-3]GalNAc-Ser/Thr	
			Neu5Ac α 2-3Gal β 1-3[GlcNAc β 1-6]GalNAc-Ser/Thr	
			Neu5Acyl α 2-6[Fuc α 1-3(4)Gal β 1-4GlcNAc β 1-3]GalNAc-Ser/Thr	



Neu5Ac α 2-6[GalNAc β 1-4[Fuca1-3]Gal β 1-4GlcNAc β 1-3]GalNAc-Ser/Thr

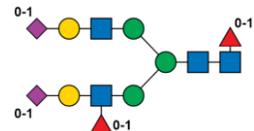


Neu5Ac α 2-6[Fuca1-3(4)Gal β 1-4[Fuca1-3]GlcNAc β 1-3]GalNAc-Ser/Thr



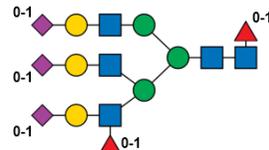
Orosomucoid
(α -1-acid
glycoprotein)

N- ; di-, tri- and
tetra-antennary
LacNAc sialylated
N-glycans

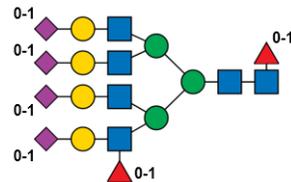


(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4[(Fuca1-3)₀₋₁]GlcNAc β 1-2Man α 1-3[(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4GlcNAc β 1-2Man α 1-6]Man β 1-4GlcNAc β 1-4[(Fuca1-6)₀₋₁]GlcNAc-Asn

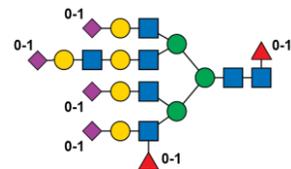
Imre *et al.*, 2008



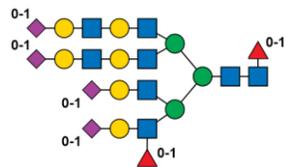
(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4[(Fuca1-3)₀₋₁]GlcNAc β 1-2[(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4GlcNAc β 1-4]Man α 1-3[(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4GlcNAc β 1-2Man α 1-6]Man β 1-4GlcNAc β 1-4[(Fuca1-6)₀₋₁]GlcNAc-Asn



(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4[(Fuca1-3)₀₋₁]GlcNAc β 1-2[(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4GlcNAc β 1-4]Man α 1-3[(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4GlcNAc β 1-2[(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4GlcNAc β 1-4]Man α 1-6]Man β 1-4GlcNAc β 1-4[(Fuca1-6)₀₋₁]GlcNAc-Asn



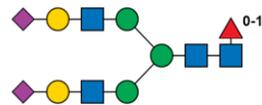
(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4[(Fuca1-3)₀₋₁]GlcNAc β 1-2[(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4GlcNAc β 1-4]Man α 1-3[(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4GlcNAc β 1-4Gal β 1-4GlcNAc β 1-2[(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4GlcNAc β 1-4]Man α 1-6]Man β 1-4GlcNAc β 1-4[(Fuca1-6)₀₋₁]GlcNAc-Asn



(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4[(Fuc α 1-3)₀₋₁]GlcNAc β 1-2[(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4GlcNAc β 1-4]Man α 1-3[(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4GlcNAc β 1-4Gal β 1-4GlcNAc β 1-2[(Neu5Ac α 2-6(3))₀₋₁Gal β 1-4GlcNAc β 1-4Gal β 1-4GlcNAc β 1-4]Man α 1-6]Man β 1-4GlcNAc β 1-4[(Fuc α 1-6)₀₋₁]GlcNAc-Asn

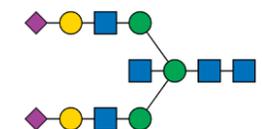
ALCAM/CD166

N- : bi-, tri- and tetra-antennary sialylated N-glycans (α 2.6- and α 2.3-Sia)



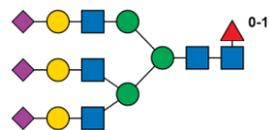
Neu5Ac α 2-6Gal β 1-4GlcNAc β 1-2Man α 1-3[Neu5Ac α 2-6Gal β 1-4GlcNAc β 1-2Man α 1-6]Man β 1-4GlcNAc β 1-4[(Fuc α 1-6)₀₋₁]GlcNAc-Asn

Kim and Hahn, 2015

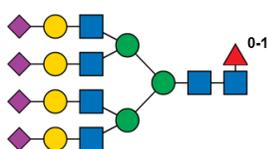


Neu5Ac α 2-6Gal β 1-4GlcNAc β 1-2Man α 1-3[GlcNAc β 1-4][Neu5Ac α 2-6Gal β 1-4GlcNAc β 1-2Man α 1-6]Man β 1-4GlcNAc β 1-4GlcNAc-Asn

Ferragut *et al.*, 2019



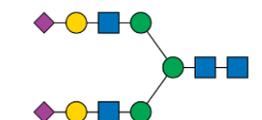
Neu5Ac α 2-3(6)Gal β 1-4GlcNAc β 1-2[Neu5Ac α 2-3(6)Gal β 1-4GlcNAc β 1-4]Man α 1-3[Neu5Ac α 2-3(6)Gal β 1-4GlcNAc β 1-2Man α 1-6]Man β 1-4GlcNAc β 1-4[(Fuc α 1-6)₀₋₁]GlcNAc-Asn



Neu5Ac α 2-3Gal β 1-4GlcNAc β 1-2[Neu5Ac α 2-3Gal β 1-4GlcNAc β 1-4]Man α 1-3[Neu5Ac α 2-3Gal β 1-4GlcNAc β 1-2[Neu5Ac α 2-3Gal β 1-4GlcNAc β 1-2]Man α 1-6]Man β 1-4GlcNAc β 1-4[(Fuc α 1-6)₀₋₁]GlcNAc-Asn

DNase I

N- : LacNAc sialylated N-glycans

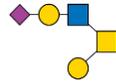


Neu5Ac α 2-6(3)Gal β 1-4GlcNAc β 1-2Man α 1-3[Neu5Ac α 2-6(3)Gal β 1-4GlcNAc β 1-2Man α 1-6]Man β 1-4GlcNAc β 1-4GlcNAc-Asn

Weide *et al.*, 2006

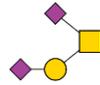
NRP-2

O- : Core-1, Core-2

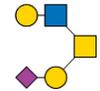


Neu5Ac α 2-3Gal β 1-4GlcNAc β 1-6[Gal β 1-3]GalNAc-Ser/Thr

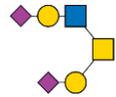
Rollenhagen *et al.*, 2013



Neu5Ac α 2-3Gal β 1-3[Neu5Ac α 2-6]GalNAc-Ser/Thr



Gal β 1-4GlcNAc β 1-6[Neu5Ac α 2-3Gal β 1-3]GalNAc-Ser/Thr



Neu5Ac α 2-3Gal β 1-4GlcNAc β 1-6[Neu5Ac α 2-3Gal β 1-3]GalNAc-Ser/Thr

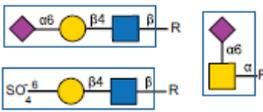
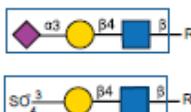
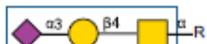
B. LECTINS AND NEURAMINIDASES

Lectins are carbohydrate-binding proteins could be useful to determine specific type of glycosylated pattern *in vitro* or *in cellulo*. Particularly about sialylation, the plant lectins *Sambucus nigra* (SNA) which mainly recognizes LacNAc motifs sialylated with α 2,6-linkages and *Maackia amurensis* (MAH) which recognizes Gal β 1-3GalNAc motifs sialylated with α 2,3-linkages (Geisler and Jarvis, 2011) make it possible to distinguish the activities of ST6Gal and ST3Gal enzymes (Table 14).

Expressed on various immune cells, siglecs are Sia binding Ig-like lectins which selectively recognize different linkage-specific Sia. Among siglecs, the endogenous ligand Siglec-11 has been identified on macrophage, microglia and ovarian fibroblasts in humans and could interact with pathogens (Angata, 2018). Siglec-11 preferentially recognize Neu5Gc over Neu5Ac (Angata, 2018) and are reported to selectively bind polySia as shown by MD simulations (Shastry *et al.*, 2020). To be noted, the recombinant soluble human microglial form of Siglec-11 binds endogenous, cell surface and immobilized polySia better than the tissue macrophage form of Siglec-11 (Hane *et al.*, 2021), suggesting its potential use as biotechnological tool to detect polySia.

Table 14: Lectin description. With their relative specificity to bind glycan motifs.

Data retrieved from Vector Labs.

Lectin	Common Name	Preferred Sugar Specificity	General Binding motif
<i>Ricinus communis I</i>	RCA I, RCA ₁₂₀	Gal	
<i>Sambucus nigra</i>	SNA, EBL	Neu5Ac α 2-6Gal/GalNAc Neu5Ac α 2-6GalNAc	
<i>Wisteria floribunda</i>	WFA, WFL	GalNAc	
<i>Maackia amurensis I</i>	MAL I, MAL	Neu5Ac α 2-3Gal β 1-3/4GlcNAc Gal β 1-4GlcNAc	
<i>Maackia amurensis II</i>	MAL II, MAH	Neu5Ac α 2-3Gal β 1-3GalNAc	

Neuraminidases (NEU), also called sialidases, could be considered as “the STs counterpart” maintaining a balance in sialylation process as they recognize and cleave Sias from glycoconjugates required for a cell and tissue physiological state. Initially characterized from bacteria or viral strains and encoded by NEU genes in eukaryota, Neu are divided in four family which don't share protein sequence identity but have conserved motifs and 3D structure similar (Monti *et al.*, 1999). NEU are located in the cytosol or in lysosome and considered as the first step for catabolism of sialylated glycoconjugates (Corfield, 1992). Neu1 and Neu4 act in lysosomal compartments to recycle sialylated glycans, even Neu1 could regulate sialylation at cell surface. Neu3 is more specific on gangliosides cleavage at cell surface and Neu2 act in cytosol (Varki *et al.*, 2015b). Compared to STs, NEU have substrate specificity towards α 2-3, α 2-6 or α 2-8 glycosidic linkages and potential O-methylation and O-acetylation cleaved on sialylated glycoconjugates. Indeed, hydroxyl group in C7, C8, C9 or acetylation in C4 notably is known to decrease or inhibit NEU activity respectively (Varki *et al.*, 2015b). Pathogens express sialidases which could help them for recognition as for immune escape. Notably, H1N1 *influenza* refers to Neu N1 strain and Neu inhibitors are currently and widely used as drugs targeting directly NEU as Relenza and Tamiflu (Rewar *et al.*, 2015).

Endosialidases (endoNs) are specific NEU identified in bacteriophages as tailspike enzymes which recognize, bind and cleave polySia (Jakobsson *et al.*, 2012). Among endoNs, endosialidase activity isolated from *E. coli* K1 phage named endoNF specifically degrade α 2,8-linked polySia (Hallenbeck *et al.*, 1987 ; Mühlenhoff *et al.*, 2003). EndoNF recognize oligoSia of DP = 5 Sia residues in a helical conformation to cleave polySia at DP = 7 Sia (Haselhorst *et al.*, 2006). Basically, this enzyme is an efficient tool to control the presence of polySia as described on NCAM (Mühlenhoff *et al.*, 1996a ; Close *et al.*, 2001), on SynCAM-1 (Galuska *et al.*, 2010a) or on NRP-2 (Rollenhagen *et al.*, 2013). For efficient cleavage by endoNF, a minimum of eight Sia residues was reported under moderate digestion conditions (Finne and Mäkelä, 1985) whereas at least three Sia are required at the non-reducing end from the cleavage site (Schwarzer *et al.*, 2009). To be noted, endoNF could cleave polySia composed by Neu5Ac and Neu5Gc as they are described as Endo-N-acyl-neuraminidases (Kitajima *et al.*, 1988).

Crystal structure of endoNF reveals a homotrimeric mushroom-shaped molecule with a catalytic part subdivided into three domains (Stummeyer *et al.*, 2005). EndoNF is mainly composed by β -structures (Fig. 25), notably lectin-like domains b1 and b2 acting as binding sites for oligoSia and polySia (Stummeyer *et al.*, 2005). Another crystal structure of endoNF has been solved in complex with triSia (PDB ID : 3GVL) and confirms the homotrimeric structure organization of the enzyme (Schulz *et al.*, 2010).

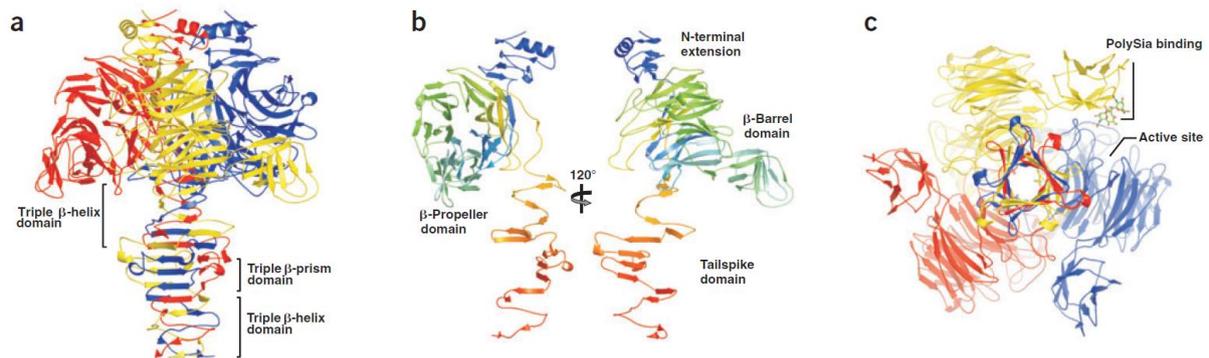


Figure 25: Homotrimeric structure of endoNF. (a) Each monomer is colored in red, yellow and blue and represented in ribbon diagram. (b) The three domains which composed endoNF. (c) Bottom view of EndoNF with its active site and polySia binding.

Illustration from Stummeyer *et al.*, 2005

Moreover, endoNF contains a C^{ter} domain (CTD) which act as an intramolecular chaperone, crucial to mediate homotrimer folding required for active enzymes (Mühlenhoff *et al.*, 2003). Indeed, CTD mutants like R837A and Q853A lead to insoluble and inactive aggregates of endoNF (Schwarzer *et al.*, 2009). After homotrimerization of EndoNF, CTD should be released by proteolysis to allow kinetically stable complex which can bind polySia (Schulz *et al.*, 2010).

To allow proper catalysis, endoNF should have a specific arrangement of its three binding sites (b2, b1 and a) in close contact (Stummeyer *et al.*, 2005). Interestingly, active site of endoNF is organized in a hydrophobic pocket which recognize C2 of Sia residues (Schulz *et al.*, 2010). Consequently, N-acyl group in C5 didn't seem to be involved for recognition and hydrolysis of polySia, suggesting an accomodation of Neu5Ac and Neu5Gc residues for EndoNF activity.

Moreover, endoNF catalyze polySia cleavage in an inverting S_N1-type mechanism with an optimal pH ~ 5 (Mühlenhoff *et al.*, 2003 ; Morley *et al.*, 2009) but still highly also active at pH 7.4 and 8.0 (Stummeyer *et al.*, 2005 ; Schwarzer *et al.*, 2009). Indeed, some aa are crucial for cleavage

process : Glu⁵⁸¹ activate a water molecule to attack the glycosidic linkage, then His³⁵⁰ allows deprotonation and the negative charge is delocalized towards Tyr³²⁵. Finally, reducing end moiety of Sia are released and EndoNF active site is regenerated to rebind polySia (Schulz *et al.*, 2010).

Various EndoNF mutants were achieved to determine aa involvement in polySia binding and cleavage. E581A, R596A and R647A mutants exhibit 95% less activity than WT (Stummeyer *et al.*, 2005). Crystal structure of H350A mutant (PDB ID: 3GVK) showed a binding of the active site with triSia residues. Co-crystallization of R647A mutant with pentameric Sia highlights involvement of binding sites b1 and b2 with five Sia residues (Fig. 26) (Schulz *et al.*, 2010).

Although exact mechanism of processivity for polySia cleavage is not known, a cyclic mechanism was suggested in four steps (Fig. 26) where interactions of the three binding sites could be the key for processivity (Schwarzer *et al.*, 2009 ; Schulz *et al.*, 2010).

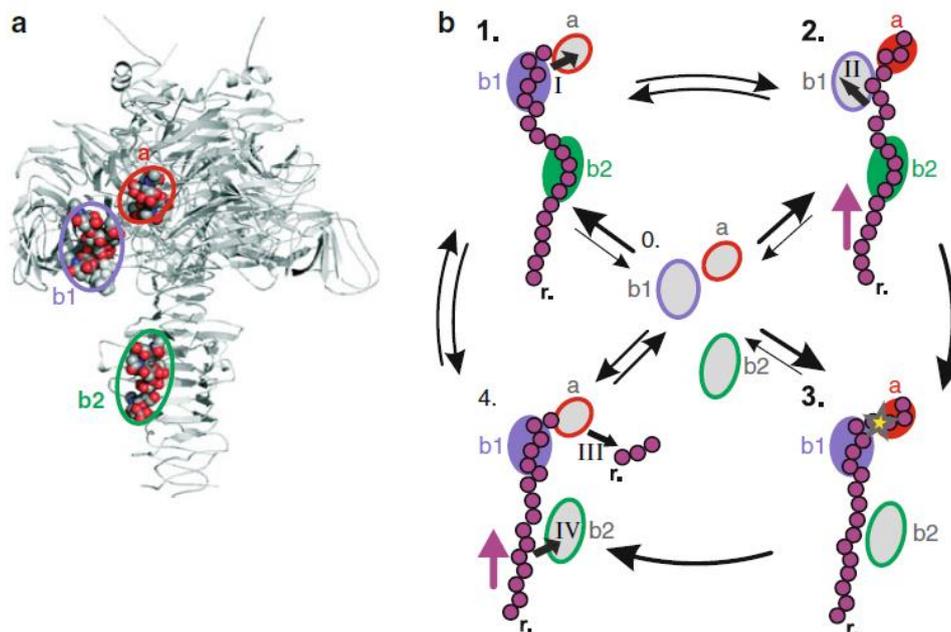


Figure 26: Suggested processivity mechanism of endoNF bound with oligo/polySia. (a) Structure of endoN with its three binding sites a, b1 and b2. (b) TriSia residues (red spheres) binds active site a (red) ; tetra and pentamer binds binding sites b1 (purple) and b2 (green). By diffusion, polySia changes its binding state (from I – IV). Notably, a conformational change of polySia is required for catalysis (asterisk). Then, polySia reorientation (IV) happened simultaneously with cleavage (III) involving a, b1 and b2 binding sites for further regeneration and processivity of EndoN.

Illustration from Jakobsson *et al.*, 2012

Finally, inactivated mutant of endoN (iEndoN) could be used as lectin to detect polySia chains. Different iEndoN mutants have been developed to investigate binding to polySia, particularly R592A/R647A mutant which displays higher binding affinity (K_D) than mAb735 specific anti-polySia antibody (*i.e.* K_D of 1.9 nM for iEndoNF vs 7 nM for mAb735) (Schwarzer *et al.*, 2009). Likewise, iEndoN fixed on beads were used to isolate, concentrate and purify soluble polySia NCAM in the lung (Ulm *et al.*, 2013). In parallel, iEndoN have been fused to GFP and used in bacterial polyST enzymatic assay (Yu *et al.*, 2014) or to detect polySia on cells by fluorescence microscopy (Jokilammi *et al.*, 2004). Interestingly, the use of iEndoN-GFP showed equal binding in both Neu5Ac and Neu5Gc supplemented SH-SY5Y cells (Davies *et al.*, 2012), suggesting the same recognition for polySia composed by these two Sia residues.

C. ANTIBODIES

Antibodies are powerful biochemical tools to detect Sia epitopes in cells or tissues like sLe^x or sT antigens. Depending on their immunospecificities, antibodies can also detect more complex sialylated structures as di-/tri-/oligo-/polySia (Kitajima *et al.*, 2015). They are divided in three different groups according their specificity of binding towards linkage, DP and nature of Sia residues (Table 15). These antibodies were oftenly produced and collected in mammalian organs resulting from immunogenicity towards Sia structures.

Indeed, group III antibodies recognize mainly diSia structures, group II are more diverse whereas group I recognize specifically α 2,8-linked polySia structures. Nevertheless, their uses are limited due to their highly specificity as for instance S2-566 which detect only Sia on G_{D3} (Sato *et al.*, 2000), 12E3 which is restricted to polysialylated-NCAM (Seki and Arai, 1991) whereas mAb735 antipolySia antibody (Häyrynen *et al.*, 1989) is widely used to detect polySia of minimum DP = 11 (Mühlenhoff *et al.*, 2001 ; Nishimura *et al.*, 2014 ; Kitajima *et al.*, 2015) (Table 15).

As detected by SPR, mAb735 have high affinity towards polySia influenced by DP, pH, temperature and ionic environment (Häyrynen *et al.*, 2002). K_D of mAb735 for oligoSia and short polySia (DP < 16) is $\sim 7.4 \mu\text{M}$ whereas K_D is dramatically increased for polySia-NCAM ($\sim 5 \text{ nM}$) and even more towards polySia of DP = 200 ($K_D \sim 0.85 \text{ nM}$) (Häyrynen *et al.*, 2002).

Table 15: Sia epitope antibodies and their specificity. Towards group I : polySia, group II : oligoSia and polySia and group III : diSia and oligoSia. Ab : antibody clone ; Ig : Immunoglobulin ; DP : Degree of Polymerization ; n.d. : not determined.

Adapted from Teinturier-Lelièvre thesis, 2006 ; Kitajima *et al.*, 2015 ; Sato and Kitajima, 2018.

Group	Ab name	Sia specificity	DP	Ig type	Immunogen	References
I	735	-(8Neu5Ac2) _n -	n ≥ 11	Mouse IgG _{2A}	<i>N. meningitidis</i> group B	Frosch <i>et al.</i> , 1985
	H.46		n ≥ 8	Horse IgM		Häyrynen <i>et al.</i> , 1989
	4F7	-(9Neu5Ac2) _n -	n.d.	Mouse IgG	<i>N. meningitidis</i> group C	Jennings <i>et al.</i> , 1985
	12F8	-(8Neu5Ac2) _n -	n.d.	Rat IgM	Mouse brain membrane	Miyata <i>et al.</i> , 2006
II	12E3		n ≥ 5		Embryonic rat forebrain	Chuang and Lagenaur, 1990
	2-2B	-(8Neu5Ac2) _n -	n ≥ 4		<i>N. meningitidis</i> group B	Seki and Arai, 1991
	5A5		n ≥ 3	Mouse IgM	Embryonic rat spinal cord membrane	Mandrell and Zollinger, 1982
	Kdn8kdn	-(8Kdn2) _n -	n ≥ 2		Kdn-gp	Dodd <i>et al.</i> , 1988
	OL.28	-(8Neu5Ac2) _n -	n ≥ 4		Rat oligodendrocyte	Kanamori <i>et al.</i> , 1994
	2-4B	-(8Neu5Gc2) _n -	n ≥ 2		Oligo/polyNeu5Gc-PE	Martersteck <i>et al.</i> , 1996
III	1E6		n = 2		(Neu5Ac) ₂ -glycopolymer	Sato <i>et al.</i> , 1998
	S2-566	-(8Neu5Ac2) _n -	n = 2	Mouse IgM	Human G _{D3}	Sato <i>et al.</i> , 2000
	2A11		n = 2-3		Sea urchin sperm gangliosides	Ohta <i>et al.</i> , 2000
	AC1	-(8Neu5Gc2) _n -	n = 2-4	Mouse IgG3	Rat thymuses (Neu5Gc) ₂ G _{D1c}	Nohara <i>et al.</i> , 1997
	A2B5	-(8Neu5Ac2) _n -	n = 3	Mouse IgM	G _{Q1c} , G _{P1c} , G _{H1c} , OAcG _{T3}	Dubois <i>et al.</i> , 1990
	FS1	-(8Neu5Ac2) _n -	n = 2-3	Mouse IgG3	G _{D3} , G _{Q1b}	Koga <i>et al.</i> , 2005

Furthermore, mAb735 single chain variable fragment (scFv735) has been crystallized in complex with eight Sia residues (PDB ID: 3WBD) (Nagae *et al.*, 2013). Crystal structure revealed that two scFv735 molecules are associated with one octaSia. Each molecule is composed by variable heavy V_H and light V_L chain domain arranged in β -sandwich fold. One molecule interacts with Sia2-Sia4 and the other interacts with Sia6-Sia8 residues whereas Sia5 interaction is weak (Fig. 27). The antigen binding site is formed by six complementary-determining regions (CDR) which are all involved in water-mediated interactions, except L3 (Fig. 27b). Eleven structured water molecules are located in the space between protein and ligand. To be noted, Tyr³⁷ and Arg⁵⁵ in V_L and Tyr¹⁵⁹ and Asp²³² in V_H regions are critical aa for antigen recognition (Nagae *et al.*, 2013).

Interestingly, mAb735 preferentially binds long polySia as CDR regions formed a loop to fit with helical segments of long polySia (Nagae *et al.*, 2013). Nevertheless, no structural information was available about mAb735 recognition of polySia composed by Neu5Gc or Kdn residues.

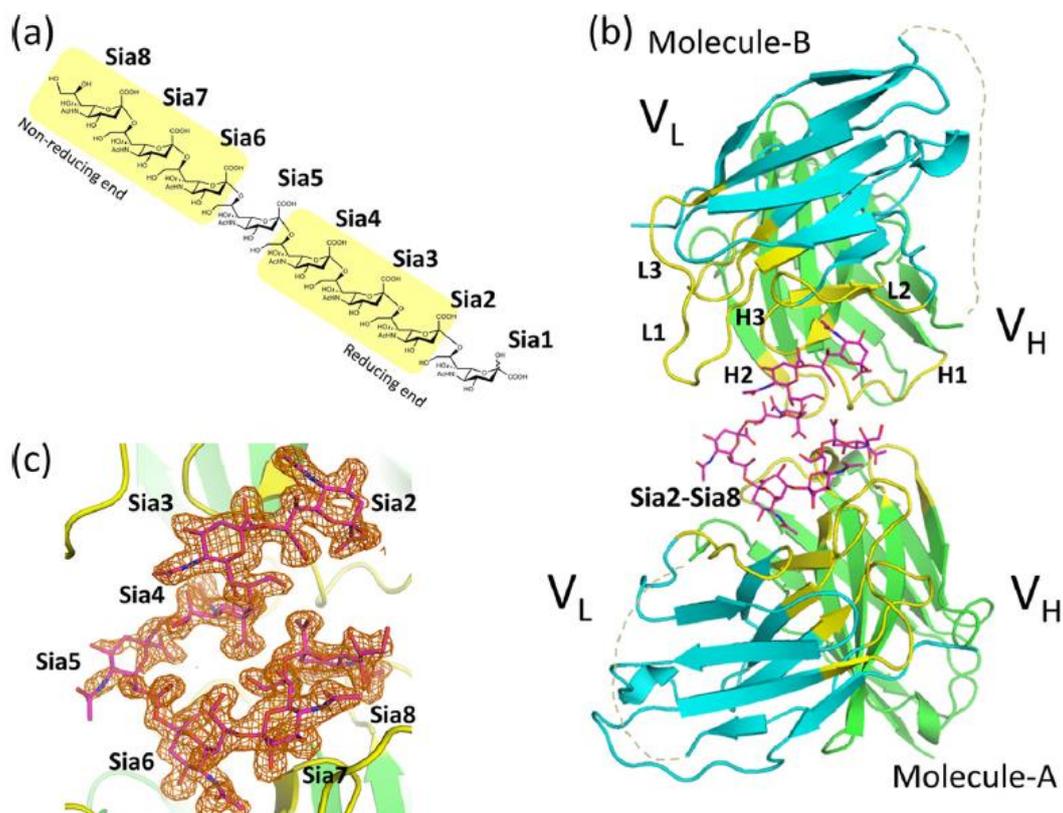


Figure 27: Crystal structure of scFv735 in complex with oligoSia. (a) Chemical structure of the eight Sia residues co-crystallized with scFv735. (b) Co-crystallized structure of two scFv735 molecules with octaSia (purple). V_L and V_H domains are colored in cyan and green. (c) Omit map showed interaction with octaSia.

Illustration from Nagae *et al.*, 2013

D. ENZYMATIC ASSAYS

i. Use of radiolabeled CMP-Neu5Ac

After the first molecular cloning of STs (Weinstein *et al.*, 1987), many enzymatic assays were developed in order to characterize their enzymatic activity and specificities. First *in vitro* enzymatic assays in the 90's were based on the use of radiolabeled donor substrates CMP-[¹⁴C]Neu5Ac as example on microplate onto gangliosides with STs (Nakamura *et al.*, 1991) or with bacterial polyST to generate polySia onto G_{D3} (Cho and Troy, 1994). Our team also used this radiolabeled donor substrate to identify STs activity (Harduin-Lepers *et al.*, 2000 ; Vallejo-Ruiz *et al.*, 2001).

ii. Detection of sialylation side products

Nevertheless, the study of STs enzymatic activity remains limited by the lack of appropriate biochemical tools, radioactivity-free and high sensitive with low background. Notably, strategies are based on colorimetric assays like the universal phosphate assay for GTs developed in 2011 (Wu *et al.*, 2011) and currently commercialized by R&D Biotech. Briefly, it relies on the detection of the quantity of CMP side product released from CMP-Sia donor after Sia transfer onto acceptor substrate. This kit contains phosphatase which cleaves CMP into Cytidine and inorganic phosphate (Pi) residues, then Pi reacts with Malachite green and intensity can be measured by colorimetry at 620 nm. This assay was used recently to measure mouse, zebrafish and human ST6Gal I activities (Houeix and Cairns, 2019). However, it suffers from important limitations as it is an indirect detection and it is not suitable for non-purified enzymes or complex media including phosphate.

iii. Use of tagged-Sia and lectin detection

In vitro enzymatic assays were developed this last ten years with bioorthogonal labeling and fluorophore-conjugated CMP-Sia transferred onto asialofetuin using human STs (Wu *et al.*, 2019), using BODIPY fluorescent-tagged CMP-Sia transferred onto asialofetuin and asialo- α 1 anti-trypsin with human ST3Gal I and ST6Gal I and bacterial Cst-II (Abukar *et al.*, 2021). Microplate-based assay was also developed to evaluate kinetic parameters of various bacterial α 2,6-STs and human ST6Gal I with CMP-Neu5Ac. Then, biotinylated-SNA lectin was used to bind Neu5Ac transferred by α ,6-linkage onto asialofetuin, biotins were revealed with streptavidin-peroxidase and finally

detected by chemiluminescence (Janesch *et al.*, 2019b). As previously described, iEndoN could be used to detect polySia. Consequently, plate-based assay was developed with bacterial polySTs and CMP-Neu5Ac to detect polyNeu5Ac formed onto coated oligosaccharides acceptor using iEndoN-GFP (Yu *et al.*, 2014). Nevertheless, this assay is so far limited to one bacterial polySTs and have to be optimized to study other enzymatic activity like for vertebrate STs.

iv. Detection of sialylation by analytical chemistry

Electrochemical analyses were also applied to study sialylation processes and STs activity. Platform based on impedimetric biosensing and electrochemical impedance spectroscopy (EIS) is currently optimized to monitor desialylation and sialylation processes, allowing attachment of N-glycans on a glassy carbon electrode surfaces and generating sialylation with bacterial α 2-6-STs and CMP-Neu5Ac (Alshanski *et al.*, 2021 ; Alshanski *et al.*, 2022). Likewise, acceptor substrates linked to CUPRA-tag allowed measurement of STs activity with electrospray ionization mass spectrometry (ESI-MS) (Li *et al.*, 2020). CUPRA-ZYME was recently adapted to measure kinetic parameters and substrate specificities of ST6Gal I onto different N- and O-glycosylproteins like fetuin, asialo- α 1 anti-trypsin and asialo-orosomucoid (Li *et al.*, 2020 ; Li *et al.*, 2021).

v. Use of unnatural modified Sia

First modifications of Sia were explored in the 1980s by Brossmer and collaborators that synthesized various 9-substituted Sia derivatives and used them *in vitro* as donor substrates with bovine brain CSS. They detected quite comparable affinity for Neu5Ac and Sia derivatives, particularly with 9-azido-Neu5Ac ($K_m = 1.2 - 1.4 \mu\text{M}$) (Gross *et al.*, 1987). Among these derivatives, 9-fluoresceinyl-Neu5Ac was activated with CSS and its use as donor substrate for sialylation reaction was compared with CMP-Neu5Ac. Interestingly, K_m values obtained for five different purified STs was comparable in magnitude for both donor substrates (Gross *et al.*, 1990).

Some years later, Reutter and Bertozzi groups have investigated chemical reporter strategy particularly with N-propanoyl precursors administrated in rat which led to N-propanoyl-Sia detected in different rat organs (Kayser *et al.*, 1992) and with N-levulinoyl mannosamine (ManLev) to detect N-levulinoyl Sia (SiaLev) at the cell surface of three human cell lines (Mahal *et al.*, 1997).

Click-chemistry concept was introduced by Sharpless and co-workers in 2001 to describe chemical reactions which must be modular, stereoselective, water-tolerant, give very high yields, generate inoffensive byproducts and require only simple reaction conditions thermodynamically advantageous with no protective groups required (Kolb *et al.*, 2001). Remarkably, Nobel Prize 2022 in Chemistry has been awarded to Pr Meldal, Pr Bertozzi and Pr Sharpless for their contribution about bioorthogonal chemistry and click-chemistry. Among these reactions, Copper-catalyzed azide–alkyne cycloaddition (CuAAC) are based on copper reduction and allowed bioorthogonal conjugation of Sia residues with modified probes like azido-biotin (Fig. 28).

In the laboratory, our glycochemists optimized the synthesis of alkyne-C5-modified-Sia (SiaNAI) and showed that SiaNAI could be activated in CMP-SiaNAI with integrated workflow that takes less than an hour (Gilormini *et al.*, 2016a). In the team, we took advantage of this chemo-enzymatic synthesis of CMP-SiaNAI to develop enzymatic assay to study STs activity. Notably, we showed that the activated donor CMP-SiaNAI could be tolerated by human ST3Gal I and ST6Gal I with the same kinetic parameters than CMP-Neu5Ac with the substrate specificities onto O- and N-glycosylproteins respectively (Noel *et al.*, 2017). Moreover, we used CuAAC ligation to detect specifically the transferred SiaNAI by coupling it with azido-biotin (Noel *et al.*, 2017).

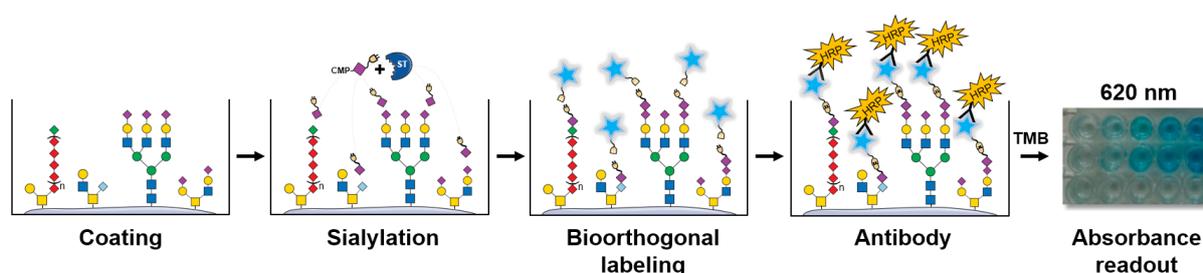


Figure 28: MPSA enzymatic assay with different steps described to detect sialylation proportional to colorimetric intensity detected at 620 nm. Firstly, acceptor substrate is coated on a microplate well. After saturation, acceptor glycoconjugate is sialylated using ST and CMP-SiaNAI. SiaNAI transferred is coupled to a biotin by click-chemistry reaction which is then recognized by an anti-biotin antibody coupled to horseradish peroxidase. Finally, the chromogenic substrate TMB recognized by peroxidase takes on a blue color whose absorbance is measured at 620 nm.

Recently, we developed a new *in vitro* rapid, specific and efficient plaque sialylation assay (MPSA) (Noel *et al.*, 2018) which allows a highly sensitive detection of sialylation using chemo-enzymatically synthesized non-natural donor substrate CMP-SiaNAI and CuAAC. The principle

of this assay is illustrated in Figure 28. To be noted, MPSA allows the use of crude STs produced in the culture medium of transfected cells to obtain significant enzymatic activity onto various acceptor substrates (Noel *et al.*, 2018). Moreover, we demonstrated the impact of different ions on enzymatic activity of ST3Gal I and ST6Gal I. Interestingly, Zn^{2+} have higher inhibitory effects at 100 μ M on ST6Gal I compared to ST3Gal I activity (Noel *et al.*, 2018). Likewise, MPSA was adapted to compare human ST8Sia VI and zebrafish ST8Sia VIII activities with CMP-SiaNAI onto different N- and O-glycosylated acceptor substrates (Chang *et al.*, 2019).

E. ANALYTICAL APPROACHES

Complementary to biochemical approaches like enzymatic *in vitro* assays, use of antibodies or lectins, the rise of analytical approaches allowed these 20 last years to bring out diversity and complexity of sialylated glycoconjugates, notably polysialylated proteins which upset beliefs that polySia is a rare event (Sato *et al.*, 2002 ; Sato, 2004). Due to its nature of long anionic polymer, polySia is difficult to study as it interferes with protein mobility and antibody detection. In addition, it non-specifically binds proteins and is difficult to analyze *via* mass spectrometry (MS) (Galuska *et al.*, 2010b). Although, isolation and enrichment of polySia is not necessary for immunoassays (Guo *et al.*, 2019), analytical detection and enzymatic assays requires improved sensitivity, particularly when there is a low-abundance of polySia (Guo *et al.*, 2019). Moreover, polySia is characterized by a fragility of internal α 2-8 glycosidic bonds leading to polySia degradation. A recent review summarized all available approaches for extraction and characterisation of polySia from biological matrices (Guo *et al.*, 2019). The most efficient methods to detect polySia are fluorescent labeling and analytical approaches although they display advantages and limits (Nishimura *et al.*, 2014).

Periodate fluorometric C7/C9 analysis of polySia is definitely the most sensitive and selective analytical tool for internal Sia residues measurements of oligo/polySia from glycoproteins (Sato *et al.*, 1998 ; Guo *et al.*, 2019). Briefly, sodium periodate and borohydride treatments of polySia lead to C7-Sia for the terminal Sia at the non-reducing end of polySia. After full hydrolysis, oxidized periodate treatment, all free Sia are derivatized with 1,2-diamino-4,5-methylenedioxybenzene (DMB) (Fig. 29) and Sia-labelled residues can be analyzed on anion-exchange chromatography or

high-performance liquid chromatography (DMB-HPLC) (Galuska *et al.*, 2011). This method allowed the evaluation of total internal Sia in comparison to terminal Sia with the corresponding fluorescence peak areas of C7 and C9 Sia (Sato *et al.*, 1998) and has the advantage to be relatively sensitive, with the detection of low amounts of internal Sia residues until ng.

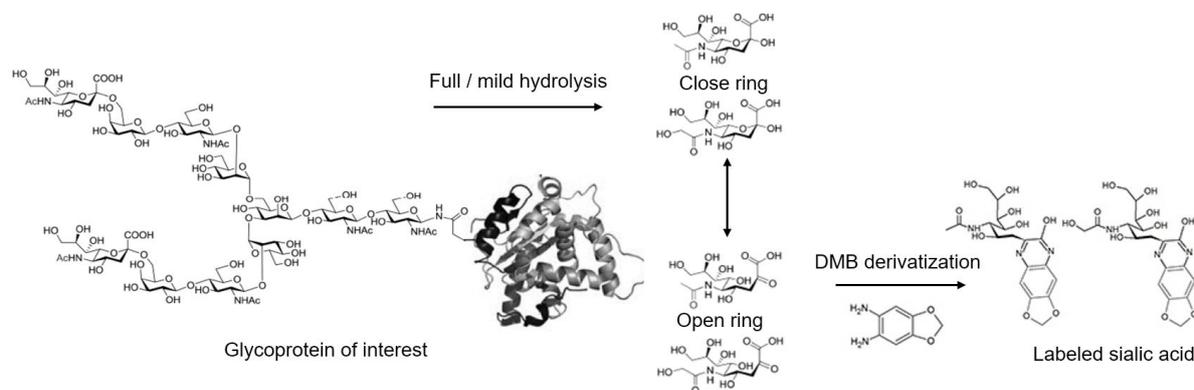


Figure 29: Process of DMB labelling of Sia prior analytical approaches. Glycoprotein of interest is treated for full or mild hydrolysis to release free Sia or oligo/polySia chains respectively. Then, open ring forms could be derivatized with DMB fluorescent compound and detected on HPLC analysis.

DMB-HPLC could also be used after mild-acid hydrolysis to quantify polySia chains length with a DP detectable of 60 Sia residues at least from biological matrices (Guo *et al.*, 2019) although accelerated intramolecular self-cleavage of polySia have been observed under these conditions (Manzi *et al.*, 1994). Moreover, acidic conditions could increase polySia lactonization resulting from condensation of the carboxyl group in C1 position of one Sia with the hydroxyl group in position C9 of the adjacent Sia residues (Kakehi *et al.*, 2001). To be noted, polylactonization could generate resistance to acid hydrolysis reducing charge and flexibility of polySia (Zhang et Lee, 1999). However, almost all *in vivo* polySia is lactone free at neutral pH (Azurmendi *et al.*, 2007).

Other analytical approaches based on charged groups allowed good detection as fluorescent anion-exchange chromatography analysis (FAEC) or with UV (HPAEC-UV) but requires higher quantities as described with at least 10 μg of purified polySia. Reverse Phase Liquid Chromatography (RPLC-FD) is not suitable for polySia as they are very sensitive to acidic pH and high temperature and can be lactonized and/or hydrolyzed (Galuska *et al.*, 2010b).

On the other hand, Western Blot (WB) can be used for polySia detection with specific antibodies (Table 15) in biological samples. However, these strategies are indirect and semi-quantitative methods because of low binding of long polySia chains compared to short chains onto PVDF membranes (Nishimura *et al.*, 2014). As mentioned previously, iEndoN could also be used to binds polySia with strong affinity, without hydrolytic activity (Guo *et al.*, 2019) allowing notably detection of polySia cell surface by flow cytometry (Jokilammi *et al.*, 2007). Some ELISA strategies were also developed to quantify polySia from serum, proteins and extravesicles (Tajik *et al.*, 2020). Alternative study of polySia combining X-ray crystallography, NMR and docking allowed the way of binding onto human viral pathogen and how virus interact with polySia through transient electrostatic interactions (Lenman *et al.*, 2018).

Enrichment methods should be explored in the future for qualitative and quantitative assesment of polysialylated glycoconjugates as the use of mAb735 or iEndoN on column (Sato, 2004 ; Sato, 2013) and/or by click-chemistry, lectin affinity and immunoaffinity prior to analyse them by sensitive strategies (Song and Mechref, 2015). Development of more rapid, highly-sensitive techniques to conduct accurate quantitative analysis remain important goals for the future, particularly when challenged by low polySia abundance, without affecting polySia chemical structure integrity and DP (Guo *et al.*, 2019).

F. SIALYLATION LABELLING IN CELLULO AND ON CELL SURFACE

Bertozzi's group has demonstrated the incorporation of unnatural derivative of Sia precursor with a ketone group, converted to the corresponding Sia residues and displayed on cell surface sialoglycoconjugates (Mahal *et al.*, 1997). In order to detect polySia onto NT2 neurons, they feed cells with ManLev and detected SiaLev at the cell surface after bioorthogonal coupling with an hydrazine-biotin (Charter *et al.*, 2000). However, modification on unnatural Sia have a great importance as they shown that N-propanoyl Sia preserve polySia synthesis in HeLa cells whereas N-butanoyl and N-pentanoyl modified Sia are inhibitors of these syntheses (Charter *et al.*, 2002).

Modified precursors of Sia have been oftenly used for glycoconjugates detection as G_{M3} on Jurkat cell surface (Bussink *et al.*, 2007). Also called metabolic glycoengineering and

oligosaccharide engineering (MGE and MOE) (Gilormini *et al.*, 2018), this method is regularly used in our lab with Sia precursor ManNAc modified with an alkyn (ManNAI) or azyde (ManNAz) group for bioorthogonal labeling on cell surfaces (Gilormini *et al.*, 2016b ; Noel *et al.*, 2018 ; Scache *et al.*, 2022 ; Rigolot *et al.*, 2023). In the future, MGE could be a promising strategy for therapeutics as in regenerative medicine to detect and target specific cells. A study on mouse model after implantation of alkyn-modified biofilm has shown an efficient capture of azido-labelled monocytes on the biofilm, 6 days after intravenous injection of monocytes (Mao *et al.*, 2019).

Nevertheless, it seems important to note that so far no data shows how these incorporated unnatural Sia are recognized and taken up by endogenous STs and that this is currently a real lack to understand the functioning of this enzymatic machinery.

On the other hand, other strategies tend to focus on enzymatic process outside cell context as chemoenzymatic glycoengineering (CGE) to label, visualize and modify specifically glycans at the cell surface using non natural Sia (Almaraz and Li, 2017) and employing STs specificities into more “physiological-like conditions” than *in vitro*.

Indeed, this process of exo-enzymatic labeling is quite similar to extrinsic sialylation occurring physiologically in the periphery as described in mice model (Manhardt *et al.*, 2017). The same strategy was also named Selective Exo-Enzymatic Labeling on cell surface (SEEL) to label cell surface N-glycans using human and rat ST6Gal I (Mbua *et al.*, 2013 ; Sun *et al.*, 2016) and recently to imaging lipooligosaccharides of pathogens (Fig. 30) (de Jong *et al.*, 2022).

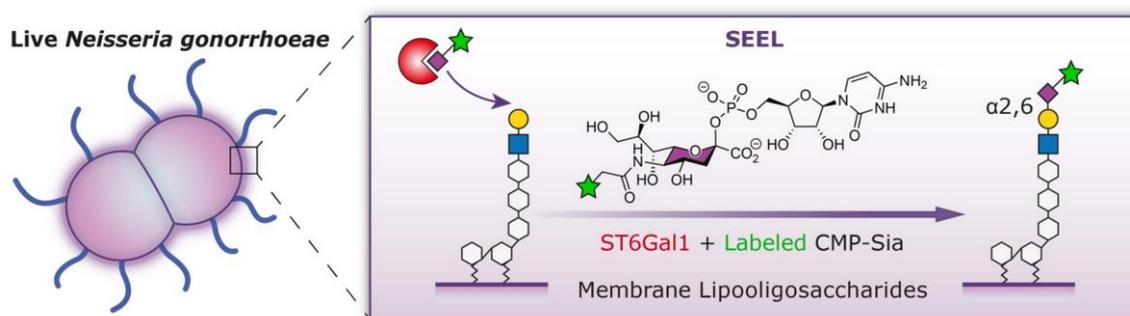


Figure 30: Selective ExoEnzymatic Labeling (SEEL) strategy adapted for modification and imaging of bacterial lipooligosaccharides.

Illustration from de Jong *et al.*, 2022

Promisingly, Sackstein's group in Harvard Medical School applied CGE strategy to modify cell surface of its CAR T-cells adding a sLe^x epitope to optimize drug delivery in the tumor beds of cancer. CGE modification significantly improve homing and targeting of CAR T-cells in mice model (Mondal *et al.*, 2019).

To summarize, *in cellulo* and *in vivo* imaging approaches of glycans on living cells are extremely interesting for the study of enzymatic activities and specificities, to determine new STs activities or of great interest in therapy to monitor sialylated glycoconjugates and remodel cell surface.

G. STs INHIBITORS DEVELOPMENT

These last few years, increasing glycan-based strategies have been tested in translational glycobiology. Glycotherapies and glycomimetics development are focused on novel cancer glycan-modified immunotherapy agents and nanoparticles, anti-glycan vaccines and neuraminidase inhibitors like Oseltamivir or Zanamivir to prevent *Influenza* infections. Indeed, the development of cell-permeable, non-toxic ST specific small-molecule inhibitors is much needed and represents a field of intense investigations (Harduin-Lepers *et al.*, 2012 ; Heise *et al.*, 2018 ; Munkley, 2022 ; Perez *et al.*, 2021 ; Rodrigues and Macauley, 2018 ; Szabo and Skropeta, 2017 ; Wang *et al.*, 2003 ; Bowles and Gloster, 2021) notably because sialylation has a key role in cancer progression and metastasis (Dobie and Skropeta, 2021).

CMP is known to inhibit all STs activity as this compound is in balance with CMP-Sia during their exchanged by SLC35A1 for CMP-Sia entrance in the Golgi apparatus. Notably, CMP inhibits ST8Sia II activity leading to a decrease of cancer cells migration *in vitro*, interesting to limit metastases dissemination (Al-Saraireh *et al.*, 2013). Moreover, natural products as soyasaponin I, ginsenosides and lithocholic acid and derivatives decreased sialylation modifying invasive behavior of tumor cells (Chang *et al.*, 2006 ; Hsu *et al.*, 2005 ; Huang *et al.*, 2020 ; Chen *et al.*, 2011).

Sialic acid mimetics (SAMs) with modified groups on CMP-Sia and Sia residues have been explored as inhibitors. Notably, N-cyclopropylcarbonyl group (Neu5Cyclo) on modified CMP-Sia is reported as inhibitor of ST8Sia II to prevent from polySia elongation (Ehrit *et al.*, 2017). A recent study also demonstrated that 8-keto-Neu5Ac decreases polySia elongation, suggesting importance

of hydroxyl group on C8 Sia residue for activity of polySTs (Hunter *et al.*, 2022). Otherwise, C-7 modification with fluoro group seems to be faster transferred by bacterial $\alpha 2,3$ -STs and improves therapeutic glycoprotein half-life, slowly degraded by sialidases with hydrolysis rate 100 times lower (Geissner *et al.*, 2021). SAMs were also developed with modification of C-2 and C-9 positions and have a better binding capacity towards Siglecs to target Sia-Siglecs axis (Büll *et al.*, 2016). Furthermore, various C-5 carbamate derivatives significantly enhanced inhibitory potency *in vitro* and *in cellulo*, particularly C-5 propargyl carbamate group (Fig. 31) (Moons *et al.*, 2022).

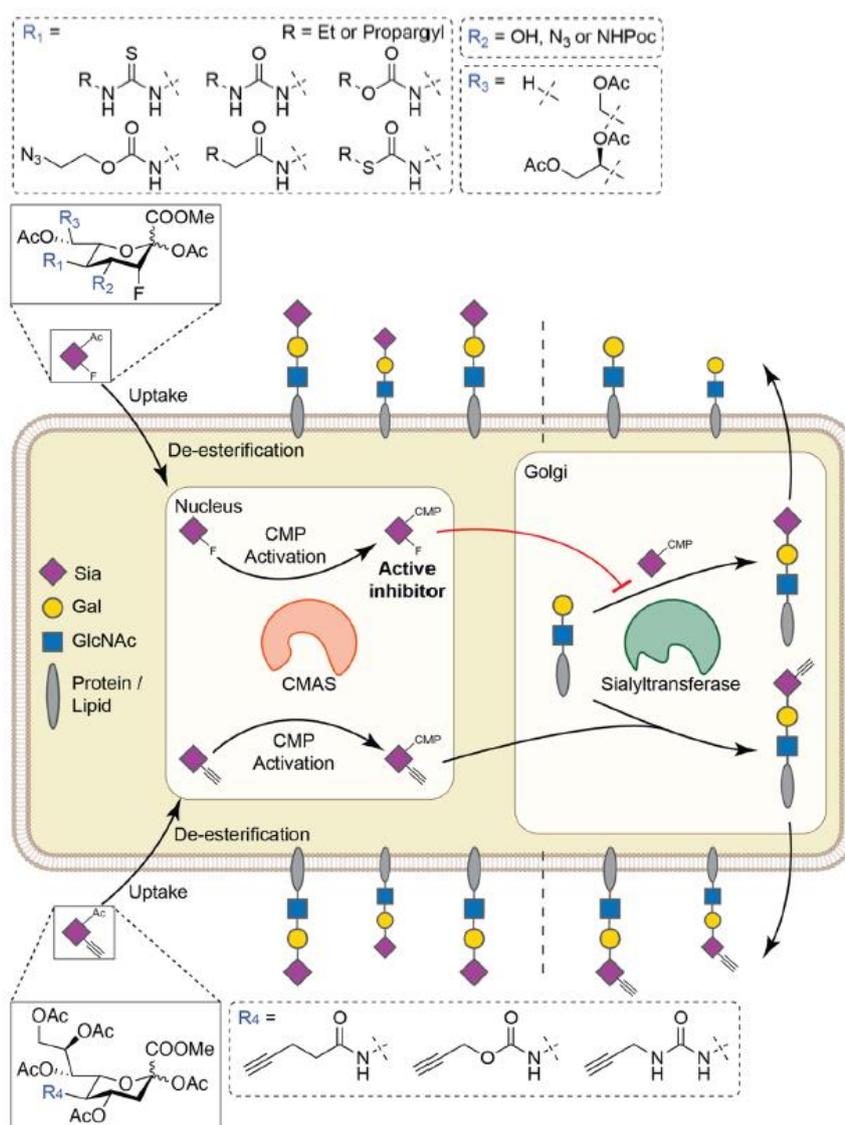


Figure 31: Structure and mode of action of metabolic sialylation inhibitors and metabolic labeling reagents. After uptake in cells, Sia derivatives are activated in the nucleus, transported in the Golgi apparatus and could have inhibitor effects if they are modified with a fluorinated group or could be transferred by STs onto glycoconjugates if they are modified with an alkyne group.

Illustration from Moons *et al.*, 2022

Currently, the most studied ST inhibitor is a peracetylated SAM named P3F_{ax}-Neu5Ac for 2,4,7,8,9-pentaacetyl-3F_{ax}-Neu5Ac-CO₂Me (Rillahan *et al.*, 2012 ; Macauley *et al.*, 2014) which has been used as drugs onto nanoparticles and shown to decrease adhesion, migration and delays tumor growth *in vivo* (Büll *et al.*, 2013). However, other type of inhibitors were developed notably based on transition state of sialylation process with oxocarbenium structure, notably with an aryl group (Skropeta *et al.*, 2004) although it appears hard and requires long times to chemically synthesize specific transition-state analogs with optimized affinity and stability (Wang *et al.*, 2003) and to improve pharmacokinetic properties with adapted linkers (Montgomery *et al.*, 2017b).

Specific inhibitors of α 2-3 and α 2-6 sialylations were also explored for drug design on transition-state molecules but remains a great challenge for the future (Perez *et al.*, 2021 ; Guo *et al.*, 2017 ; Szabo and Skropeta, 2017). Docking and molecular dynamics will be useful to optimize the design of SAMs and transition state analogues of Sia as tried with human ST8Sia III based on its crystal structure (Dobie *et al.*, 2018 ; Montgomery *et al.*, 2017b). However, the limited numbers of 3D structures available for STs and the lack of information on inhibitor effect on individual ST remains a limit for docking and designing of new selective and specific inhibitor.

H. BIOINFORMATICS TOOLS AND MODELING

These last 15 years, more and more online applications and computer softwares are developed to face questions of glycosciences, represented glycans, glycomolecules and complex structures, their involvements and functions in physiopathology (Lal *et al.*, 2020). Resulting from workshop discussions in 2014, a detailed document “A Roadmap for Glycoscience in Europe” was created to address some of the challenges faced by Europe and where glycosciences could be of major interest. In comparison to proteomics and genomics analysis, guidelines named MIRAGE (Minimum Information for A Glycomics Experiment) were proposed by the Consortium for Functional Glycomics in aim to enhance the value and reproducibility of glycoanalytical data. Some years later in 2018, the GlySpace Alliance (Lisacek *et al.*, 2022) was formed to gather glycoscience community through three principal investigators: Glyco@Expasy from Switzerland, GlyCosmos from Japan and GlyGen from USA.

During my thesis, lots of different tools, databases and software developed and shared by the scientific community inspired me and have been useful to try to answer the issues surrounding my thesis subject about STs three dimensional structures, designs, studies of activities and functions, studies of their products formed. For instance in the laboratory, we have internal tools available like our storage server dedicated to accumulate data from AlphaFold modeling and our private GT-database which collects protein sequences and subfamily classification based on sequences motifs and phylogeny studies, particularly in GT29 and GT31 families.

In addition, I proposed an exhaustive list which gathers existing informations, databases, softwares and websites dedicated for glycans and GTs (Appendix 1). Particularly, AlphaFold tool brings a new era in molecular biology to visualize, analyze and interpret structures and functions of all proteins solely from their primary sequences (Guo *et al.*, 2022). AlphaFold2 produces 3D models and also a serie of confidence metrics, such as predicted local distance difference test (pLDDT), template modelling score (pTM), and Aligned Error (PAE). There is also an extension of AlphaFold2 specifically built to predict protein-protein homodimeric and heterodimeric protein complexes (Basu and Wallner, 2016 ; Bryant *et al.*, 2022 ; Evans *et al.*, 2022). Combining 3D structural data available for STs and AlphaFold2 will be promising tools to understand STs interaction with other GTs and their integration in macromolecular complex in the Golgi apparatus.

VIII. SCIENTIFIC OBJECTIVES OF THESIS

Previous studies of polySia diversity in fish (Sato and Kitajima, 2013) and the particular distribution of polySTs gene in salmonid species (Venuto *et al.*, 2020b) were the starting point to define my thesis objectives. All these observations developed in the introductive part of my manuscript raise questions about acceptor and donor substrate recognition by polySTs and how aa substitutions in diverse species like in salmonid could impact their enzymatic specificities ? To explore these hypotheses, my thesis project was centered on the characterization of enzymatic activity of fish polySTs, to compare their specificities with their human homologs in aim to better

understand how polySTs work to generate such a biodiversity of polySia with diverse Sia occurring in the nature. The objectives were as follows:

- **Biochemical characterization of the three *Coregonus maraena* polySTs activities.**

First of all for this objective, I had to clone and produce each recombinant enzymes in cultured cells using different strategies and tools described above. In parallel, I had to adapt and optimize our innovative biotechnological tools to study polyST activity. This allowed me to determine kinetic parameters for the biochemical characterization of each of these enzymes.

- **Highlighting differences of enzymatic specificity using various donor substrates.**

From chemoenzymatically synthesis of donor substrates CMP-SiaNA1, CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn and using acceptor substrates of various nature and sensitive tools as bioorthogonal chemistry and mAb735 polySia antibody.

- **Synthesis and analysis of polySia generated by the different polySTs.**

Finally, I was able to use these recombinant enzymes to synthesize polySia chains onto various acceptors (glycoproteins fixed on microplates, in suspension or on cell surface platform) and to analyze them by different strategies (mAb735, endoN, WB, MPSA, DMB-HPLC, SEEL) to confirmed their differences of enzymatic specificities responsible for polySia biodiversity in salmonid. Particularly, polySia analyses were conducted in collaboration with our german partner expert in polySia analyses.

- **Explore the impact of various factors on enzymatic activities of polySTs.**

I studied autopolysialylation process producing enzymes in different cell types and measuring influences on their productions and activities. I used strategies of computational structural biology to model polySTs 3D structure with AlphaFold2 and highlight importance of conserve peptide motifs or aa for a potential cooperation between enzymes.

RESULTS

I. SALMONID POLYSIALYLTRANSFERASES SUBSTRATE SPECIFICITIES: CONTRIBUTION TO DIVERSITY OF FISH POLYSIA?

To better understand the structure-function relationships and substrate specificities of salmonid polySTs, each polyST gene identified in the *C. maraena* genome was cloned and produced as recombinant and active enzymes prior exploring their enzymatic activities.

A. PRODUCTION OF THE *CMA* POLYSIALYLTRANSFERASES

i. Analysis of the *Cma* polySTs sequences

Using a BLAST approach on NCBI databases, three polySTs *st8sia2-r1*, *st8sia2-r2* and *st8sia4* genes were previously identified in the salmonid whitefish (*Coregonus maraena*, *Cma*) genome (Venuto *et al.*, 2020b). The deduced protein sequence of *Cma* ST8Sia II-r1 and *Cma* ST8Sia II-r2 shared 63.8% and 62.7% sequence identity with their human homolog *Hsa* ST8Sia II whereas there is 78.8% sequence identity between *Cma* ST8Sia IV and *Hsa* ST8Sia IV. To compare protein sequences, a multiple sequence alignment (MSA) was performed with the three salmonid polySTs (378, 378 and 357 aa) and the two human polySTs (375 and 359 aa) (Fig. 32).

In comparison to human polySTs, salmonid polySTs are also composed by the four sialylmotifs L, S, III and VS conserved in all GT29 STs (Harduin-Lepers, 2010) and the PBR and PSTD motifs, characteristic of polySTs (Nakata *et al.*, 2006; Colley *et al.*, 2014) (Fig. 32). Further sequence analyses using TMHMM, indicated the presence of a hydrophobic region likely corresponding to the transmembrane (TM) domain and, using NetNGlyc (Appendix 1), the presence of six, six and five potential *N*-glycosylation sites for *Cma* ST8Sia II-r1, *Cma* ST8Sia II-r2 and *Cma* ST8Sia IV respectively. Among them, two positions are conserved between the five polySTs and, particularly in the PBR motif, one is described to be an autopolysialylation site for the two human polySTs (Asn⁸⁹ for ST8Sia II and Asn⁷⁴ for ST8Sia IV) (Table 5). Described in chapter VI-D-i to be essential for polyST structure stability and enzymatic activity, the two pairs of cysteine located in SM L, SM S and in C^{ter} are conserved for the three coregone polySTs (Fig. 32).

		$\Delta 28$ $\Delta 35$ $\Delta 50$ ↓ ↓ ↓	
	TM		
<i>CmaST8SIA2_r1</i>	MQLEFRTVMFGIVTLLVIFLIADIAEIEEEIA-NIGGSRTLY--LHSLIPKPNR-NVAVKANPTPLISEGEDKSPA	73	
<i>CmaST8SIA2_r2</i>	MQLEFRTLTMFGIVTVLVIFLIADIAEVEEEIA-NIGGSRKLY--MHSLIPKPNR-NVAVKANPKPLVSEGEDKSPA	73	
<i>HsaST8SIA2</i>	MQLDQRSWMLAALTLLVVFLIFADISEIEEEIG-NSGGRGTIRSAVNSLHKSNSRAEVVINGSSSPAVVDRSNES--	74	
<i>CmaST8SIA4</i>	MRLSRKRWTICTISILVIFYKTKETRSEEHQEAQVTGDSELD-----TSRLMV-----NSSEKSSR	57	
<i>HsaST8SIA4</i>	MRSIRKRWTICTISLLLIIFYKTKETIARTEEHQETQLIGDGELS-----LSRSLV-----NSSDKIIR	57	
	PBR		
<i>CmaST8SIA2_r1</i>	SPSGLNNTTRLSSDNWTFNRTLSSLIRKNILRFFDPERDISILKGTLPKGDVIHYIFDRQSTTNISENLYRLLPTAS	150	
<i>CmaST8SIA2_r2</i>	SPSYSNNTTKLSSDNWTFNRSLSNSIGKNILRFFDPERDISILKGTLPKGDVIHYIFDRQSTTNISENLYRLLPTVS	150	
<i>HsaST8SIA2</i>	----IKHNIQPASSKWRHNQTLRLRIRKQILKFLDAEKDISVLKGTLPKGDIIHYIFDRDSTMNVSQONLYELLPTS	147	
<i>CmaST8SIA4</i>	---SGPSFFQHSVEGWRLNSSLVLMIRKDVLRFLDAERDVSVVKSSFKPGDTIHYVLDRRRTLNI SHTLHSLLPDVS	131	
<i>HsaST8SIA4</i>	K--AGSSIFQHNVEGWKINSSLVLEIRKNILRFLDAERDVSVVKSSFKPGDVIHYVLDRRRTLNI SHDLHSLLPDVS	132	
	SM L		
<i>CmaST8SIA2_r1</i>	PMKNQHRRRC CAIVGNSGILLNSSCGPEIDSHDFVIRCNLAPVEEYAGDVGRRTNLV TMNPSVVQRAFHDLASEQWRE	227	
<i>CmaST8SIA2_r2</i>	PMKNQHRRRC CAIVGNSGILLNSSCGPEIDSYDFVIRCNLAPVEEYAGDVGRRTNLV TMNPSVVQRAFQDLASEEWRE	227	
<i>HsaST8SIA2</i>	PLKNKHFGT CAIVGNSGVLLNSGCGQEIDAHSFVIRCNLAPVQEYARDVGLKTDLV TMNPSVIQRAFEDLVNATWRE	224	
<i>CmaST8SIA4</i>	PLKNKRFKT CAVVGNSGVLLNSGCGKEIDRHDFVIRCNLAPLAEFAEDVGLRSDFT TMNPSVIQRVYGGGLKNATDTE	208	
<i>HsaST8SIA4</i>	PMKNRRFKT CAVVGNSGILLDSECGKEIDSHNFVIRCNLAPVVEFAADVGTKSDFI TMNPSVVQRAFGGFRNESDRE	209	
	PSTD SM S		
<i>CmaST8SIA2_r1</i>	RFLQRLRGLSGSVLWIPAFMAKGGEERVEWAIRLII LHTVDVHTAFPSLRLLIHAVRGYWLTNNVQIKR PTTGLLMYT	304	
<i>CmaST8SIA2_r2</i>	RFLQRLRGLSGSVLWIPAFMAKGGEERVEWAIRLII LHTVDVHTAFPSLRLLIHAVRGYWLTNNVQIKR PTTGLLMYT	304	
<i>HsaST8SIA2</i>	KLLQRLHSLNGSILWIPAFMARGGKERVEWVNELII KHHVNVRTAYPSLRLLIHAVRGYWLTNKVHIKR PTTGLLMYT	301	
<i>CmaST8SIA4</i>	RFVQRLRMLNDSVLWIPAFMVKGGERHVESVNELIV KRKLVRTAYPSLRLLIHAVRGYWLTNKINIKR PSTGLLMYT	285	
<i>HsaST8SIA4</i>	KFVHRLSMLNDSVLWIPAFMVKGGEKHVEWVNALII KNKLVRTAYPSLRLLIHAVRGYWLTNKVPIKR PSTGLLMYT	286	
	SM III SM VS		
<i>CmaST8SIA2_r1</i>	MATRF C EEIHLYGFWPFPRDSQGI PVKYHYD TLTYEYTSHAS P HTMPLEFRTLSSLHRQALRLNTGSCDAGM	378	
<i>CmaST8SIA2_r2</i>	MATRF C EEIHLYGFWPFQDSQGK PVKYHYD TLTYTYTSHAS P HTMPLEFRTLSSLHRQALRLHTGSCDAGT	378	
<i>HsaST8SIA2</i>	LATRF C KQIYLYGFWPFPLDQNPV KYHYD SLKYGYTSQAS P HTMPLEFKALKSLHEQGALKLTVGQCDGAT	375	
<i>CmaST8SIA4</i>	LATRF C DEIHLYGFWPFPRDSNGNV VKYHYD LLKYRYFSNAG P HRMPLEFKTLKMLHSGALKLTTSKCES--	357	
<i>HsaST8SIA4</i>	LATRF C DEIHLYGFWPFKDLNGKAV KYHYD DLKYRYFSNAS P HRMPLEFKTLNVLHNRGALKLTTGKCVKQ-	359	

Figure 32: Multiple sequence alignment showing the conserved domains of the salmonid and human polySTs proteins. MSA of the three polySTs sequences of *C. maraena* (*i.e.* *Cma* ST8Sia II_r1, *Cma* ST8sia II_r2 and *Cma* ST8Sia IV) with the two human polySTs sequences (*i.e.* *Hsa* ST8Sia IV and *Hsa* ST8Sia II) were performed using ClustalW (EMBL-EBI). Numbering of the last aa per sequence is indicated at the end line. The transmembrane (TM) domain appears with the 20 aa framed in black on N^{ter}. The five and six potential *N*-glycosylation (NxS/T) sites for ST8Sia II and ST8Sia IV respectively are shown in bold (Mühlenhoff *et al.*, 2001) and cysteins involved in disulfide bonds are underlined. Sialylmotifs L, S, III and VS common to all animal STs (Harduin-Lepers, 2010) are framed in light green background. The specific domains PBR and PSTD of the polySTs (Colley *et al.*, 2014) are in white letters with light brown background. Finally, the black arrow indicates the place where proteins were truncated to obtain the *Cma* Δ28ST8Sia IV, the *Cma* Δ28ST8Sia II-r1, Δ35ST8Sia II-r1, Δ50ST8Sia II-r1 and the *Cma* Δ28ST8Sia II-r2, Δ35ST8Sia II-r2, Δ50ST8Sia II-r2 recombinant proteins.

Protein sequences were submitted to compute pI/MW on ExPASy (Appendix 1) to determine their theoretical molecular weight (MW), without considering any post-translational modification such as glycosylation. Indeed, *Hsa* ST8Sia IV and *Cma* ST8Sia IV display a MW of 41.3 and 41.1 kDa whereas *Hsa* ST8Sia II, *Cma* ST8Sia II-r1 and *Cma* ST8Sia II-r2 display a MW of 42.4, 42.9 and 42.9 kDa. All these informations have been further considered for the molecular cloning and the production of active salmonid polySTs.

ii. Molecular cloning of polySTs in expression vectors

Firstly, we have shown that polySTs genes exhibit a distinct expression profile in tissues. Indeed, *st8sia2-r1* and *st8sia2-r2* genes displayed their highest mRNA expression level in gonads and telencephalon and in gonads and spleen respectively, whereas *st8sia4* gene is only expressed at low mRNA levels in the brain, head kidney, gills, gonads and spleen (Venuto *et al.*, 2020b).

As detailed in “Material and Methods” section, we performed PCR to amplify cDNA of the polyST genes from *C. maraena* gonads and brain using Q5 polymerase and sense and antisense oligonucleotide primers containing *Hind*III and *Kpn*I restriction sites respectively. *St8sia2-r1* and *st8sia4* cDNAs were successfully amplified, recovered and inserted into the multiple cloning site (MCS) of the p3×FLAG-CMV10 vector (Appendix 2) using *Hind*III and *Kpn*I restriction enzymes. Indeed, this vector is oftenly used as an expression vector for the monitoring of full-length recombinant 3xFLAG-tagged proteins using an anti-FLAG antibody (Noel *et al.*, 2017).

Since no cDNA amplification was detected for *st8sia2-r2* gene in gonads and spleen, it prompted us to synthesize its full length sequence chemically into pcDNA3.1 vector using GeneArt Instant Designer (ThermoFischer). To achieve functional analyses, soluble forms of the three enzymes lacking their cytoplasmic and TM domains were constructed in the p3×FLAG-CMV9 expression vector, which allows secretion of fusion proteins (Appendix 2). From MSA analysis between polySTs (Fig. 32), we removed the first 28 aa residues ($\Delta 28$) on N^{ter} side of the three polySTs, an interesting position located in the stem region which preserves the catalytic domain of each polyST.

Consequently, we designed primers (Appendix 3) and used the previous constructs as DNA templates (i.e. p3×FLAG-CMV10 for *st8sia4* and *st8sia2 r-1* and pcDNA3.1 for *st8sia2 r-2*) to amplify and insert $\Delta 28$ truncated forms of the three polySTs in the p3xFLAG-CMV9 vector.

In parallel, we designed two other more truncated forms ($\Delta 35$ and $\Delta 50$) (Fig. 32) in the stem region for *Cma* ST8Sia II-r1 and *Cma* ST8Sia II-r2. Indeed, ProtParam tool from ExPASy (Appendix 1) provides instability index (II) resulting from statistical analysis of dipeptides composition in protein sequences. Although any polyST sequence could be considered as stable proteins (II < 40.00) , $\Delta 28$ ST8Sia IV appeared more stable than FL ST8Sia IV whereas $\Delta 35$ and $\Delta 50$ deletions for ST8Sia II-r1 and ST8Sia II-r2 displayed better II than $\Delta 28$ forms (Table 16).

Table 16: Stability of the recombinant *Cma* polySTs sequences designed. Instability index (II) were determined for 3xFLAG-tagged full-length (FL) and 3xFLAG-tagged truncated (Δ) polySTs sequences.

	<i>Cma</i> ST8Sia II-r1				<i>Cma</i> ST8Sia II-r2				<i>Cma</i> ST8Sia IV	
Δ	FL	28	35	50	FL	28	35	50	FL	28
II	47.70	49.08	46.08	47.26	49.35	52.55	49.61	48.30	45.81	40.52

To summarize, the study of *Cma* polySTs sequences and comparison of their characteristics with their human homologs allowed us to design interesting constructs in expression vectors to produce the three polySTs and to study their enzymatic activity after cellular transfections.

iii. Cellular transfection for Cma polySTs production

Initially, we produced full length tagged *Cma* polySTs from p3xFLAG-CMV10 vectors, however this strategy didn't lead to sufficient production of active enzymes (data not shown).

In contrast to CMV10 vector, constructs done in CMV9 vector allowed the production of truncated and soluble forms of the recombinant polySTs, secreted outside the cells, in order to obtain sufficient quantities after cell transfection to study their enzymatic activity separately.

Thus, we decided to optimize the secretion of *Cma* polySTs using lipofectamine for transiently cellular transfection in different cell types and varying different parameters of transfection (Fig. 33). Since no good transfection yield could be achieved in the homologous salmonid fish cell line CHSE-214 system as shown for *Cma* Δ 28 ST8Sia IV (Fig. 33A), secretion of the recombinant proteins was optimized outside transiently transfected HEK293 cells (Fig.33 B), which remain the preferred host to produce difficult-to-express GTs (Moremen *et al.*, 2018). To be noted, CHO-K1 cell system was also developed by R&D for the commercial production of human polySTs. Heterologous transfection of *Cma* Δ 28 ST8Sia IV in CHO-K1 led to lower level of secretion than in HEK293 cells, with only 20% of enzymes detected on WB using anti-FLAG antibody (Fig. 33A).

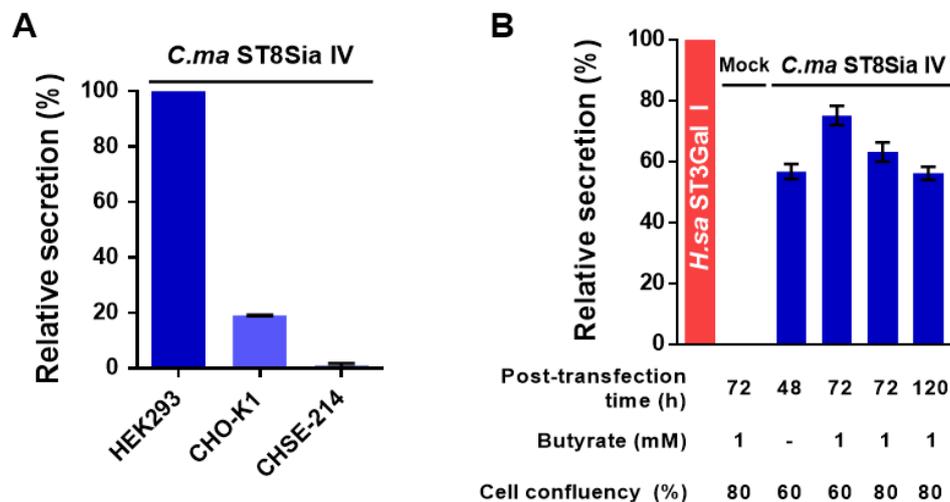


Figure 33: Optimized conditions for the production of *Cma* ST8Sia IV. (A) In different cell types. p3×FLAG-CMV9 Δ 28ST8Sia4 plasmid was transiently transfected using lipofectamine on CHO-K1, HEK293 and CHSE-214 cells. Media culture were collected 72 hours post-transfection, loaded on SDS-PAGE and analyzed onto WB. Bands intensity was measured with ImageJ and normalized towards HEK293 production used as standard. Data were reported onto GraphPad Prism 6. (B) With different parameters. The relative secretion rate of *C.ma* Δ 28ST8Sia IV transfected in HEK293 cells was investigated using different initial cell confluence levels, post-transfection collection times (24, 48, 72 and 120 h) and butyrate supplementation (1 mM). The intensity of the signals obtained was analyzed by densitometry (ImageJ) in comparison to a previous ST3Gal I production used as a control. These figures are representative to 3 biological replicates.

In order to obtain the highest quantities of recombinant enzymes in the medium culture, we investigated different parameters for cellular transfection in HEK293 cells. We used different initial levels of cell confluency to improve transfection efficiency, varied post-transfection collection times and supplemented medium culture using sodium butyrate to stimulate the production of integrated proteins (Bassens *et al.*, 2005). Reproduced three times for each parameter of transfection, the best conditions allowed us to obtain *Cma* Δ 28 ST8Sia IV secreted around 75% compared to our standard (Fig. 33B), a recombinant 3xFlag-tagged human ST3Gal I used previously in our enzymatic assays (Noel *et al.*, 2017 ; Noel *et al.*, 2018).

In parallel, we analyzed the quality of the enzymes produced. After cell transfection in HEK293 cells, media culture were collected 72 hours post-transfection, cells were collected and lyzed for further protein quantification using BCA assay. A control with an empty p3xFLAG-CMV9 plasmid was produced and collected at the same time. To determine the presence of N-glycans and their nature on *Cma* Δ 28 ST8Sia IV, we used *N*-glycosidase F (PNGase-F), which removes all N-glycans by cleaving bond between Asn and first GlcNAc residues and endoglycosidase H (EndoH) which cleaves the glycosidic bond after the first GlcNAc residue of a mannoside or hybrid glycosidic chain, but complex types are preserved.

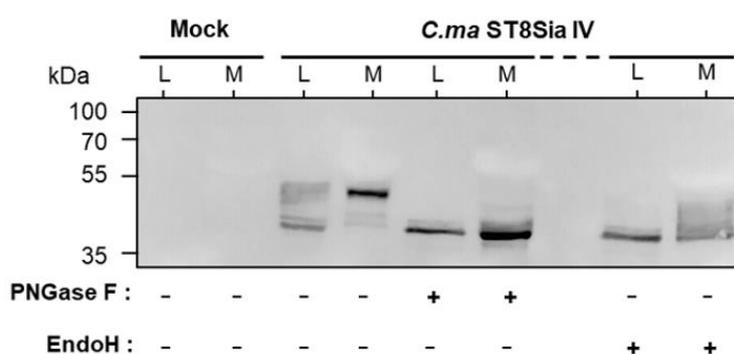


Figure 34: Quality of production for *Cma* Δ 28ST8Sia IV in transfected HEK293 cells. HEK293 cells were transiently transfected using lipofectamine with an empty p3xFLAG-CMV9 plasmid (Mock) or with the p3xFLAG-CMV9- Δ 28ST8Sia IV allowing truncated (Δ 28) form of the enzyme. 72 hours post-transfection, 7 μ g of proteins from cell lysates and 20 μ L of cell culture media were denatured in SDS buffer, Triton X-100 (10%) and treated overnight at 37°C with 20U of PNGase-F or EndoH in acetate buffer (pH 6.0, 50 mM). Samples were loaded on 8% SDS-PAGE and proteins were transferred onto nitrocellulose membrane. WB analyses were carried out using anti-3xFLAG antibody BioM2 (1 μ g/mL). Predicted molecular weight markers are indicated on the left side, cell lysate (L) and culture medium (M) are indicated on top.

20 μ L of medium culture or 7 μ g of cell lysates were run on 8% SDS-PAGE. Coomassie blue and Ponceau red stainings were respectively done on the gel and the nitrocellulose membrane after transfer, to check equal samples loading and complete proteins transfer onto the membrane. Then, we performed a WB with an anti-FLAG antibody and we detected the 3xFLAG-tagged recombinant *Cma* Δ 28 ST8Sia IV in the medium culture and in the cell lysate whereas no band was detected with the Mock control (Fig. 34).

Indeed, several bands ranging from 41 to 52 kDa were detected in cell lysate (L) whereas a major band around 52 kDa was observed in the medium (M), suggesting different glycoforms of ST8Sia IV (Fig. 34). These observations are consistent with the presence of five N-glycosylated sites onto the enzyme (Fig. 32). Moreover, PNGase-F and EndoH treatments induced a shift of the bands to the expected size of 41.1 kDa, although various bands are still present after EndoH treatment of the medium culture compared to the cell lysate (Fig. 34), suggesting that the enzyme carries both high Mannose and complex types of N-glycans.

Concerning the two other polySTs *Cma* ST8Sia II-r1 and *Cma* ST8Sia II-r2, transfections with Δ 28 forms didn't lead to enough quantities for further significant enzymatic activities. To be noted, we tried to decrease temperature to 30, 32 and 35°C during transfection to slow cell metabolism and increase secretion rate in HEK293 cells (Lin *et al.*, 2015), however no significant increase was detected (data not shown). Therefore, we designed and used other constructs (i.e. Δ 35 and Δ 50) to optimize stability of the recombinant proteins (Table 16). Compared to *Cma* Δ 28 ST8Sia IV, we obtained a maximum of 21-25% for the best productions in HEK293 cells, *Cma* Δ 35 ST8Sia II-r1 and *Cma* Δ 35 ST8Sia II-r2 (Fig. 35A ; Fig. 35B) whereas lower levels were obtained for Δ 28 and Δ 50 forms of the two enzymes (Fig. 35B). As for ST8Sia IV, we analyzed the quality of the productions for *Cma* Δ 35 ST8Sia II-r1 and *Cma* Δ 35 ST8Sia II-r2 on WB (Fig. 35C). In contrast to ST8Sia IV, we observed that both enzymes are accumulated in the cell lysates and lower secreted in the culture media. Moreover, PNGase F treatment confirms N-glycosylation of the enzymes and EndoH treatment precises a difference for their production in the media culture : *Cma* ST8Sia II-r1 carried mainly high-mannose whereas *Cma* ST8Sia II-r2 has more complex N-glycans (Fig. 35C).

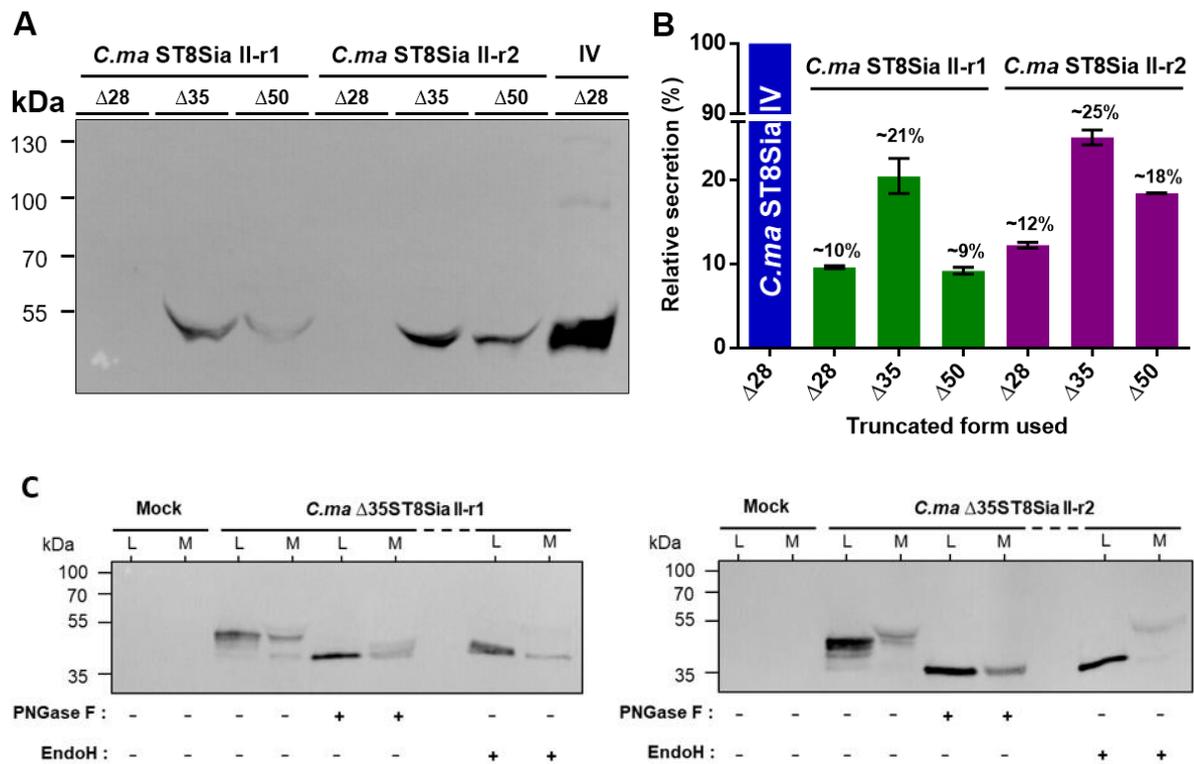


Figure 35: Production of the recombinant *Cma* ST8Sia II-r1 and ST8Sia II-r2 in HEK293 cells. Plasmids for the expression of $\Delta 28$ ST8Sia IV (blue), $\Delta 28$ ST8Sia II-r1, $\Delta 35$ ST8Sia II-r1, $\Delta 50$ ST8Sia II-r1 (green) and $\Delta 28$ ST8Sia II-r2, $\Delta 35$ ST8Sia II-r2, $\Delta 50$ ST8Sia II-r2 (purple) were transfected during 4 hours at 37°C using Lipofectamine 2000 into HEK293 cells. Culture media and cell lysates were collected 72 h post-transfection. **(A)** Representative WB with 20 μ L of each cell culture medium loaded, separated on SDS-PAGE (8 %) and transferred on nitrocellulose membrane. WB was carried out with the anti-3 \times FLAG antibody (1 μ g/mL). Molecular weight markers are indicated on the left side. **(B)** Quantification of polySTs production. Signal intensity was analyzed by densitometry (ImageJ). Secretion of each production is reported to *Cma* ST8Sia IV (blue) production as indicated above each bar (n=2). **(C)** 7 μ g of cell lysates and 20 μ L of culture media of HEK293 transfected cells with $\Delta 35$ ST8Sia II-r1 (left panel) and $\Delta 35$ ST8Sia II-r2 (right panel) were treated overnight with 20U of PNGase F and EndoH and analyzed as described for *Cma* ST8Sia IV in Fig. 34.

To conclude on this glycosyltransferase engineering part, each step appears crucial from sequence analyses and design of the constructs to optimize and obtain the best productions of the enzymes for further biochemical analyses of enzymatic activities.

Listed below, I summarized all the constructs, productions and enzymatic activities of the various vertebrate STs used and studied during my thesis (Table 17).

Table 17: Summary table of the production and use of different truncated STs during my thesis project.

Name	Species	Plasmid	Deletion	Production level	Enzyme Activity
ST8Sia II-r1	<i>Coregonus maraena</i>	3XFLAG-CMV9	Δ28	Very low	-
			Δ35	Low / Medium	-/+
			Δ50	Very low	-
ST8Sia II-r2	<i>Coregonus maraena</i>	3XFLAG-CMV9	Δ28	Very low	-
			Δ35	Low / Medium	-/+
			Δ50	Very low	-
ST8Sia II	<i>Homo sapiens</i>	6-His	Δ23	Commercial	+
ST8Sia III	<i>Danio rerio</i>	3XFLAG-CMV9	Δ83	Very high	+
ST8Sia IV	<i>Homo sapiens</i>	6-His	Δ39	Commercial	+
	<i>Coregonus maraena</i>	3XFLAG-CMV9	Δ28	Medium	+
ST8Sia VI	<i>Homo sapiens</i>	6-His	Δ27	Commercial	+
ST8Sia VIII	<i>Danio rerio</i>	3XFLAG-CMV9	Δ33	Low / Medium	-/+
ST6Gal I	<i>Homo sapiens</i>	3XFLAG-CMV9	Δ56	Medium	+
	<i>Danio rerio</i>	3XFLAG-CMV9	Δ135	Very high	+
ST3Gal I	<i>Homo sapiens</i>	3XFLAG-CMV9	Δ56	High	+

B. PREPARATION OF DONOR AND ACCEPTOR SUBSTRATES FOR ST ACTIVITY

As described in the introductory part, STs use Sia donor substrates corresponding to an activated form of the glycosyl nucleotide (i.e. CMP-Sia). In order to study enzymatic activity of polySTs, we carried out the chemo-enzymatic synthesis of various CMP-Sia, notably the three natural major forms in vertebrates: CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn (Fig. 4). Based on previous developments in our laboratory (Gilormini *et al.*, 2016a), we have also synthesized CMP-SiaNA1, a non-natural donor substrate carrying an alkyn group. Interestingly, CMP-SiaNA1 could be coupled to a probe by bioorthogonal click-chemistry reaction for further specific detection.

We performed chemo-enzymatic synthesis of each CMP-Sia using each Sia in equimolarity with CTP and a bacterial CMP-Sialic acid synthetase (CSS), except for the synthesis of CMP-Kdn which was done with rainbow trout CSS, more efficient to activate Kdn into CMP-Kdn (Wu *et al.*, 2022).

The formation of all CMP-Sia was followed for 1 hour at 37°C by ^{31}P NMR (Fig. 36), which has the advantage to control the purity of the products formed and avoid potential degradations. Compared to the CTP control spectrum which exhibits three peaks corresponding to each phosphate group (Fig. 36), we observed a progressive decrease of these peaks during synthesis in favor to the formation of new peaks characteristic of inorganic phosphate (Pi) and CMP-Sia residues formed as previously described (Gilormini *et al.*, 2016a).

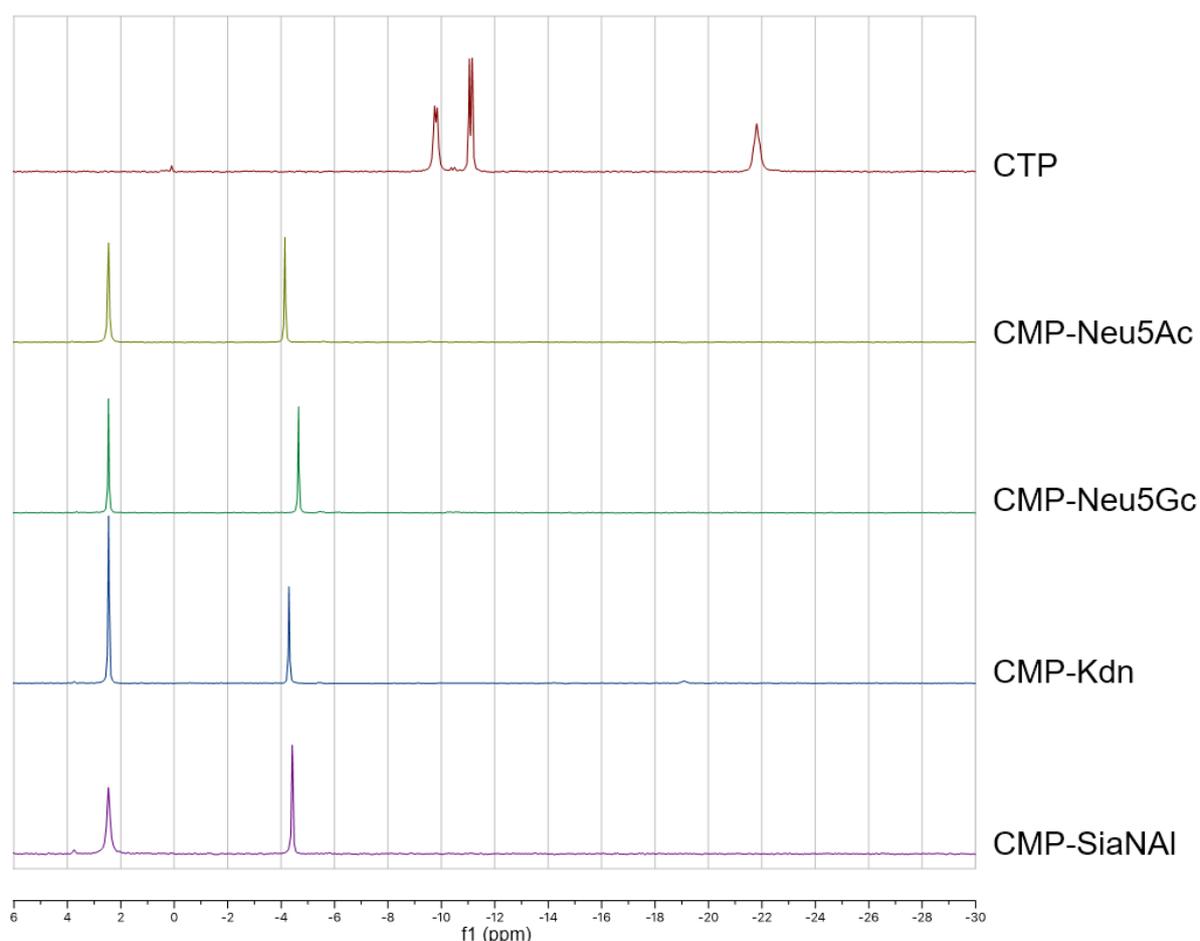


Figure 36: ^{31}P NMR spectra of chemo-enzymatically synthesized activated CMP-sialic acids. Enzymatic reactions were performed in equimolarity of CTP and each Sia (1:1) with 0.3 U CMP-Sialic acid Synthase (CSS) from *N. meningitidis* for CMP-SiaNAI, CMP-Neu5Ac and CMP-Neu5Gc and from rainbow trout for CMP-Kdn and 0.5 U PPase in Tris-HCl 100 mM, MgCl₂ 20 mM buffer (pH 8.5) during 1 h at 37 °C. Products integrity was checked by 1D NMR spectrometry of ^{31}P in a Brüker Avance II 400 MHz NMR spectrometer. Peaks shift (ppm) from CTP spectrum (red) after one hour of synthesis to CMP-Neu5Ac (olive), CMP-Neu5Gc (green), CMP-Kdn (blue) and CMP-SiaNAI (purple) spectra indicates 100 % formation and purity of the four products formed.

Furthermore, CMP-Sia could be degraded over the time at RT due to the repulsion of neighboring anionic charges present on the carboxylate and phosphate groups of this compound. Its hydrolysis leads to the release of CMP compound, reported to be inhibitor of STs activities as described in introductory part, and of free Sia which might generate aspecific detection in enzymatic assays. Therefore, once synthesized, these compounds were stored at -80°C and we ensured their stability over time. After three months of storage at -80°C , each CMP-Sia was checked again by ^{31}P NMR. As shown for CMP-SiaNAI (Fig. 37), we did not detect any significant change on the NMR spectra, confirming a good stability of our compounds stored at -80°C .

The precautions taken to synthesize CMP-Sia were important to enhance detections of polySTs activities presented below. Among these donor substrates, only CMP-Neu5Ac is commercially available but is not completely pure (Gilormini *et al.*, 2016a). Notably, my thesis work was the first to report the chemo-enzymatic synthesis of various and pure natural and unnatural CMP-Sia in order to use them as donor substrates for the study of polySTs enzymatic activities and specificities.

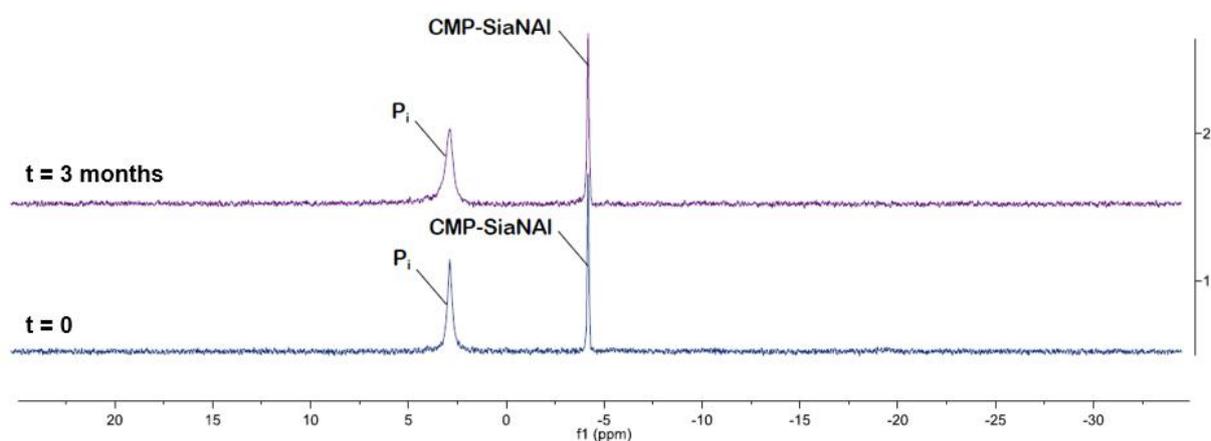


Figure 37: Analysis of CMP-SiaNAI stability by ^{31}P NMR after 3 months of storage at -80°C . The peaks represent the chemical shift (ppm) of the phosphates of P_i and CMP-SiaNAI detected by ^{31}P NMR. The spectrum at t=3 months is compared with the spectrum at t=0 after the chemo-enzymatic synthesis.

On the other hand, we investigated for potential acceptor substrates to study polySTs enzymatic activity. We prepared different types of glycoproteins coming from vertebrates origin such as bovine fetuin, which is a glycoprotein carrying three complex N-glycans of the triantennary type with sialylated LacNAc epitope (Gal β 1-4GlcNAc) and three mucin-like O-glycans (disialylated core 1: Neu5Ac α 2-3Gal β 1-3[Neu5Ac α 2-6]GalNAc-Ser/Thr) (Baenziger and Fiete, 1979),

orosomucoid and bovine submandibular mucin (BSM), carrying respectively N- and O-glycans (Table 13). In addition, we expressed a truncated and tagged version of zebrafish ALCAM/CD166 (Appendix 4), kindly obtained from Dr J. Bentrop (KIT, DEU). Indeed, ALCAM belongs to adherence molecules family, has a role in zebrafish brain development (Diekmann and Stuermer, 2009) and is notably an homolog of human NCAM oftenly associated with human polySTs activities (Mühlenhoff *et al.*, 1996a ; Close *et al.*, 2003).

Transfections of ALCAM were performed in HEK293, CHO-K1 and CHO-Lec2 cells in order to obtain various glycosylation profiles on the glycoprotein. Among these cells, CHO-K1 do not present ST6Gal activity and α 2-6 sialylation (Table 12) whereas CHO-Lec2 mutants are deficient in CMP-Neu5Ac transporter (SLC35A1) and display asialylated glycans. Each ALCAM production was collected 48 hours post-transfection in minimum medium (UltraMEM) and was coated on microplate for lectin recognition (Fig. 38). For this assay, we used various lectins like SNA, RCA, MAL I and MAL II which detect specific glycan motifs (Table 14).

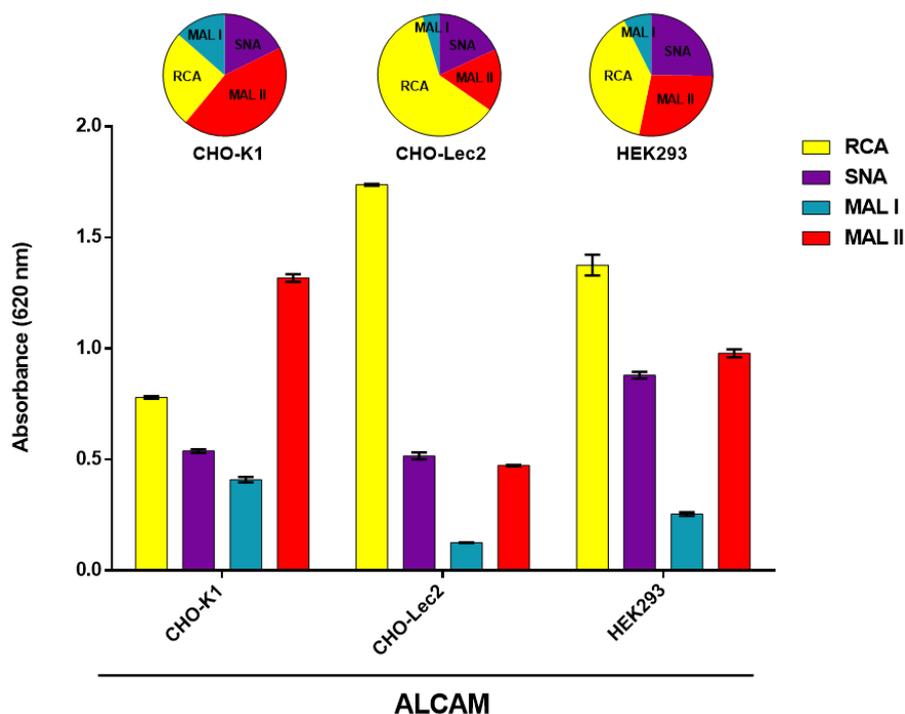


Figure 38: Lectin assay to determine ALCAM glycosylation profiles on microplates. 400 ng of diverse ALCAM versions were coated on microplate and recognized by various biotinylated-lectins described in Table 14 (RCA₁₂₀, SNA, MAL I and MAL II) for 30 min. After washing steps, anti-biotin-HRP antibody was incubated for 1 hour. After washing steps, TMB chromogenic reagent allowed colorimetric measurement at 620 nm. Parts of whole representation is used on the top for lectin recognition proportion.

Each ALCAM version was not recognized in the same way by the four lectins used. In contrast to ALCAM produced in CHO-K1, those produced in CHO-Lec2 and HEK293 are highly recognized by RCA₁₂₀ (Fig. 34), which binds Gal residues. ALCAM produced in HEK293 cells is well recognized by SNA lectin, which binds Neu5Ac α 2-6GalNAc whereas lower levels are detected for productions into CHO cells. As reported by Thermofisher company, CHO cells are not fully deficient for α 2-6 sialylation in their glycoproteins although they don't have ST6Gal enzymes, which might explain that CHO versions of ALCAM are recognized by SNA lectin. Among the three ALCAM version, the one produced in CHO-K1 cells have the highest recognition level by MAL I and MAL II lectins, suggesting an α 2,3-sialylated profile on their N-glycans and their O-glycans respectively (Fig. 38).

Finally, the preparation of all these donor and acceptor substrates allowed us to develop enzymatic assays to study the three *Cma* polySTs enzymatic activity with their human homolog.

C. DEVELOPMENT OF ENZYMATIC ASSAY FOR POLYSTS ACTIVITY

First of all, we have taken advantage of the MPSA, a rapid and sensitive enzymatic assay recently developed in our team to study human ST3Gal I, ST6Gal I and ST8Sia VI activities using CMP-SiaNAI as donor substrate (Noel *et al.*, 2018 ; Chang *et al.*, 2019).

Thus, we investigated if MPSA could be applied for polySTs activity starting with the recombinant commercial human polySTs Δ 23 ST8Sia II and Δ 39 ST8Sia IV (Table 17). We coated 400 ng of bovin fetuin on microplate and achieved sialylation reactions during 2 hours at 37°C with 100 μ M of CMP-SiaNAI and 250 ng of both human polySTs in mixture into cacodylate buffer prepared at different pH from 4.4 to 7.0. Then, azido-PEG3-biotin was covalently attached to the alkyne group of transferred SiaNAI, biotin was detected using an HRP-conjugated anti-biotin antibody and absorbance was measured after addition of the chromogenic reagent TMB (Fig. 28).

We obtained an optimal pH close to 5.8 – 6.0 for *Hsa* ST8Sia II, consistent to data obtained in a recent british study (Guo *et al.*, 2020) and a higher optimal pH for *Hsa* ST8Sia IV (Fig. 39), around 6.4. These data clearly indicated that MPSA is a good enzymatic assay to study polySTs activities.

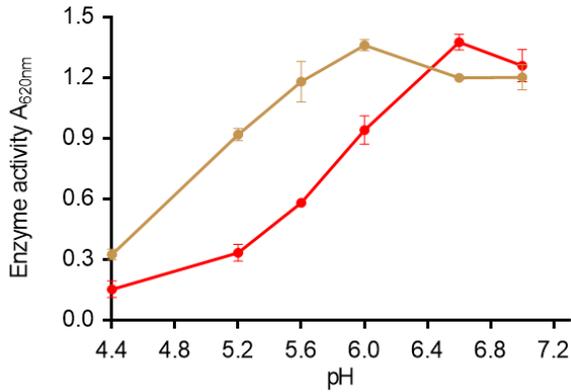


Figure 39: Use of MPSA to study human polySTs activities. After coating of 400 ng of fetuin, sialylations were achieved during 2 hours at 37°C with 100 μM of CMP-SiaNAI, 250 ng of Δ23 *Hsa* ST8Sia II (light brown) and Δ39 *Hsa* ST8Sia IV (red) in cacodylate buffer at pH from 4.4 to 7.0. After CuAAC click-chemistry with azido-PEG3-biotin and biotin detection with specific HRP-coupled antibody, TMB substrate was added and absorbance was measured at 620 nm.

Afterwards, we extended MPSA to verify the use of CMP-SiaNAI as donor substrate by the recombinant *Cma* polySTs. We coated 400 ng of fetuin and the various productions of ALCAM on microplate and we performed sialylation reactions with 100 μM of CMP-SiaNAI and using cell culture media as a crude enzyme sources for each *Cma* polyST.

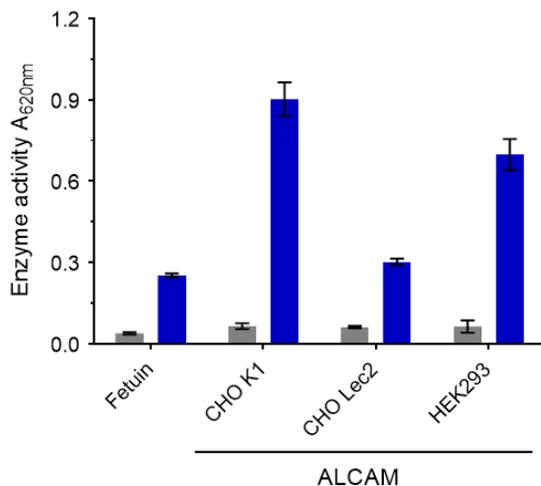


Figure 40: Use of MPSA to study *Cma* ST8Sia IV enzymatic activity. 400 ng of fetuin or various CD166/ALCAM versions were coated. Sialylations were achieved during 4 hours at 27°C with 100 μM of CMP-SiaNAI and 20 μL of *Cma* ST8Sia IV (blue) or Mock (grey) in cacodylate buffer at 27°C. Each step after sialylation is the same as described in Fig. 39. Absorbance was measured at 620 nm. Error bars in black represent SEM n=5.

As shown in Figure 40, significant and reproducible data were obtained on both acceptor substrates compared to the Mock control, showing that *C.ma* Δ28 ST8Sia IV is active and efficiently uses CMP-SiaNAI on both fetuin and CD166/ALCAM acceptor substrates. Interestingly, this enzyme exhibits a higher activity onto ALCAM produced in CHO-K1 than on the other version or on fetuin, suggesting a preferential activity onto α2,3-sialylated acceptor substrate (Fig. 40).

In contrast, *C.ma* ST8Sia II-r1 and ST8Sia II-r2 demonstrated no chemoenzymatic modification of these acceptor substrates under the same conditions in MPSA (Fig. 41), suggesting that these enzymes are either not active or do not use CMP-SiaNAI as donor substrate in MPSA.

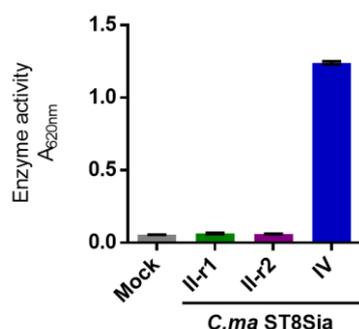


Figure 41: Use of characterization of *Cma* ST8Sia II-r1 and *Cma* ST8Sia II-r2 on CD166/ALCAM using the MPSA. Sialylation reactions were performed in MPSA as Fig. 40 using 55 μ L of *Cma* Δ 35ST8Sia II-r1 (green), *Cma* Δ 35ST8Sia II-r2 (purple) or 20 μ L of *Cma* Δ 28ST8Sia IV (blue). Other steps are the same as described in Fig. 40. Error bars represent SEM (n=3).

To conclude, we demonstrated here that MPSA could be extended to study polySTs activities. Notably, we showed that the human polySTs and *Cma* ST8Sia IV efficiently used CMP-SiaNAI as donor substrate and that *Cma* ST8Sia IV as a preferential activity onto α 2,3-sialylated ALCAM.

D. BIOCHEMICAL CHARACTERIZATION OF THE *CMA* ST8SIA IV ENZYMATIC ACTIVITY

At the same time, we optimized sialylation conditions to study enzymatic activity of polySTs. In MPSA, we used human Δ 39 ST8Sia IV to compare activities with *C.ma* Δ 28 ST8Sia IV and setting up several parameters: pH, temperature, quantities of enzymatic source or even sialylation time. In an effort to keep the sialylation reaction velocity in the same range for the two enzymes, we first evaluated the initial rate of the two enzymes. Initial velocity was proportional to the enzyme concentration up to 20 μ L for *Cma* ST8Sia IV and at least up to 400 ng for *Hsa* ST8Sia IV. A similar level of activity was detected for 150 ng of the human ST8Sia IV and 20 μ L of the fish enzyme (Fig. 42A) and product appearance was linear up to 4 hours of sialylation (Fig. 42B).

The influence of pH (4.4 to 6.8, measured at 27°C) and temperature (0 to 60°C) was assessed on each enzyme reaction velocity. An optimal temperature of 30°C and 37°C was found for the fish and human ST8Sia IV, respectively. To be noted, the fish enzyme remained highly active at high temperatures (Fig. 42C) suggesting a highly stable enzyme, whereas the human enzyme showed a peak activity between 30°C and 40°C. Our data also demonstrated a conserved optimum pH at 6.4 for both polysialyltransferase activities (Fig. 42D), comparable to those determined in another enzymatic assay for the human polyST ST8Sia II (Guo *et al.*, 2020).

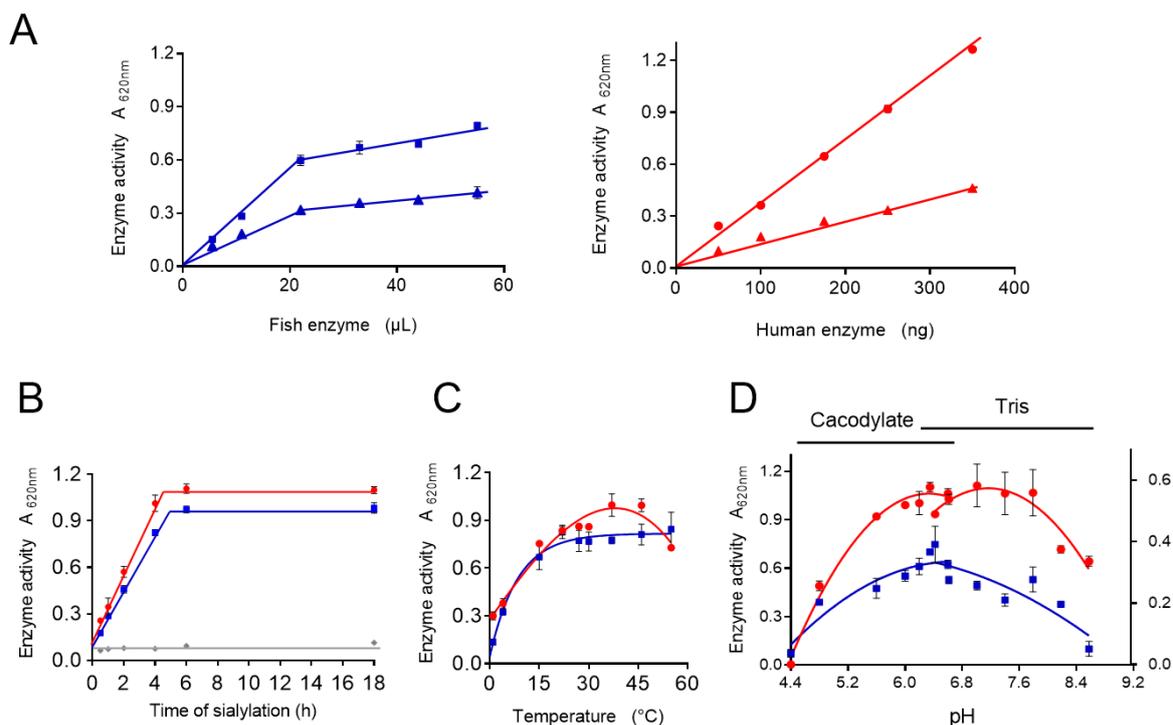


Figure 42: Optimization of conditions for ST8Sia IV activities in MPSA. A) Initial linear rate of sialylated product formation of both *Cma* and *Hsa* ST8Sia IV. The initial velocity of each ST8Sia IV was determined using either 400 ng of CD166/ALCAM (squares and circles) or fetuin (triangles) coated on microplates. Various quantities of the fish enzyme (blue, left side) ranging from 5.5 μL to 55 μL and of the human enzyme (red, right side) ranging from 50 ng to 350 ng were incubated with 100 μM of CMP-SiaNAI donor substrate in cacodylate buffer for 4 h at 27 $^{\circ}\text{C}$. After CuAAC cycloaddition of biotin and anti-biotin-HRP antibody detection, TMB substrate was used and absorbance was measured at 620 nm. The initial velocity remains linear up to 22 μL for the *Cma* ST8Sia IV and up to 350 ng for the human ST8Sia IV. Error bars in black represent SEM (n=3). **B) Time course behavior of sialylation reaction catalyzed by ST8Sia IV.** Sialylation reactions were conducted with the MPSA using 400 ng of CD166/ALCAM coated on microplates with 100 μM of CMP-SiaNAI using 150 ng of the human enzyme *Hsa* ST8Sia IV (red circle) or 20 μL of the fish enzyme *Cma* ST8Sia IV (blue squares) or of the Mock medium (grey diamonds) for 0.5, 1, 2, 4, 6 or 8 h at 27 $^{\circ}\text{C}$. After sialylation reaction, azido-PEG3-biotin was attached to the alkyne group of each transferred SiaNAI, biotin was detected by the anti-biotin antibody coupled with HRP and absorbance was measured at 620 nm. In these conditions, the product appearance was linear after 2 h of sialylation (n=2). **C) Effect of temperature on the sialylation reaction catalyzed by ST8Sia IV.** Sialylations were performed 4 h on 400 ng of CD166/ALCAM with 100 μM of CMP-SiaNAI and with 150 ng of *Hsa* ST8Sia IV (red circle) or 20 μL of *Cma* ST8Sia IV (blue squares) at different temperatures from 0 $^{\circ}\text{C}$ to 60 $^{\circ}\text{C}$ (n=2). **D) Effect of pH on the sialylation reaction catalyzed by ST8Sia IV.** Sialylation reactions were performed 4 h at 27 $^{\circ}\text{C}$ on 400 ng of CD166/ALCAM with 100 μM of CMP-SiaNAI and with 150 ng of *Hsa* ST8Sia IV (red circle) or 20 μL of *Cma* ST8Sia IV (blue squares) at different pH from 4.4 to 7.0 in cacodylate buffer (Cacodylate sodium 100 mM, MnCl_2 10 mM) and from 6.6 to 9.0 in Tris buffer (Tris-HCl 100 mM, Triton X-100 0.2%, MnCl_2 10 mM) (n=2).

Using assay conditions set at pH 6.4 and 27°C, we determined the kinetic parameters of each recombinant fish and human ST8Sia IV towards CMP-SiaNAI and CD166/ALCAM. We calculated Apparent Michaelis constant (K_m) values of $0.6 \pm 0.06 \mu\text{M}$ and $17.60 \pm 2.52 \mu\text{M}$ towards CMP-SiaNAI for the human and fish ST8Sia IV, respectively (Fig. 43A). These data pointed to 30 times lower affinity of *Cma* ST8Sia IV for the unnatural donor substrate CMP-SiaNAI compared to the human ST8Sia IV. In contrast, the two enzymes showed a relative similar affinity for ALCAM with the same order of magnitude (ratio of 3.6) and apparent K_m of $17.19 \pm 1.15 \text{ ng}$ and $61.09 \pm 4.85 \text{ ng}$ calculated for the human and fish ST8Sia IV, respectively, according to the Michaelis-Menten model (Fig. 43B).

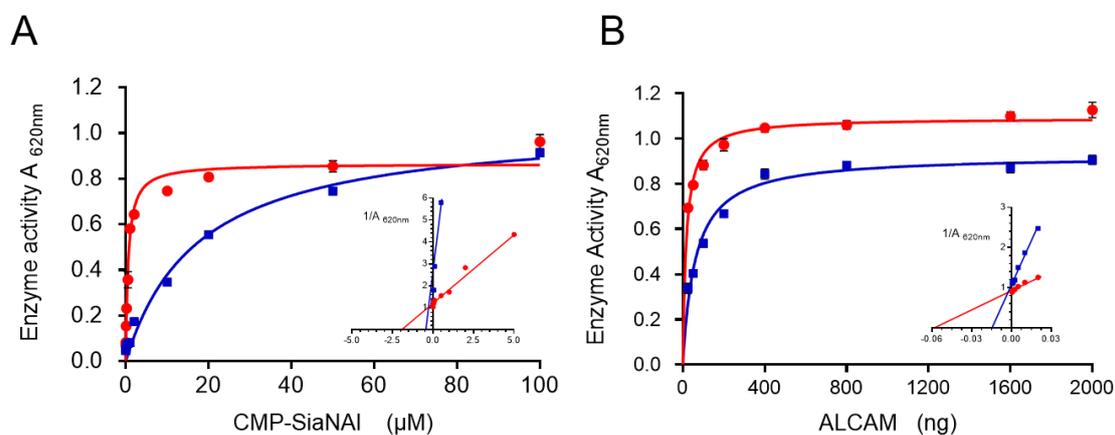


Figure 43: Determination of kinetic parameters of ST8Sia IV towards donor and acceptor substrates. Sialylation reactions were carried out in cacodylate buffer pH = 6.4, at 27°C for 4 h with 150 ng of human ST8Sia IV or 20 μL of fish ST8Sia IV. **A) With increasing concentrations of donor substrate.** CMP-SiaNAI was used at various concentrations ranging from 0 to 100 μM with a fixed amount of 400 ng of ALCAM (n=3). **B) With increasing quantities of acceptor substrates.** Various quantities of acceptor substrate ALCAM ranging from 0 to 2000 ng were fixed on microplates. Sialylation reactions were conducted with 100 μM CMP-SiaNAI (n=3). Michaelis-Menten constants were determined graphically on Lineweaver-Burk representation drawn using GraphPad Prism 6. *Hsa* ST8Sia IV (red circle) and *Cma* ST8Sia IV (blue squares) have an apparent K_m of $0.6 \pm 0.06 \mu\text{M}$ and $17.60 \pm 2.52 \mu\text{M}$ towards donor, respectively and of $17.19 \pm 1.15 \text{ ng}$ and $61.09 \pm 4.85 \text{ ng}$ towards acceptor substrate, respectively.

To assess enzyme specificity and acceptor substrate preference of each polyST, 400 ng of several glycoproteins with various *N*- and *O*-glycan structures (Table 13) were coated on microplates for sialylation assays. Both ST8Sia IV were equally active on fetuin, CD166/ALCAM, bovine submaxillary mucin (BSM), DNase I and NRP2, whereas the human ST8Sia IV exhibited

higher activity onto the fish polysialoglycoprotein PSGP-H, PSGP-L and orosomucoid (Fig. 39A). Likewise, enzymatic activities of *Hsa* ST8Sia II, *Cma* ST8Sia II-r1 and *Cma* ST8Sia II-r2 were also assessed with these various acceptors. *Hsa* ST8Sia II displayed an activity mainly on PSGP-H, PSGP-L, DNase I and CD166/ALCAM, whereas *Cma* ST8Sia II-related enzymes were not significantly active in MPSA. These observations indicate that these fish enzymes might not use the unnatural Sia donor CMP-SiaNAI although it must be noted that *Hsa* ST8Sia II has lower enzymatic activity (Fig. 44B) than *Hsa* ST8Sia IV (Fig. 44A).

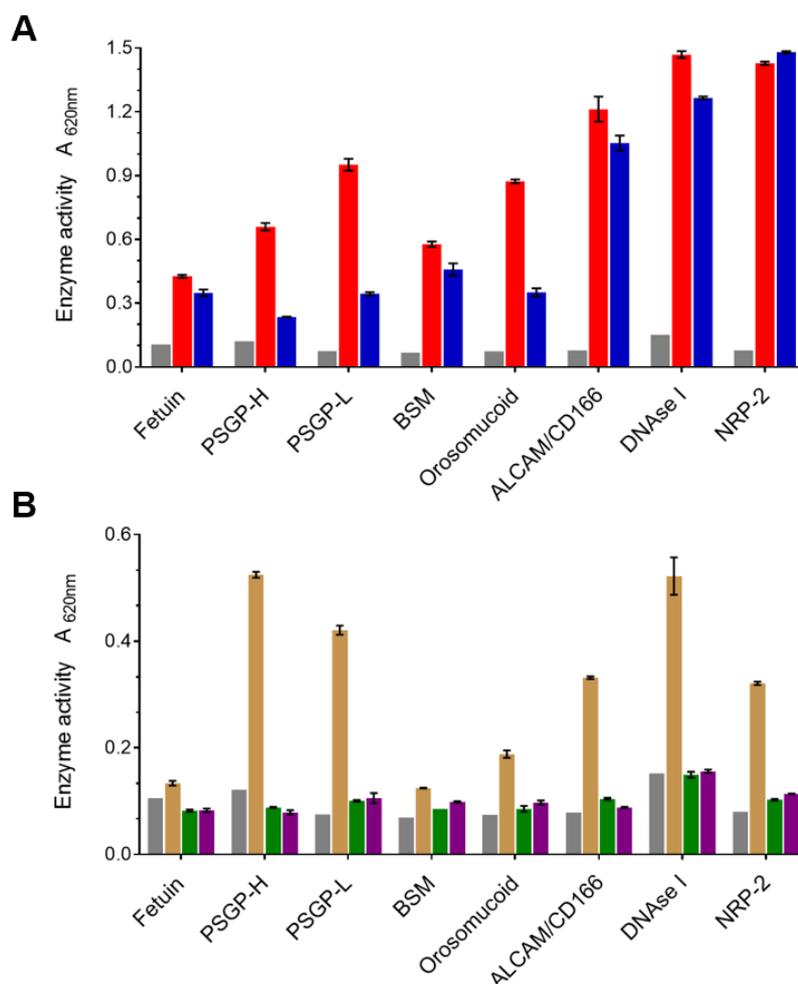


Figure 44: Functional characterization of *Hsa* and *Cma* polySTs Cma on various acceptor substrates using the MPSA. Four hundred ng of various sialylated *N*- and *O*-glycoproteins (*i.e.* fetuin, polysialylated glycoproteins PSGP-H and PSGP-L, bovine submaxillary mucin (BSM), orosomucoid, CD166/ALCAM, DNase I and neuropilin-2 (NRP2)) were coated on microplates. Sialylation reactions were performed 4 h at 27°C using 100 μ M of CMP-SiaNAI and (A) 150 ng of *Hsa* ST8Sia IV (red) or 20 μ L of *Cma* ST8Sia IV (blue) or Mock (grey) or (B) 150 ng of *Hsa* ST8Sia II (light brown) or 55 μ L of *Cma* ST8Sia II-r1 (green), 55 μ L of *Cma* ST8Sia II-r2 (purple) or Mock (grey). After each characteristic step for MPSA, absorbance data at 620 nm corresponding to the extent of chemoenzymatic modification are mean of two replicates \pm SEM (n=2).

To summarize, we extended the use of MPSA and unnatural donor substrate CMP-SiaNAI for the functional characterization of various human and fish polySTs. We optimized detection of their enzymatic activity varying conditions and choosing their best acceptor substrates. These precisions helped us to demonstrate that *Cma* ST8Sia II-r1 and *Cma* ST8Sia II-r2 have no significant activity with CMP-SiaNAI whereas human ST8Sia II could use this donor substrate, although in lower level compared to human ST8Sia IV. Furthermore, we highlighted significant differences of specificities for each ST8Sia IV towards CMP-SiaNAI, suggesting distinct recognition of the donor substrate during the mechanism of transfer catalysis at the molecular level. Finally, since only a limited set of natural polysialyated proteins like NCAM have been reported, it appears interesting to know that additional substrates could serve as suitable acceptor for polySTs.

However, MPSA has limitations since this enzymatic assay do not tell us if polySTs are processive and are able to catalyze the transfer of several Sia residues to form polySia chains. For this purpose, we performed *in vitro* sialylation reactions and achieved structural assessment of polySia using alternative approaches as detailed below.

E. POLYSIA ASSESSMENT GENERATED BY POLYSTS ON WB

As described in introduction, polySia chains are really sensitive to different parameters such as temperature and pH (Sato and Kitajima, 2004 ; Guo *et al.*, 2019). Firstly, we used 5 μ L of HEK293 *C.ma* ST8Sia IV-transfected cell lysates and boiled these samples in Laemmli buffer at different temperatures and times prior loading on WB (Fig. 45). In parallel, we also performed a control with endoN treatment of cell lysates to cleave polySia chains (Jakobsson *et al.*, 2012).

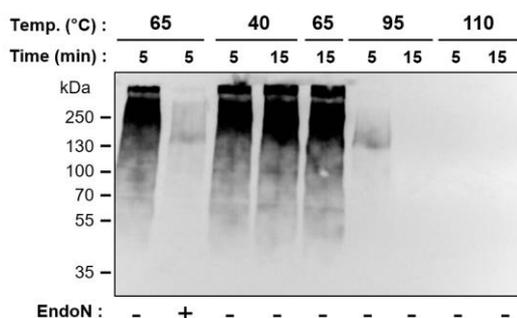


Figure 45: PolySia stability is temperature-dependent.

5 μ L of HEK293 *C.ma* ST8Sia IV-transfected cell lysates were boiled in Laemmli buffer at different temperatures (65 to 110°C) and times (5 to 10 min). EndoN was used 1h, 37°C (10 μ g/mL) as a control onto sample. After blocking step, membrane was incubated with mAb735 antibody (1:1000), secondary rabbit antibody (1:10000) and incubated 5 min with femtoECL before revelation.

To detect polySia chains on WB, we used the specific anti-polySia mAb735 antibody which recognizes polyNeu5Ac with a DP ≥ 8 residues (Mühlenhoff *et al.*, 2013). We clearly demonstrated that classical protocol for WB, boiling samples at 95°C in Laemmli buffer, is not suitable to prevent polySia from degradation and should be adapted to lower heating temperature at 65°C (Fig. 45).

Afterwards, we explored the polysialylation activity of polySTs using natural donor substrates. We firstly checked the polysialylation status of each reagent prior sialylation reaction on WB using mAb735 antibody (Fig. 46). Neither the enzymes nor the ALCAM collected in minimum medium nor the Mock control are initially polysialylated in contrast to efficient polySia detected after sialylation reaction (Fig. 46). Indeed, we performed *in vitro* sialylation starting with 100 μ M of chemo-enzymatically synthesized CMP-Neu5Ac and 20 μ L of *Cma* ST8Sia IV on 6 μ g of ALCAM.

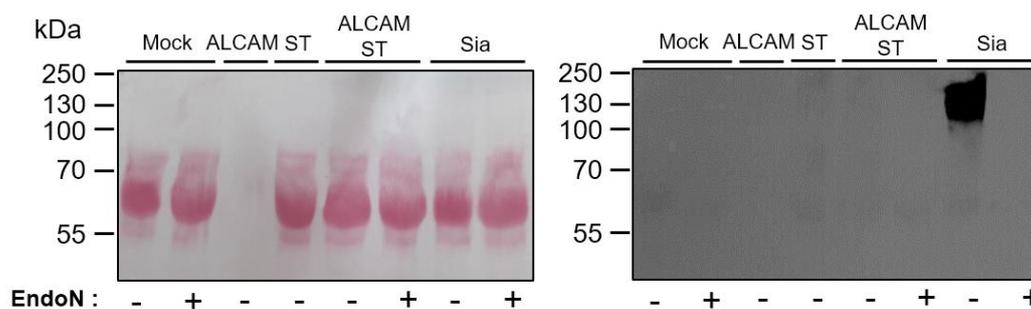


Figure 46: Initial polysialylation status of polySTs and ALCAM prior sialylation reaction. Each reagent for sialylation reaction was analyzed by SDS-PAGE and WB using mAb735 antibody prior sialylation. Ponceau staining is shown on the left side and molecular weight markers are indicated on the left top. On the WB (right side), no polySia can be detected by the mAb735 on the various actors of the sialylation reaction *i.e.* ALCAM, HEK293 cell culture medium (Mock) or *Cma* ST8Sia IV (ST). PolySia is detected only after 4 h of polysialylation on CD166/ALCAM with the *Cma* ST8Sia IV (Sia). Molecular weight markers are indicated on the right side of the WB.

Then, we were interested from how long sialylation reaction is required to detect polysialylation on WB ? Therefore, we carried out a time-course from 30 min to 8 h of sialylation using 100 μ M of CMP-Neu5Ac, 150 ng of *Hsa* ST8Sia IV or 20 μ L of *Cma* ST8Sia IV on 6 μ g of ALCAM. Within 30 min, an extensive smear ranging from 70 to more than 250 kDa could be detected with the mAb735, which confirms that the human and fish ST8Sia IV are both able to synthesize polyNeu5Ac chains (Fig. 47). To be noted, increase of ALCAM polysialylation status over time is more marked for the *Cma* ST8Sia IV (Fig. 47A) compared to *Hsa* ST8Sia IV (Fig. 47B).

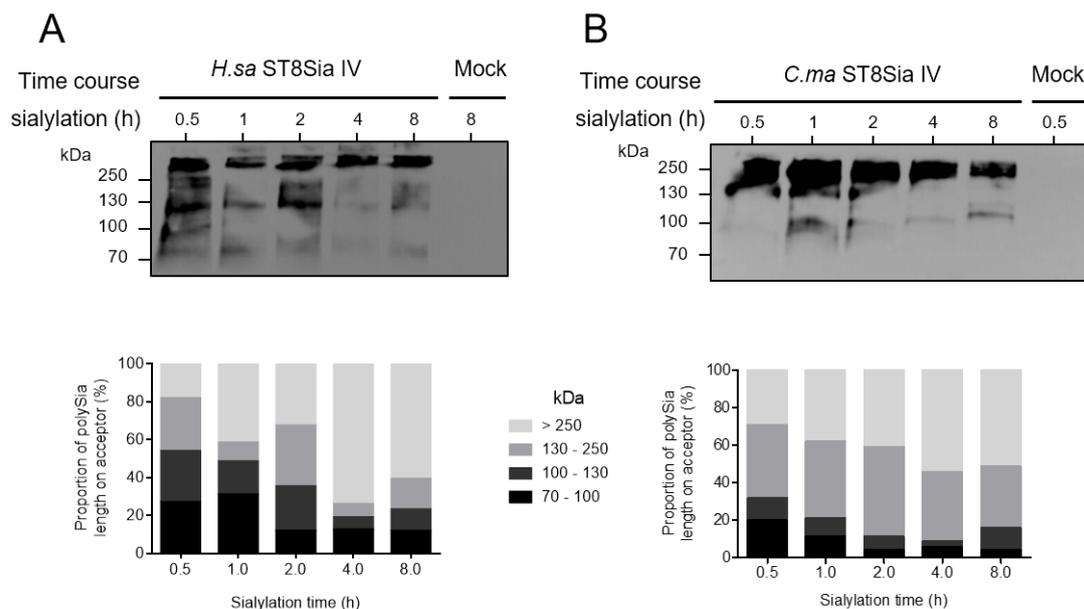


Figure 47: Time course polysialylation with human and fish ST8Sia IV showing an increasing polysialylation status of ALCAM. Sialylation reactions were performed from 0.5 to 8 h with 100 μ M of CMP-Neu5Ac on 6 μ g of ALCAM with 150 ng of *H.sa* ST8Sia IV or 20 μ L of *C.ma* ST8Sia IV or Mock. Sialylated products were heated for 5 min at 65°C in Laemmli, separated on SDS-PAGE (6%) and analyzed on WB using the anti-polySia mAb735. Molecular weight markers are indicated on the left side. Densitometry analyses were performed with ImageJ and the percentage of polySia chains length according to band sizes is reported below for A) the human and B) the coregone enzyme (n=2).

Since we used the three natural CMP-Sia chemoenzymatically synthesized, we investigated if mAb735 is adapted to detect polySia containing Neu5Gc or Kdn residues. Although this antibody has been crystallized (Fig. 27) and described by Nagae group in 2013, there were no direct evidence on polyNeu5Gc and polyKdn detection. Using UCSF Chimera with Dr Elin Teppa, we visualized the mAb735 structure (PDB: 3WBD) composed by 2 molecules, 6 complementary-determining regions (CDR) and the 11 water molecules involved in direct and indirect H-bonds interactions with polyNeu5Ac (Fig. 48A). To investigate to what extent analogous interactions could be found between the antibody and polySias made of eight Neu5Gc or Kdn, we modified the Neu5Ac to NeuGc/Kdn molecules in PDB. Then, complex was minimized using the AMBER force field and charges were computed using ANTECHAMBER. We showed that in addition to the six direct hydrogen bond interactions described for polyNeu5Ac (Fig. 48B), five others could be established between Tyr³⁹, Ala²³⁰, Gly²²⁷, Arg²²⁵ and Asn³³ for polyNeu5Gc (Fig. 48C), whereas five out of the six aa critical for polyNeu5Ac recognition were retrieved for polyKdn (Fig. 48D ; Fig. 49).

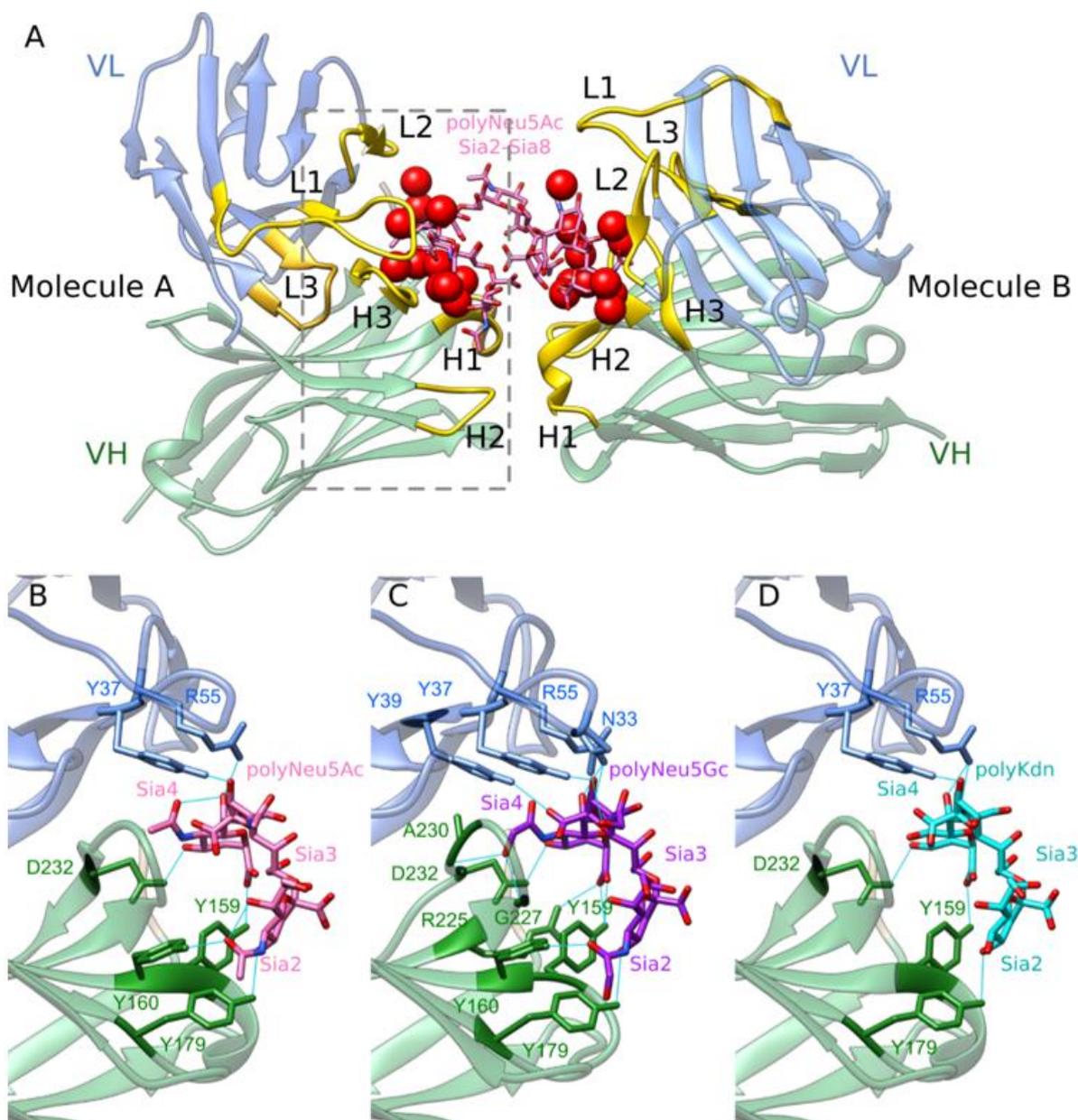


Figure 48: Crystal structure of scFv735 (PDB: 3WBD) in complex with α 2,8-linked octasialic acids. A) Overall crystal structure of scFv735 in complex with polyNeu5Ac, variable light domain (V_L) and variable heavy domain (V_H) and CDRs are coloured in blue, green and yellow respectively. The Sia residues interacting with two scFv735 molecules are highlighted in pink. Water molecules are shown in red spheres. The grey rectangle indicates the region of the close-up view. Close-up view of the antigen recognition site of scFv735 in complex with trisialic acid of B) Neu5Ac, C) Neu5Gc and D) Kdn. Hydrogen bonds between Sia and antibody are shown in cyan. Residues involved in direct H-bonds are shown in stick representation.

Moreover, we confirmed the reliability of each polymer interaction with mAb735 through the presence of the 11 water molecules conserved in the same CDR area of mAb735 and crucial for interaction with polyNeu5Ac (Fig. 49). Thus, we provided here for the first time, computational model for the recognition of various polySia chains using the mAb735 antibody.

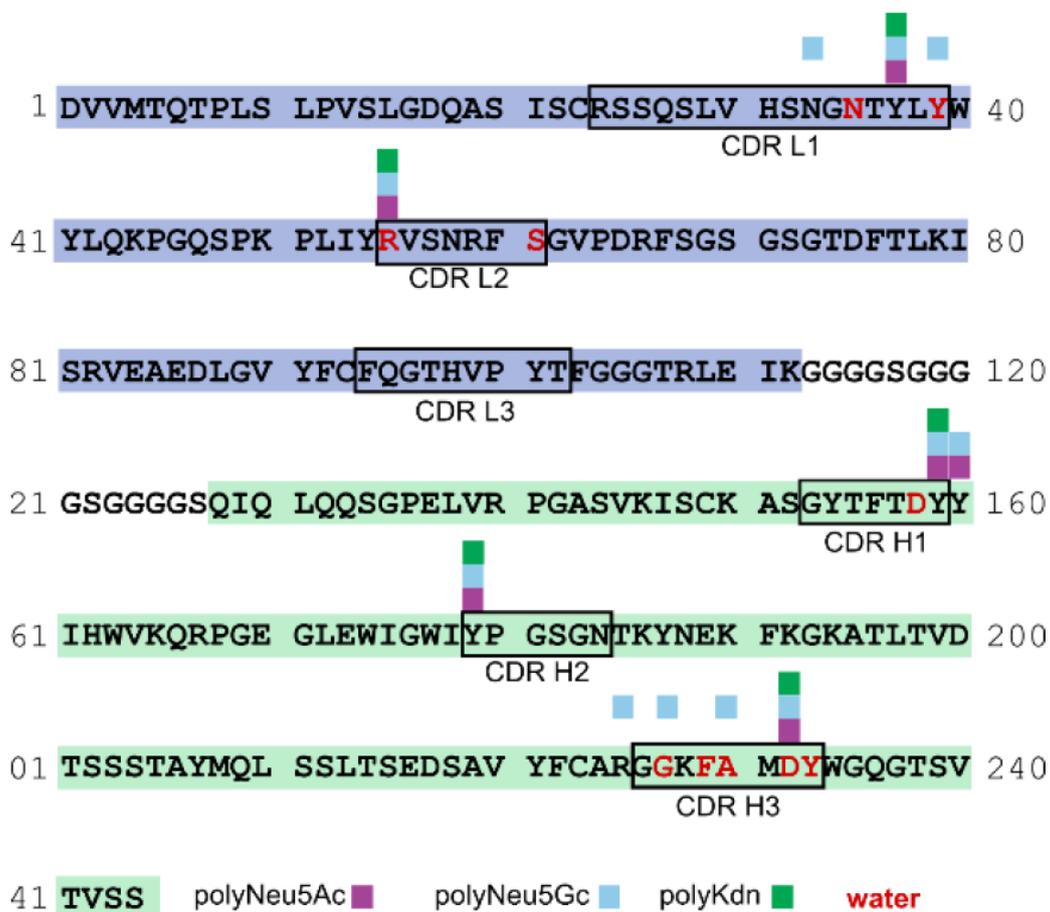


Figure 49: Protein sequence of mAb 735 and localization of hydrogen bound interactions. V_L and V_H domains are colored in blue and green background, respectively. The complementarity-determining regions (CDRs) L1, L2, L3, H1, H2, H3 are indicated with black boxes. Colored boxes above the sequence mark residues that make direct interaction with polyNeu5Ac (purple), polyNeu5Gc (light blue) and polyKdn (green), whereas residues involved in water-mediated indirect interactions with the antigen are shown in red.

Consequently, we evaluated the fish and human ST8Sia IV ability to generate polySia with various donor substrates and recognized by mAb735. We performed sialylation reactions with 100 μ M of the chemo-enzymatically synthesized CMP-Neu5Ac, CMP-Neu5Gc or CMP-Kdn (Fig. 36) on ALCAM. After protein separation on SDS-PAGE and WB with mAb735 and as checked using endoN treatment, broad polyNeu5Ac signals between 100 and 250 kDa were visualized for both ST8Sia IV with CMP-Neu5Ac whereas no signal could be detected in the Mock (Fig. 50Aa).

Remarkably, polysialylation with CMP-Neu5Gc (Figure 50Ab) and CMP-Kdn (Figure 50Ac) was detected after *in vitro* sialylation with the fish ST8Sia IV, but not with the human ST8Sia IV, highlighting distinct specificities depending on the nature of donor substrate for each ST8Sia IV.

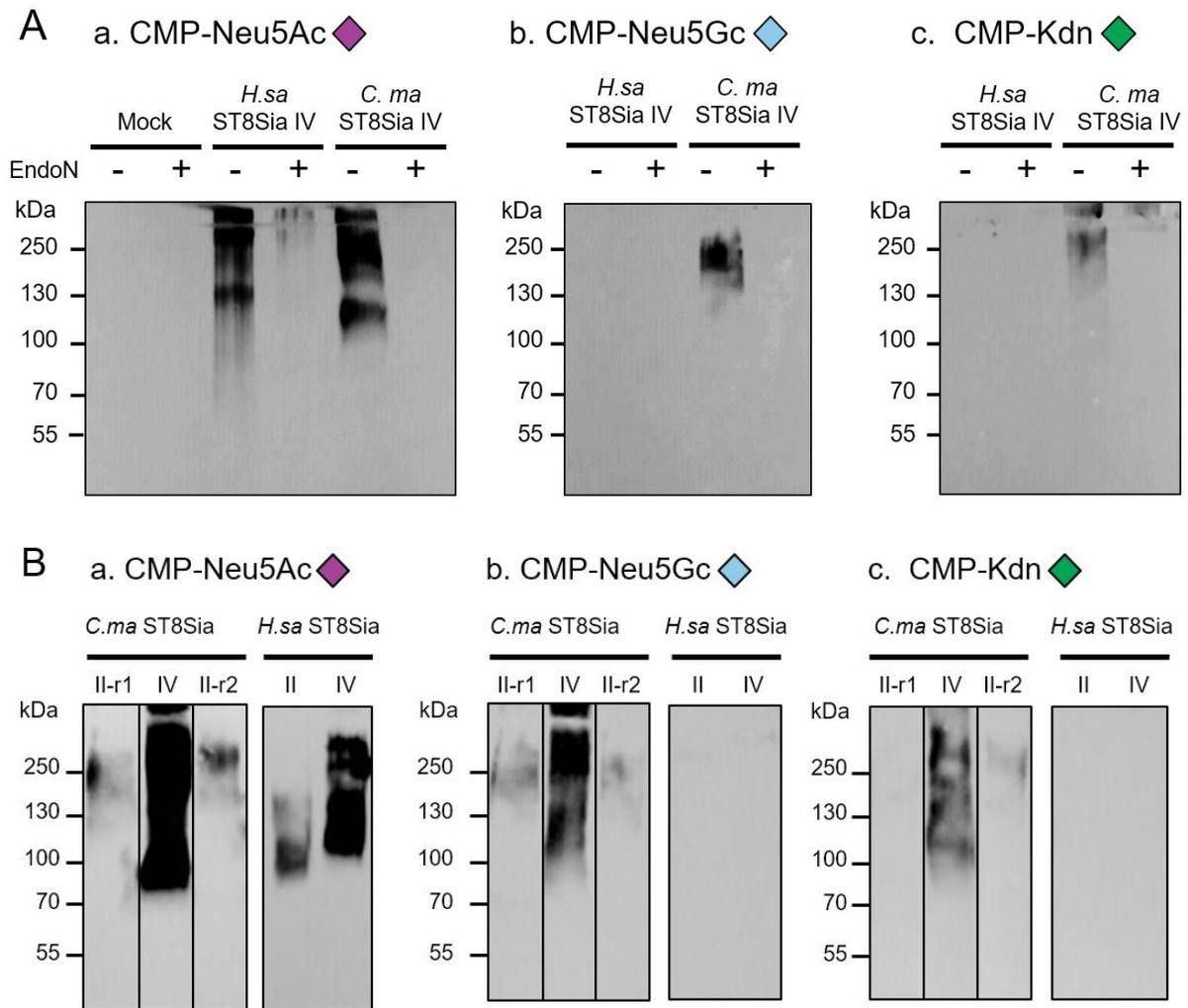


Figure 50: ALCAM polysialylation achieved with human and *Cma* polySTs and different CMP-Sia donors visualized on WB. A) Polysialylation with ST8Sia IV enzymes. Sialylation reactions were performed 4 h at 27°C with 100 μ M of the various natural Sia donors CMP-Neu5Ac, CMP-Neu5Gc or CMP-Kdn (purple, light blue and green diamonds, respectively) on 6 μ g of CD166/ALCAM with 150 ng of *Hsa* Δ 39ST8Sia IV or 20 μ L of either *Cma* Δ 28ST8Sia IV or Mock enzyme sources. Treatment with endoN for 1 h at RT was used as a control to remove polysialylation. Sialylation mixtures were heated 5 min at 65°C in Laemmli buffer, separated on 6% SDS-PAGE and proteins were transferred onto PVDF membrane. Blots were probed with the mAb735 antibody (0.5 μ g/mL). Immunoblots were incubated with femtoECL and detected after 5 sec for a) CMP-Neu5Ac, 30 sec for b) CMP-Neu5Gc and 180 sec for c) CMP-Kdn panels. **B) Polysialylation with ST8Sia II enzymes compared to ST8Sia IV.** Sialylation reactions were performed 4 h at 27°C with 100 μ M of the various natural sialic acid donors CMP-Neu5Ac (purple diamond), CMP-Neu5Gc (light blue diamond) or CMP-Kdn (green diamonds) on 6 μ g of CD166/ALCAM with 150 ng of *Hsa* Δ 23ST8Sia II or 55 μ L of either *Cma* Δ 35ST8Sia II-r1 or *Cma* Δ 35ST8Sia II-r2. Polysialylation with ST8Sia IV enzymes and detection on WB were conducted as mentioned in A). Immunoblots were incubated with picoECL and detected at 60 sec for a) CMP-Neu5Ac and with femtoECL detected at 60 sec for b) CMP-Neu5Gc and c) CMP-Kdn. Molecular weight markers are indicated on the left side.

Only low levels of polysialylation were detected with the *Cma* ST8Sia II-r1 and ST8Sia II-r2 compared to the *Cma* ST8Sia IV, whatever the Sia donor substrate used (Figure 50B). In contrast to fish polySTs, no polySia could be detected after polysialylation with CMP-Neu5Gc (Fig. 50Bb) or CMP-Kdn (Fig. 50Bc) by human polySTs denoting distinct specificities between each enzyme.

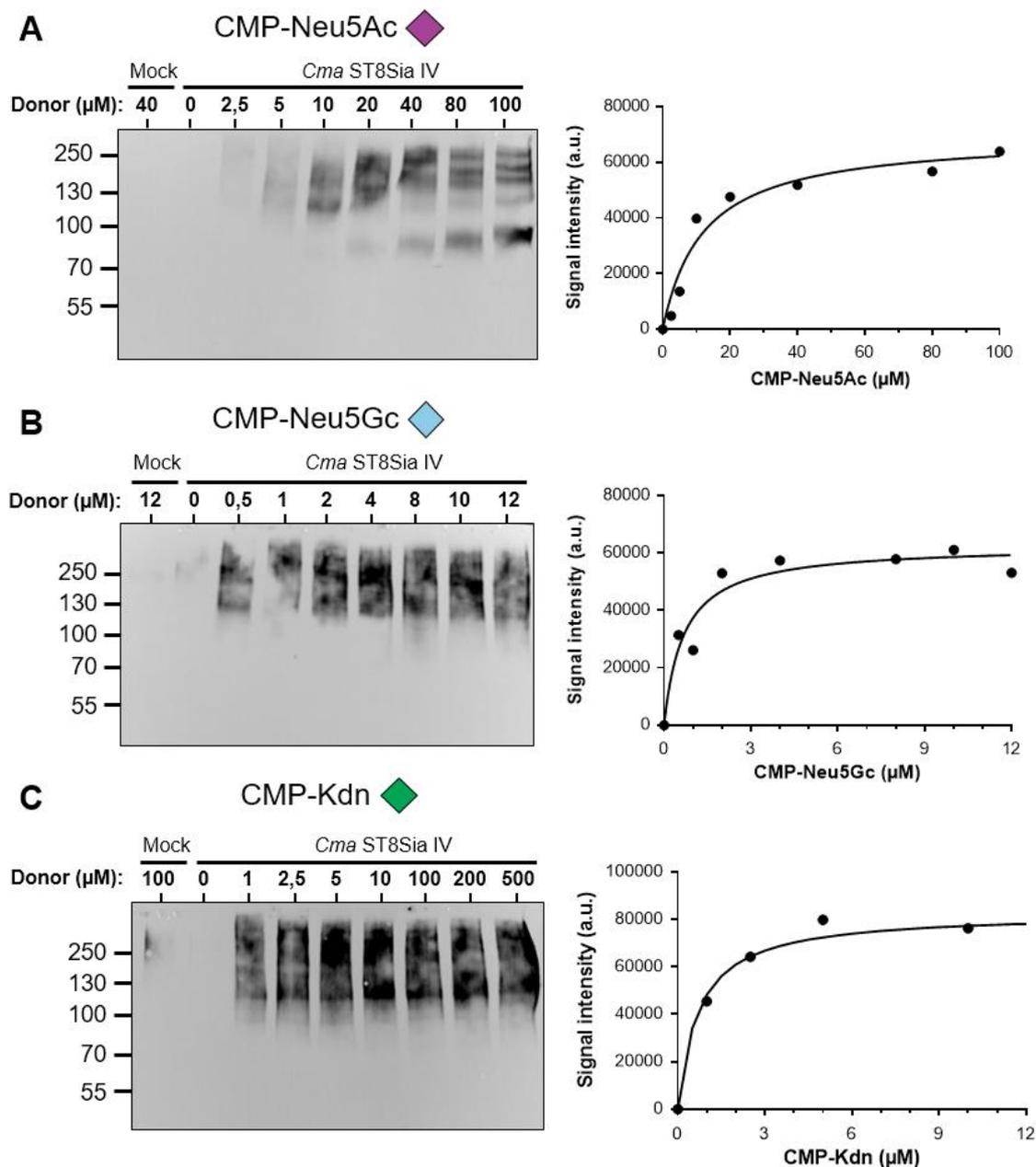


Figure 51: Relative kinetic parameters of *Cma* ST8Sia IV towards CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn determined from polysialylation on ALCAM. Sialylation reactions were performed 4 h at 27°C with various concentration of natural Sia donors CMP-Neu5Ac, CMP-Neu5Gc or CMP-Kdn on 6 μg of ALCAM with 20 μL of either *Cma* Δ 28ST8Sia IV or Mock enzyme sources. WB (left panels) were done as described in Fig. 50 and detected at 5 sec for A) CMP-Neu5Ac, 120 sec for B) CMP-Neu5Gc and 300 sec for C) CMP-Kdn with femtoECL. Signal intensity was quantified using ImageJ (arbitrary units) and reported on GraphPad Prism 6 (right panels).

In the light of these results, we focused on *Cma* ST8Sia IV which efficiently generate polySia using CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn. We studied its affinity towards these three donor substrates varying their concentrations during *in vitro* sialylation for further detection on WB. Although we should consider that WB allowed only relative quantification, we observed that the profile of polySia is different varying each CMP-Sia concentrations (Fig. 51). Indeed, polySia using CMP-Neu5Ac display an incremental effect (Fig. 51A) whereas signals appeared more diffuse and rapidly reaching a maximum with CMP-Neu5Gc (Fig. 51B) and CMP-Kdn (Fig. 51C). Nevertheless, we established Michaelis Menten profile for these donor substrates with apparent K_m of $12.04 \pm 3.59 \mu\text{M}$ for CMP-Neu5Ac and higher affinities with $0.68 \pm 0.24 \mu\text{M}$ for CMP-Neu5Gc and $0.72 \pm 0.21 \mu\text{M}$ for CMP-Kdn, highlighting that these donor substrates are not recognized and transferred in the same way by *Cma* ST8Sia IV.

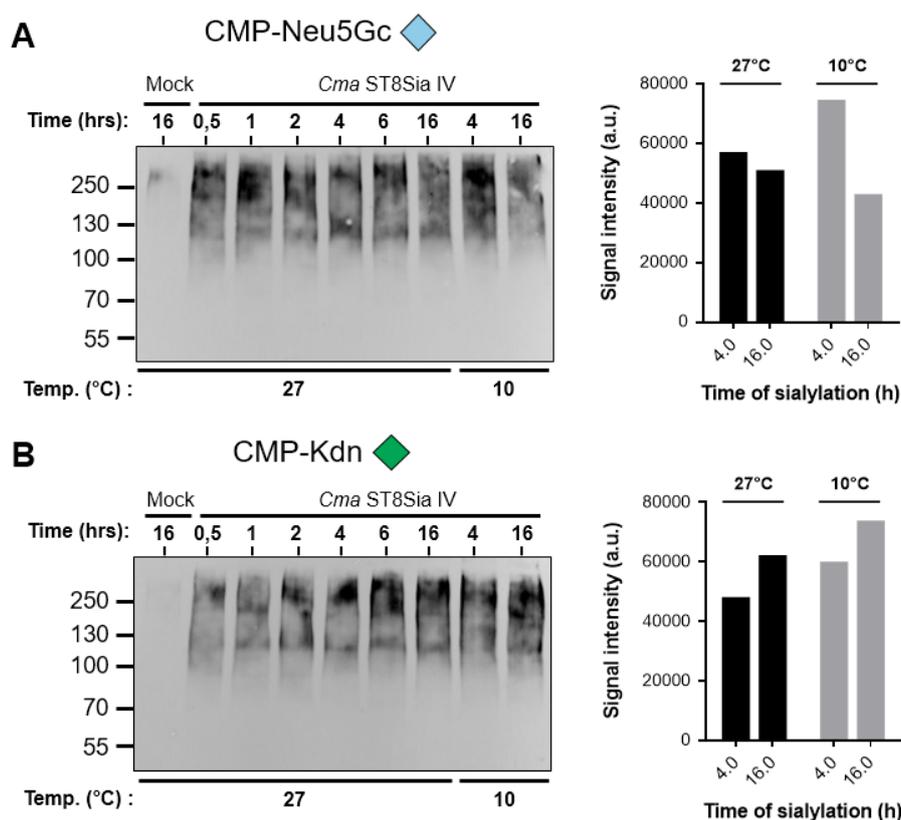


Figure 52: Influence of time and temperature on ALCAM polysialylation synthesized by *Cma* ST8Sia IV using CMP-Neu5Gc and CMP-Kdn. Sialylations were performed from 0.5 to 16 hours at 27°C and 10°C with 100 μM of CMP-Neu5Gc or CMP-Kdn on 6 μg of ALCAM with 20 μL of either *Cma* $\Delta 28$ ST8Sia IV or Mock enzyme sources. WB (left panels) were done as described in Fig. 50 and detected at 30 sec for A) CMP-Neu5Gc and 60 sec for B) CMP-Kdn with femtoECL. Signal intensity was quantified using ImageJ (arbitrary units) and reported on GraphPad Prism 6 (right panels).

As investigated for CMP-Neu5Ac in Fig. 47 and with CMP-SiaNAI on MPSA, we studied the potential influence of time and temperature on ALCAM polysialylation synthesized by *Cma* ST8Sia IV using CMP-Neu5Gc and CMP-Kdn. We observed that polySia generated at 27°C with CMP-Neu5Gc remain quite stable from 4 hours to 16 hours (Fig. 52A) whereas increased with CMP-Kdn in the same conditions (Fig. 52B).

Moreover, it is not uninteresting to think that temperature could have an impact on ST8Sia IV processivity as acute temperature rise induces metabolic changes detected in *Coregonus maraena* (Rebl *et al.*, 2018). Interestingly, we showed that *Cma* ST8Sia IV synthesized higher level of polySia with CMP-Neu5Gc and CMP-Kdn at 10°C, whereas differences between 4 and 16 hours are more pronounced than for 27°C (Fig. 52). To be noted, these experiments should be reproduced at lower concentrations of CMP-Neu5Gc and CMP-Kdn in accordance to K_m obtained in Fig. 51, to avoid potential polySia inhibition due to excess donor substrates.

To conclude on this part about polySia assessment generated by polySTs on WB, we have demonstrated an efficient polysialylation with CMP-Neu5Ac onto ALCAM for the two human polySTs and the three *Cma* polySTs. Combining computational tools and biochemical experiments, we showed that mAb735 could be used to detect polySia composed by Neu5Ac, Neu5Gc and Kdn residues. Our data using CMP-Neu5Gc and CMP-Kdn strikingly highlight that *Cma* polySTs, and particularly ST8Sia IV, have a broader enzymatic specificity towards donor substrates than human polySTs. To go further, structural assessment of these polySia have been studied and presented in the following paragraph.

F. STRUCTURAL ASSESSMENT OF POLYSIA GENERATED BY POLYSTS

As described in introduction, analytical approaches to study polySia require to consider the nature of the polymer as it interferes with protein mobility and antibody detection and to improve sensitivity, particularly when there is a low-abundance of polySia (Guo *et al.*, 2019). Here, we explored different strategies to concentrate polysialylated products and to optimize detection of polySia in terms of size (DP) and composition (nature of Sia residues).

Based on affinity for 3xFLAG tag, we immunoprecipitated *Cma* ST8Sia IV to concentrate and use it directly on anti-3xFLAG beads in enzymatic assays. As shown in Fig. 53, we succeeded to bind enzymes onto beads although we should consider that a part of ST8Sia IV remains present in the supernatant after first binding step as detected at 52 kDa, likewise beads detection on WB are composed by Fab and Fc fragment of Flag antibody at 50 and 28 kDa respectively (Fig. 53A). Only low levels of *Cma* ST8Sia IV are effectively eluted from the beads (~5%) both with 3xFLAG peptide (Fig. 53A) and glycine (data not shown) and are present after concentration on 30 kDa microcon.

Interestingly, the use of the enzyme directly fixed on beads showed significant activity with CMP-SiaNAI on MPSA although required more final concentrated quantities (Fig. 53B).

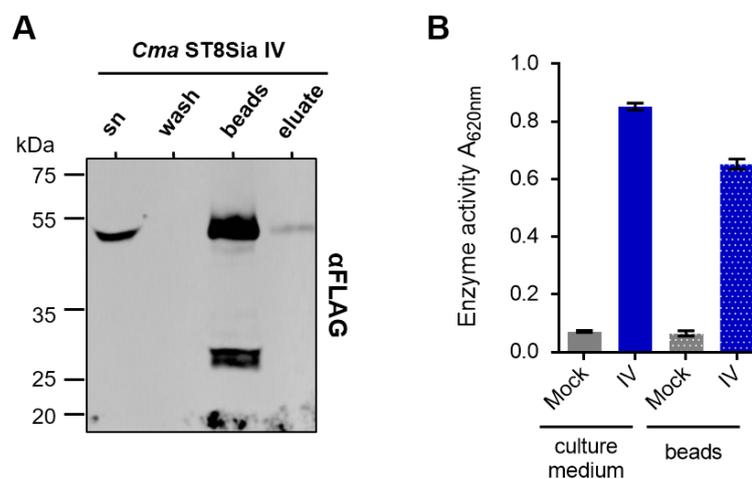


Figure 53: Immunoprecipitation of *Cma* ST8Sia IV on anti-3xFLAG beads and its use on MPSA. (A) Each reaction step was loaded on WB and detected using anti-Flag antibody. 200 μ L of enzymes were incubated on 50 μ L of slurry-beads overnight, 4°C on rotator disk on antiFLAG beads. Supernatant was collected after incubation (sn) then beads were washed 3 times (wash). Finally, elution with a 3xFLAG peptide was used overnight on rotator disk overnight at 4°C. Eluates were then concentrated on microcon 30 kDa (eluate). After elution, beads were washed 3 times and loaded (beads) on WB. Detection with picoECL, 30 sec (n=4). (B) MPSA on 400 ng of ALCAM using 20 μ L of ST8Sia IV culture medium and 10 μ L of 45X-concentrated ST8Sia IV-beads during 4 hours at 27°C and detected as previously described.

Then, we achieved sialylation reaction with CMP-Neu5Ac as previously described and attempted to fix mixture on 3xFLAG-beads after polysialylation. We detected efficient polySia with *Cma* ST8Sia IV before binding whereas only few amounts of polysialylated mixture comprising recombinant 3xFLAG-tagged ALCAM and ST8Sia IV were detected after binding on beads (Fig. 54). PNGase-F treatment of this mixture completely removes the small amount of

polySia detected on beads, suggesting that polysialylated mixture fixed on beads are composed by polysialylated N-glycans (Fig. 54). Failed attempts to fixed 3xFLAG ST8Sia IV and ALCAM on beads after polysialylation could be correlated with polySia physico-chemical properties which may hide and affect 3xFLAG affinity binding.

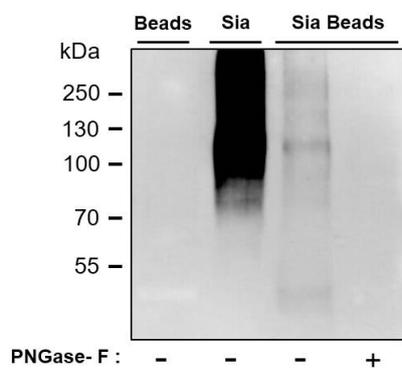


Figure 54: Preparation of polySia on beads and N-glycans release for analysis. Sialylation reactions were done 4 hours at 27°C onto 6 µg of ALCAM with 20 µL of *C.ma* ST8Sia IV and 100 µM of CMP-Neu5Ac. Samples were dialyzed in NH₄CO₃ buffer and fixed on 3xFLAG-antibody beads overnight at 4°C then treated 2 hours at 37°C with PNGase F. Control with only beads are loaded at the first lane. Detection with mAb735 and femtoECL, 10 sec.

Afterwards, we employed Fluorophore-Assisted Carbohydrate Electrophoresis (FACE) with Dr N. Szydlowski in Lille to study and separate polySia based on their electric charge differences under an electric field. Prior detection on FACE, we used aminopyrene trisulfonate (APTS) to derivatize and label monosaccharides. A first assay was done with standards of Sia (Neu5Ac, Neu5Gc and Kdn) and a mixture of them (SiaMix). After APTS derivatization and FACE injection, we noticed that each Sia displays many peaks detected at distinct time (Fig. 55), interesting to identify Sia residues in more complex samples containing polySia.

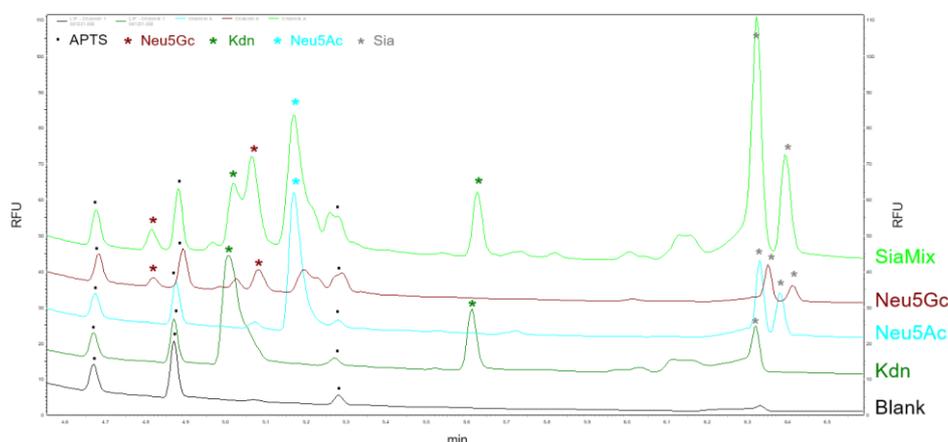


Figure 55: APTS derivatization of Sia monosaccharides and analyses in FACE. 100 µg of Neu5Ac, Neu5Gc, Kdn and 0,5 µmol of sialic acids mix (SiaMix) were derivatized 3 hours at 42°C with 2 µL of APTS 200 mM in acetic acid 15% and 2 µL of sodium cyanoborohydride 1M in tetrahydrofuran (THF). Samples were diluted 100 times in water before injections. 10 nL of samples were injected during 10 sec. Colored stars indicate significant peak detected for Sia residues.

Then, we derivatized and injected polyNeu5Ac from colominic acid (Barry and Goebel, 1957) for APTS-FACE analysis. We incubated this standard with APTS at 42°C overnight rather than 3 hours to optimize derivatization and treated samples with NaOH buffer at different pH for delactonization prior injection in FACE (Fig. 56). Indeed, lactone formation occurs during acidic treatment as used for derivatization, could impact signal and should be avoided for analysis (Kakehi *et al.*, 2001). We showed that 2 hours of colominic acid delactonization after APTS derivatization provides the best noise-to-signal ratio with sodium hydroxide buffer at pH = 10,8 (Fig. 56).

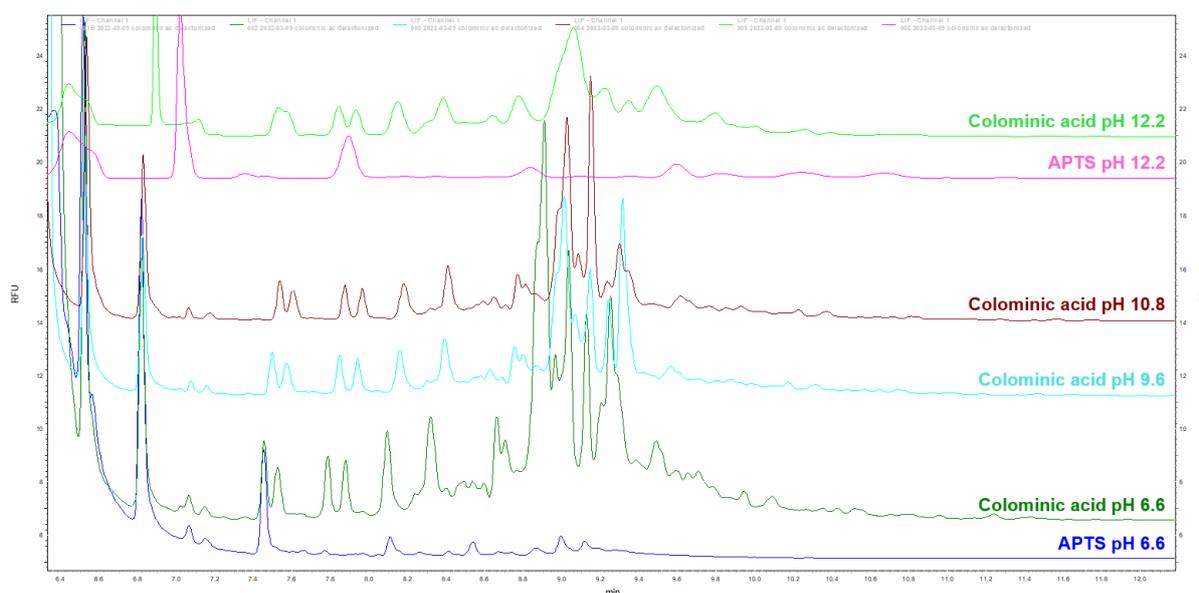


Figure 56: Optimization of colominic acid detection on FACE after APTS derivatization. 100 µg of colominic acid or water control were derivatized overnight at 42°C with 2 µL of APTS 200 mM in acetic acid 15% and 2 µL of sodium cyanoborohydride 1M in THF. Samples were delactonized 2 h at RT in buffers at different pH (6.6, 9.6, 10.8 and 12.2). 10 nL of samples 50 times diluted were injected during 10 sec.

Finally, we combined strategies used in Figures 54 and 56 to detect polySia on N-glycans after enzymatic activity of human and *Cma* ST8Sia IV with each natural CMP-Sia. After release of N-glycans, derivatization overnight with APTS and injection in FACE didn't allow us to detect significant differences between each condition in terms of DP (Fig. 57).

To go further with analytical approaches to study polySia, we have chosen other strategies to concentrate polySia samples using inactive endoN affinity to bind polySia (Jakobsson *et al.*, 2012) rather than 3xFLAG tag affinity and targeting more specifically Sia monosaccharides using DMB reagent for HPLC analysis (Fig. 29) instead of using APTS .

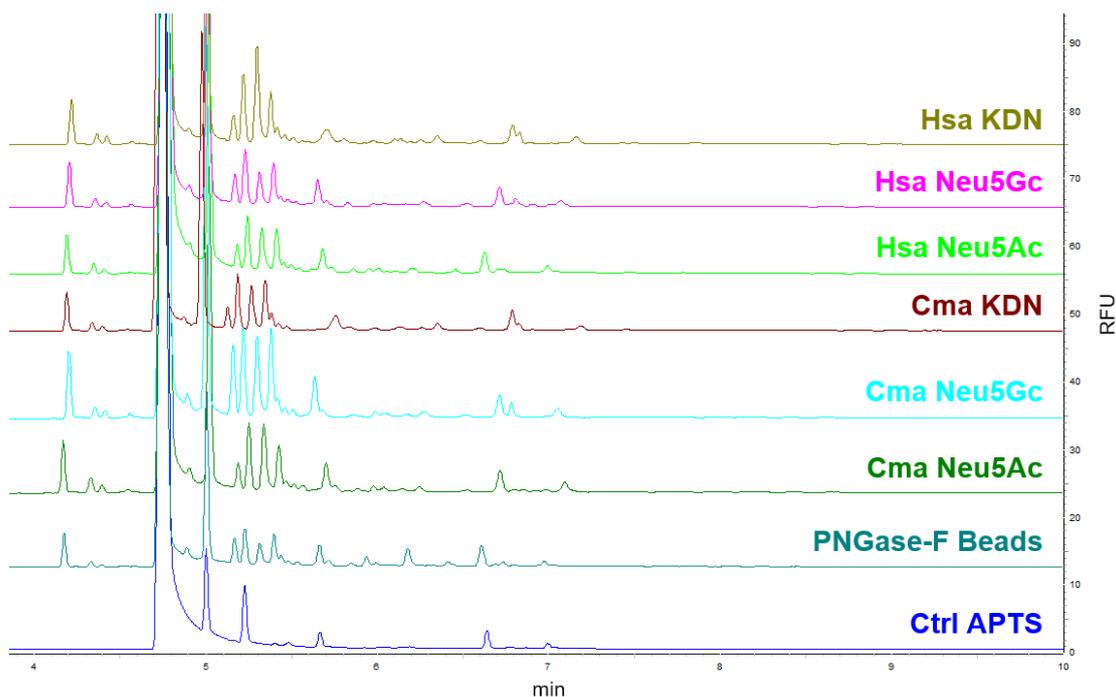


Figure 57: APTS-FACE analysis of ALCAM N-glycans after polysialylation with each ST8Sia IV. Firstly, 6 μg of ALCAM were sialylated with either CMP-Neu5Ac, CMP-Neu5Gc or CMP-Kdn (100 μM) and 20 μL of *Cma* ST8Sia IV or 150 ng of *Hsa* ST8Sia IV. After 4 hours of sialylation at 27°C, samples were dialyzed in NH_4CO_3 , lyophilized, fixed on 3xFLAG beads overnight at 4°C then treated with PNGase F for 2 hours at 37°C. Samples were derivatized overnight at 42°C with 2 μL of APTS 200 mM in acetic acid 15% and 2 μL of sodium cyanoborohydride 1M in THF. Samples were delactonized in buffers at pH = 10.8 during 2 hours at RT. 10 nL of samples 50 times diluted were injected during 10 sec.

During a stay in Dr S. Galuska laboratory (FBN, DEU), I have benefited from expertise in polySia analysis using DMB-HPLC (Galuska *et al.*, 2011). We started using each ST8Sia IV and the 3 natural donor substrates CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn in sialylation reaction onto 6 μg of DNase I, which is also a good acceptor substrate for polySTs (Fig. 44). Results on WB showed the same profile as for ALCAM (Fig. 50): *Cma* ST8Sia IV could use the three donor substrates to generate polySia whereas *Hsa* ST8Sia IV use only CMP-Neu5Ac for polySia.

Then, we analyzed the nature and composition of polySia detected in WB. We cut out the bands of interest on PVDF membranes from WB, oxidized them and released Sia after full hydrolysis with trifluoroacetic acid (TFA). Then we derivatized released Sia with DMB reagent prior to inject them into HPLC. In parallel, we prepared and injected standards containing SiaMix. Chromatograms obtained from DMB-HPLC analysis corroborated WB results: *Cma* ST8Sia IV could transfer CMP-Neu5Gc and CMP-Kdn onto polySia in contrast to *Hsa* ST8Sia IV (Table 18).

Table 18: DMB-HPLC analysis and Sia quantification from polysialylation on DNase I with various Sia and both ST8Sia IV enzymes detected on WB. Sialylation reactions were performed as described in Fig. 50. Samples were treated by TFA and derivatized by DMB to release and labeled Sia prior to DMB-HPLC analyses performed on C18 column. SiaMix containing Neu5Ac, Neu5Gc, Kdn at different concentrations were injected as standards. Sia quantities were calculated from standard curve established from SiaMix.

Donor	Enzyme	Sialic acids quantity (pmol)	%
CMP- 	Fish	3,3	100
	Human	2,7	81,8
CMP- 	Fish	1,8	54,5
	Human	0	0
CMP- 	Fish	< 0,01	< 1
	Human	0	0

Although requiring to be scaled up to detect easier each Sia residue from polySia, these results are consistent with a sialomic study in pig tissues where about 10 μ moles of total Sia (Neu5Ac/Neu5Gc/Kdn) are reported per g of proteins in pig (Ji *et al.*, 2017) which is quite comparable to 5-10 pmoles Sia detected per 0,4-1 μ g of acceptor substrates.

In parallel, we investigated polySia DP detection using mild hydrolysis and DMB derivatization on colominic acid standard. The resulting fluorescently DMB-labelled polymers were separated according to their size by anion-exchange chromatography and allowed detection of various peaks corresponding to each DP of Sia residues associated on chains (Fig. 58).

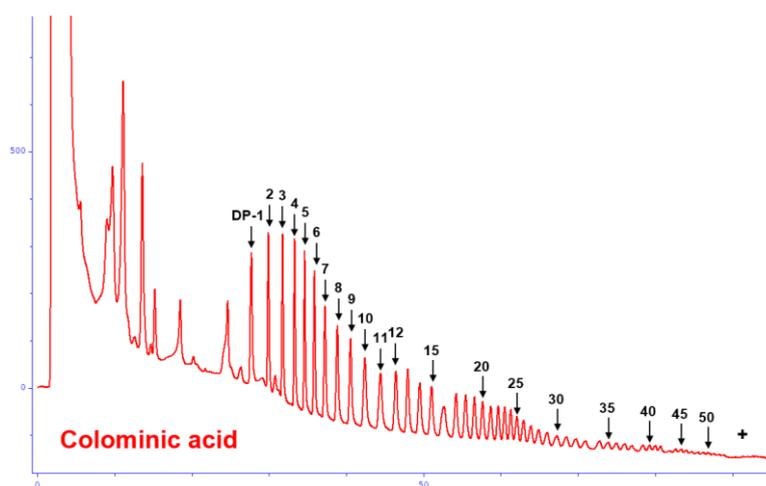


Figure 58: DMB-polymers of colominic acid analyzed by anion-exchange chromatography. PolySia degrees of polymerization (DP) were determined after DMB derivatization of 5 μ g of colominic acid injected on anion-exchange chromatography.

To extend this approach to *in vitro* polysialylation reactions, we used inactive endoN (iEndoN) on column which is able to concentrate polySia through its lectin domain (Schwarzer *et al.*, 2009). We prepared 10 times more polysialylation reactions onto ALCAM with each ST8Sia IV, each natural donor substrate, pooled and dialyzed them (10X). Then, we have enriched polysialylated fraction using affinity precipitation onto iEndoN-column. After washing steps, polySia were eluted with triethylamine and dried in a vacuum concentrator. Using 10% of eluate, we controlled affinity precipitation with mAb735 on WB. We noticed that purification on iEndoN column improves detection of 10X samples, particularly for polySia synthesized with CMP-Neu5Gc and CMP-Kdn with *Cma* ST8Sia IV (Fig. 59). Interestingly, these data indicated that endoN and iEndoN are respectively able to bound and cleave polySia generated with CMP-Neu5Gc and CMP-Kdn in addition to polyNeu5Ac (Fig. 59).

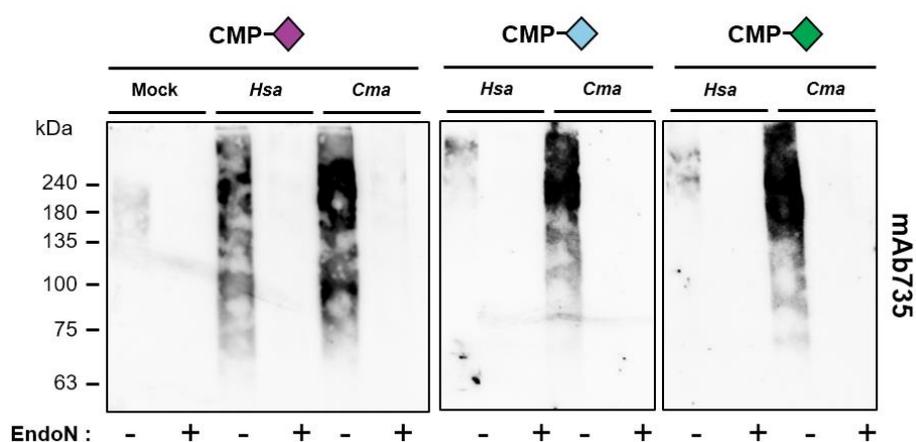


Figure 59: PolySia resulting from 10X sialylation reaction on ALCAM purified on iEndoN-column. Sialylations were done during 4 hours at 27°C onto 20 µg of ALCAM. Donor substrates were CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn (100 µM) with 1000 ng of *Hsa* ST8Sia IV and 200 µL *Cma* ST8Sia IV. Samples were dialyzed in NH₄CO₃, lyophilized and ± EndoN-treated (1 hour, 37°C) then purified on EndoN-inactive column. WB were done with 10% of samples load on SDS-PAGE 7% with mAb735 detection and using homemade ECL at 5 sec for CMP-Neu5Ac, 60 sec for CMP-Neu5Gc and 300 sec for CMP-Kdn.

Using the other 90% of eluate, polySia samples were treated with mild hydrolysis to release free polymers and labeled with DMB. The resulting fluorescent labeled polymers were separated according to their DP by anion exchange chromatography. We detected incremental peaks corresponding to diverse size until DP = 15-20 Sia residues with CMP-Neu5Ac (Fig. 60B)

consistent to polySia signals detected with mAb735 antibody on WB (Fig. 60A). Although it appears difficult to precisely quantify DP polySia due to heterogeneity, this proof of concept provides precious information about size range of polySia generated after *in vitro* polysialylation with *Cma* ST8Sia IV (Fig. 60). Otherwise, we were not able to detect significant peaks from polysialylated samples with CMP-Neu5Gc and CMP-Kdn, consistent with results obtained on WB where the detection time with these donor substrates had to be increased to visualize enough signal of polySia (Fig. 50, 51, 52, 59). In the future, the challenge to detect polySia synthesized from *in vitro* polysialylation with CMP-Neu5Gc and CMP-Kdn will be to scaled up preparations and starting with at least 100 μ g of polysialylated acceptor as recommended by Ludger Manufacturer.

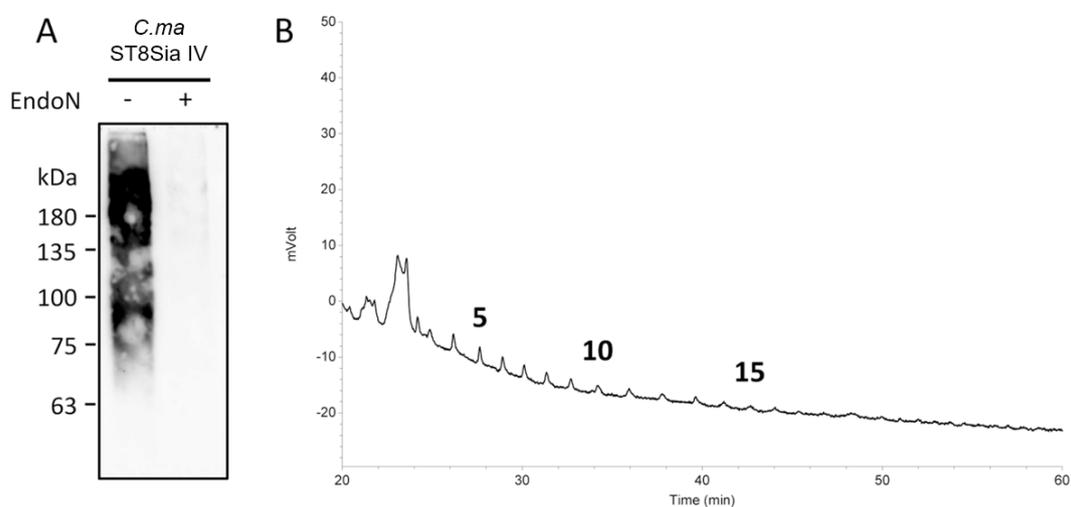


Figure 60: Chain length analysis of CD166/ALCAM after polysialylation by *Cma* ST8Sia IV. Polysialylation reactions were performed as described previously. Polysialylated ALCAM was isolated by affinity precipitation using iEndoN coupled to magnetic beads. **A)** A 10% fraction was used for WB to control the precipitation. **B)** The other 90% of the fraction was used for mild DMB labeling. Resulting fluorescent labelled Sia polymers were separated via anion-exchange chromatography. The chain length is given for selected peaks.

To proof the transfer of Neu5Ac, Neu5Gc and Kdn by *Cma* ST8Sia IV with an antibody independent and sensitive approach, we applied a HPLC-based strategy (Fig. 61A) in the laboratory of our german collaborator, Dr S. Galuska and thanks to his PhD student Anna. For this purpose, polysialylation of ALCAM was performed in centrifugal filter units. After polysialylation, remaining CMP-Sia was removed before polySia degradation by endoN.

Untreated samples (without endoN treatment) were used as negative control. Polysialylation and degradation by endoN were performed in the centrifugal filter units as checked by WB (Fig. 61B).

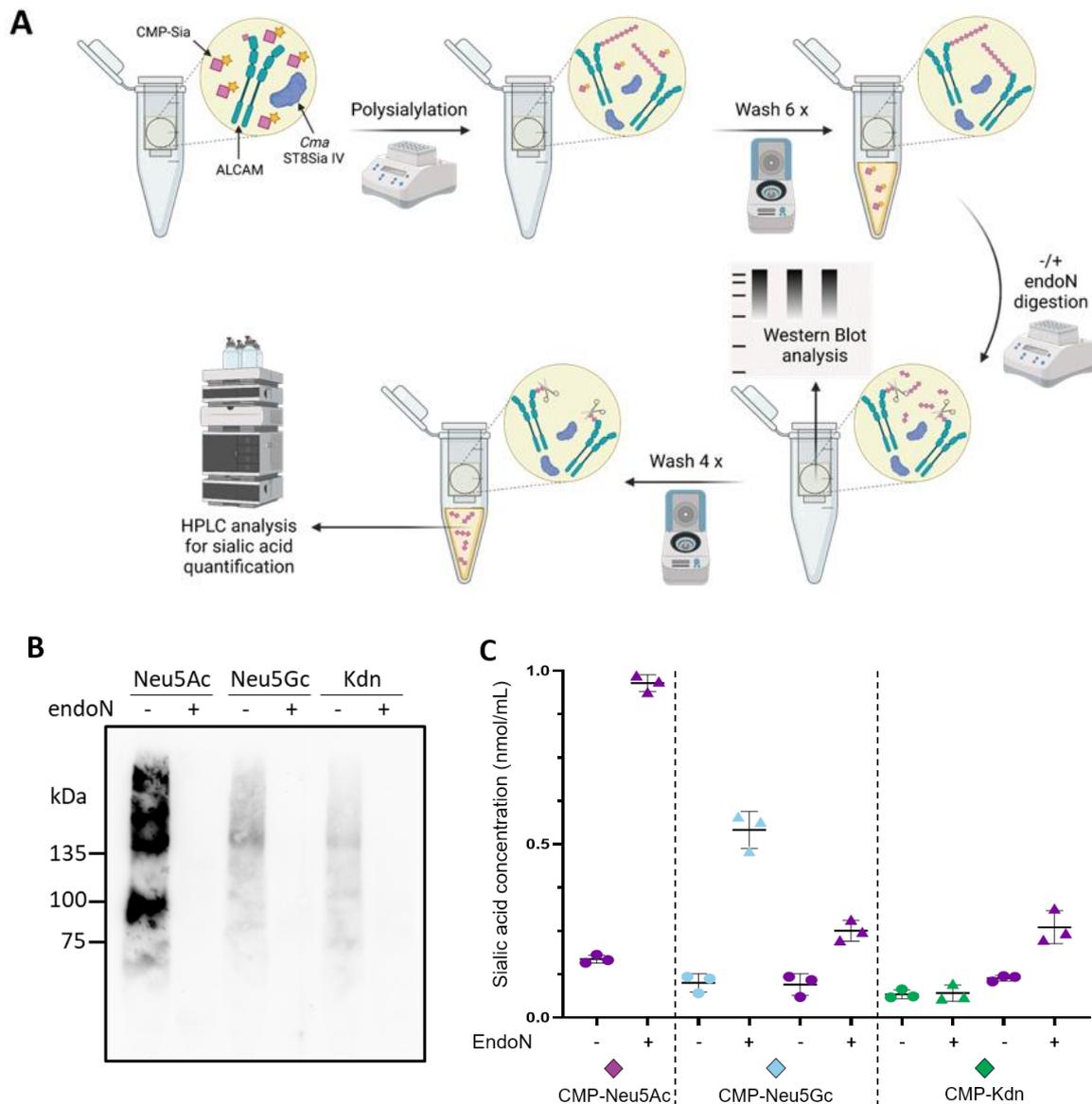


Figure 61: Composition analysis of polySia on CD166/ALCAM. **A)** Workflow of the centrifugal filter polysialylation assay to quantitatively analyse the transfer of different Sia to ALCAM with *Cma* ST8Sia IV. **B)** The polysialylation reaction using different CMP-Sia was monitored by WB. **C)** Resulting polySia fragments found in the flow-through after four washes were analyzed by HPLC. All Sia residues detected were represented according the SNFG nomenclature.

After a final filtration step, the flow-through were analyzed by HPLC. Newly added Sia residues could be determined comparing the Sia content in the flowthrough of untreated and endoN treated samples. When CMP-Neu5Ac was used as donor substrate, a strong increase of Neu5Ac was detected after endoN digestion indicating that *Cma* ST8Sia IV efficiently transferred Neu5Ac

to ALCAM (Fig. 61C). With CMP-Neu5Gc, an increase of Neu5Gc was also observed although lower than that obtained with CMP-Neu5Ac, in line with the WB data obtained (Fig. 61C). However, the transfer of Kdn could not be proven since the Kdn values of the endoN and the untreated sample were comparable. Only slightly higher amounts of Neu5Ac were detectable in endoN treated samples of the CMP-Kdn setup (Fig. 61C). These might be a result of an elongation of pre-existing oligo-Neu5Ac since a comparable increase of Neu5Ac in the endoN sample was also observed in the case of CMP-Neu5Gc (Fig. 61C). The transfer of few Kdn residues was likely not sufficient to be detected by HPLC, but enabled binding of mAb735 and degradation of polySia by endoN, a hypothesis that is further supported by the WB data.

To summarize, the structural assessment of polySia forced us to develop various strategies notably increasing quantities of polysialylated products. We concentrated tagged-acceptor substrates and tagged-enzymes by immunoaffinity and using the binding properties of iEndoN to enriched polysialylated fraction. We investigated various analytical methods from APTS-FACE, DMB-HPLC and anion-exchange chromatography in order to precise DP and composition of polySia chains generated. Although still a challenge to be optimized, these analytical approaches allowed us to confirm the polysialylation activity of *Cma* ST8Sia IV.

G. EXOSIALYLATION OF CULTURED CELLS USING ST8SIA IV

Since these last ten years have seen the emergence of glycans remodeling on cell surface using GTs like ST6Gal I and SEEL approach (Mbua *et al.*, 2013), we also investigated exogenous sialylation on cell surface using polySTs. This cell-based approach has the advantage to broaden the panel of acceptor substrates available for an efficient polysialylation in more “physiological-like conditions” than *in vitro* sialylation. Prior exogenous polysialylation, we turned to HEK293 cells and ensured they display NCAM on their cell surface and we checked their polysialylation status (Fig. 62). The immunostaining exhibited that NCAM on the cell surface was already polysialylated by cellular polySTs (Fig. 62A). Since it would be impossible to distinguish between the cellular polysialylation and the subsequently added polysialylation products of the recombinant enzymes, we fixed HEK293 cells and we used endoN to remove initial polySia (Fig. 62B).

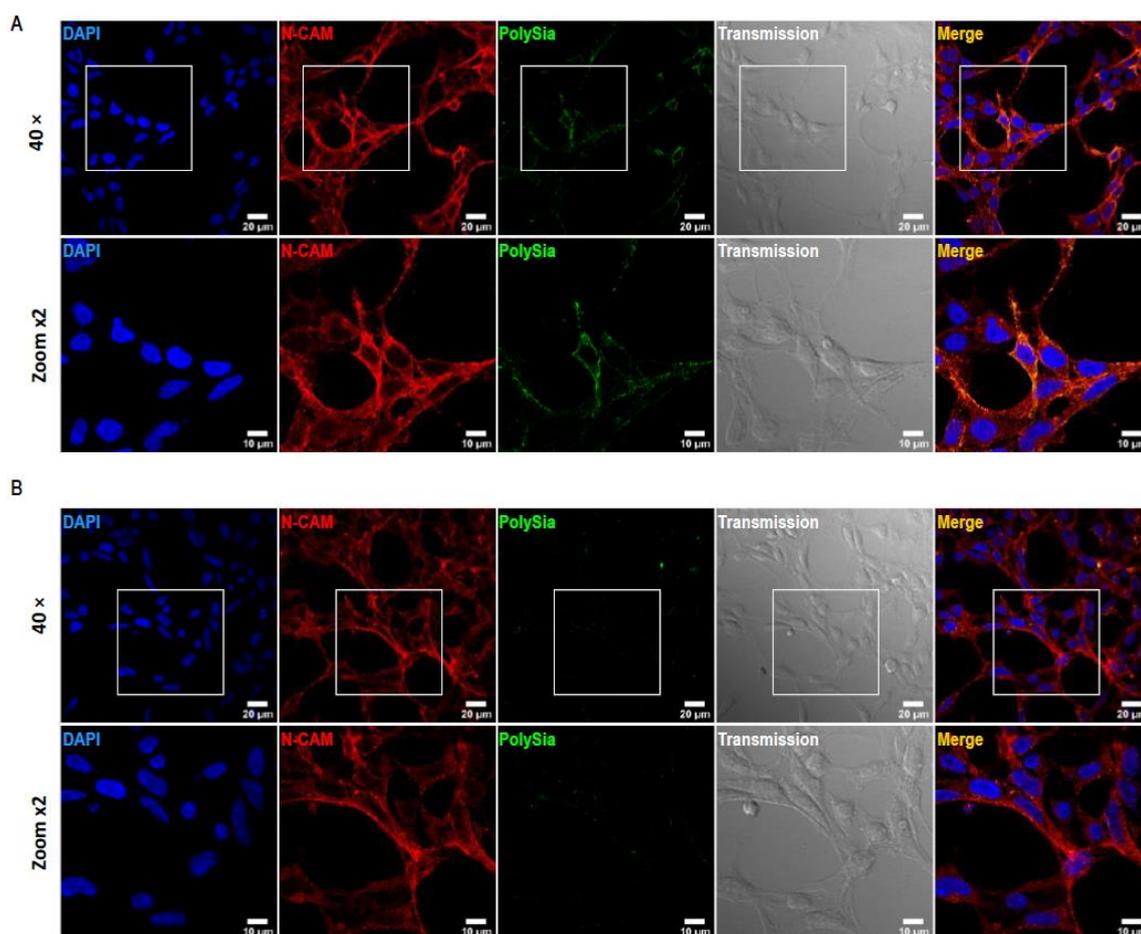


Figure 62: Polysialylation of HEK293 cells. Two days after seeding, HEK293 were fixed with 4% PFA and IF experiments were performed. Confocal microscopy images on A) WT HEK293 cells show polysialylation labelling on cell surface using mAb735 and anti-NCAM antibodies. B) After EndoN treatment, mAb735 detection signal is abolished. Nuclei are stained with DAPI. Outline of the cells is visualized by transmission in grey. DAPI (blue), NCAM (red) and polySia (green) signals are superposed in Merge. Fluorescence was detected through a Carl Zeiss confocal microscope and Transmission correspond to transmitted light microscopy images. Scale bars: 20 μm for acquisitions at 40 × and 10 μm for the Zoom x2 pictures.

To be noted, optimization of the endoN quantity used on cells was a crucial point for further efficient exogenous repolysialylation. Indeed, too much quantities on endoN (500 ng/mL) effectively remove initial polySia but high amount of endoN remain fixed on oligoSia chains and prevent their use as acceptor substrates from significant exogenous polysialylation. Thanks to our laboratory engineer Céline and to Dr M. Mühlhoff (MHH, DEU) advices and fresh-prepared endoN, we determined that 100 ng/mL of endoN is the best balance between removal of initial polySia and accessibility for further exogenous polysialylation.

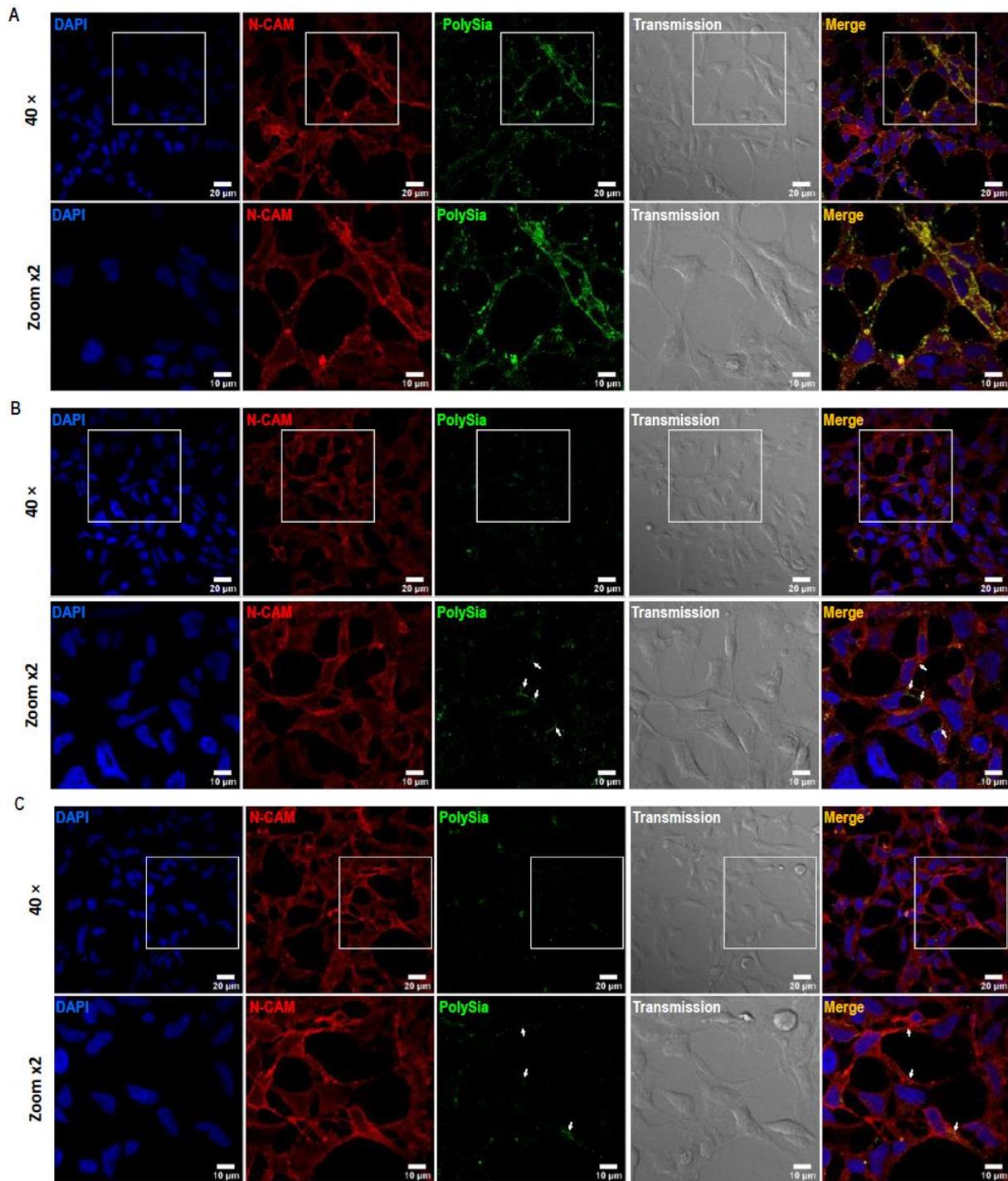


Figure 63: Exosialylation of HEK293 cells with *Cma* ST8Sia IV. After endoN treatment, sialylation reactions were carried out 4 h at 27 °C on fixed cells with 40 μ L of *Cma* ST8Sia IV and 500 μ M of donor substrates: A) CMP-Neu5Ac, B) CMP-Neu5Gc and C) CMP-Kdn. Polysialylation on cell surface is detected using anti-mAb735 antibody (green) and NCAM is detected using anti-CD56 antibody (red). Nuclei are stained with DAPI (blue). Outline of the cells is visualized by transmission in grey. DAPI, NCAM and polySia signals are superposed in Merge. Fluorescence was detected through a Carl Zeiss confocal microscope and Transmission correspond to transmitted light microscopy images. White arrows indicate polySias. Scale bars: 20 μ m for acquisitions at 40 \times and 10 μ m for the Zoom \times 2 pictures.

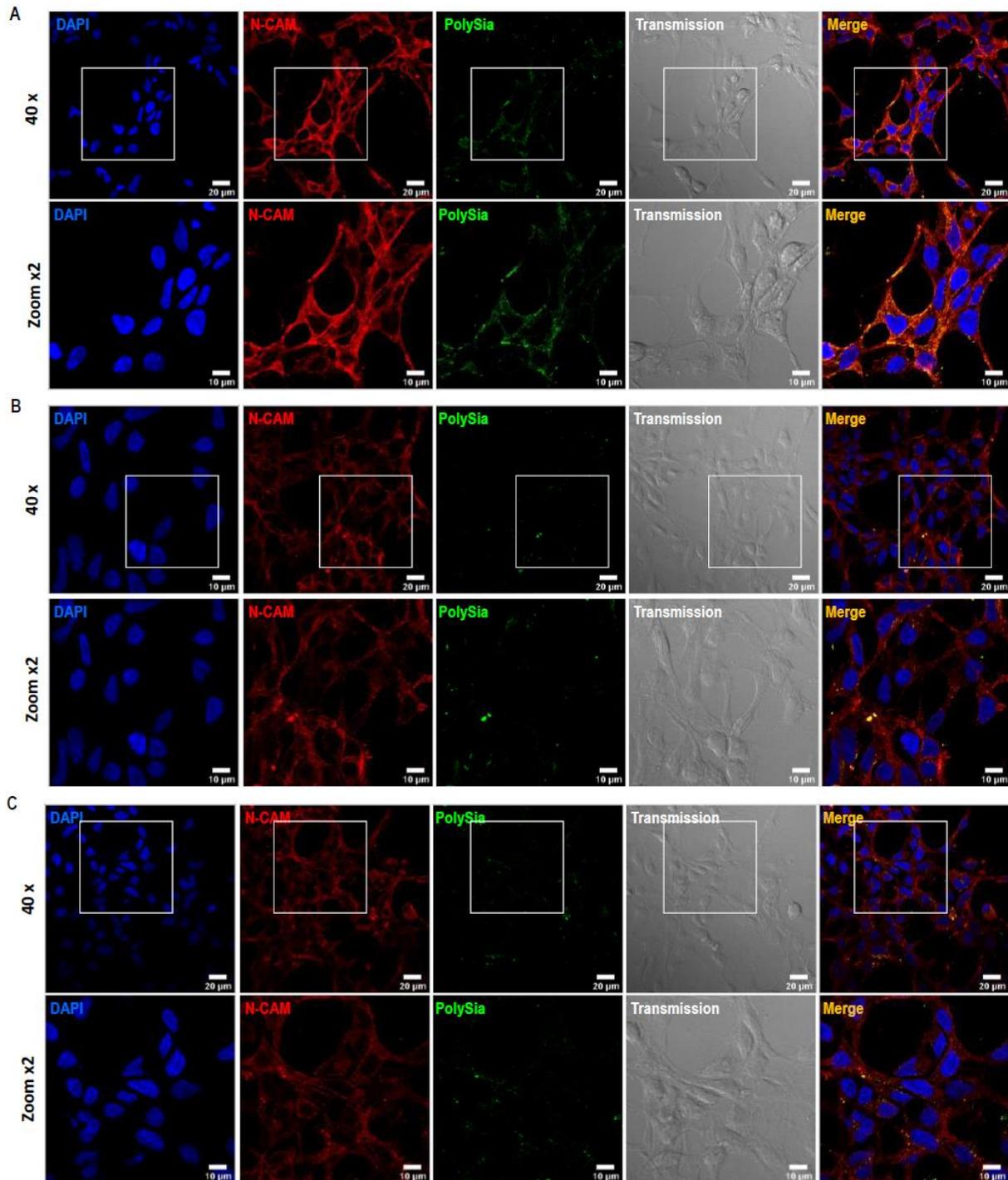


Figure 64: Exosialylation of HEK293 cells with *Hsa* ST8Sia IV. After endoN treatment, sialylation reactions were carried out 4 h at 27 °C on fixed cells with 200 ng of human ST8Sia IV and 500 μM of donor substrates: A) CMP-Neu5Ac, B) CMP-Neu5Gc and C) CMP-Kdn. Polysialylation labelling on cell surface is detected using anti-mAb735 (green) and anti-CD56/NCAM (red) antibodies. Nuclei are stained with DAPI (blue). Outline of the cells is visualized by transmission in grey. DAPI, NCAM and polySia signals are superposed in Merge. Fluorescence was detected through a Carl Zeiss confocal microscope and Transmission correspond to transmitted light microscopy images. Scale bars: 20 μm for acquisitions at 40 × and 10 μm for the Zoom ×2 pictures.

Consequently, we performed polysialylation reactions on endoN-treated HEK293 cells during 4 hours at 27°C with 500 μM of CMP-Neu5Ac, CMP-Neu5Gc or CMP-Kdn and either the human or *Cma* ST8Sia IV. We detected polySia with mAb735 using CMP-Neu5Ac for both ST8Sia IV (Fig. 63A and 64A). In contrast, using CMP-Neu5Gc and CMP-Kdn, polySia was detected only with the *Cma* ST8Sia IV (Fig. 63B,C) but not with the *Hsa* ST8Sia IV (Fig. 64B,C). Although weak, green fluorescence detected with mAb735 after sialylation with *Cma* ST8Sia IV and CMP-Neu5Gc or CMP-Kdn appears more surrounding cell surface (Fig. 64B,C) whereas fluorescence detected with human ST8Sia IV in the same conditions appears more punctiform or as aggregates, suggesting a fluorescence less specific to polysialylation on cell surface.

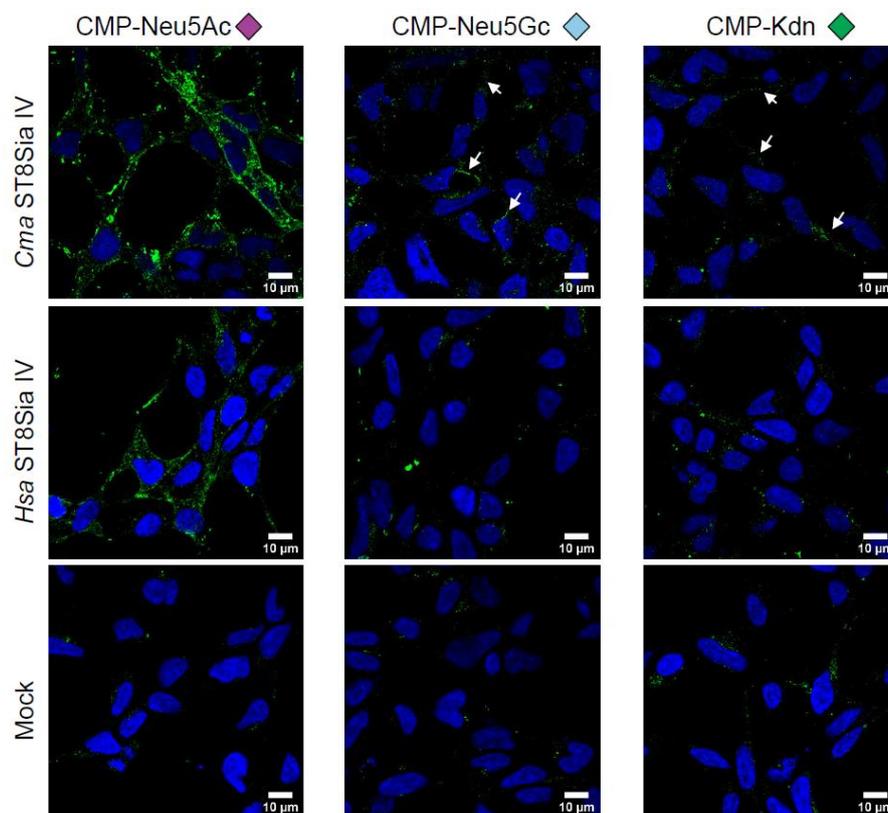


Figure 65: Summary of exosialylation on HEK293 cells using ST8Sia IV. Merge of polySia and DAPI signals obtained for both ST8Sia IV and Mock after exogenous polysialylation detailed in Figures 60 and 61. White arrows indicate polySia formed by *Cma* ST8Sia IV with CMP-Neu5Gc and CMP-Kdn.

Finally, we demonstrated using exogenous sialylation on cell surface that polySTs could effectively synthesized polySia chains on HEK293 cells (Fig. 65) and we have added new evidences correlated with our previous results and demonstrating a distinct specificity between the human and *Cma* ST8Sia IV enzymes towards donor substrates.

To conclude the main chapter of my thesis results, I developed various biomolecular, biochemical, analytical and cellular approaches in order to highlight *Cma* polySTs substrate specificities, distinct from their human homologues and which undoubtedly contribute to the diversity of polysialylation occurring in salmonids. Starting from the production of these enzymes, I had to precisely analyze their sequences to optimize their molecular cloning and to maximize their production by cellular transfection. In parallel, the preparation of various donor and acceptor substrates allowed me to compare their enzymatic activity using firstly MPSA and determining kinetic parameters in the best sialylation conditions with unnatural donor substrate CMP-SiaNA1. Then, structural assessment of polySia synthesized by polySTs gave me the opportunity to use WB and analytical approaches based on Sia and polySia detection and demonstrating new evidences of the usefulness of mAb735 and endoN towards a polySia composed by Neu5Ac, Neu5Gc and Kdn. Using cell surface as an acceptor platform, I contributed to show that polySTs could be useful for exogenous sialylation on cell surface. All these results converge on distinct specificities towards donor substrates between human and *Cma* polySTs. Since no crystal 3D-structure of vertebrate polySTs has been obtained yet, molecular phylogeny, modeling and docking strategies will be developed to shed light on the molecular basis of these differential substrate recognition by vertebrate polySTs and better understand their biological function.

Furthermore, the second part of my thesis was dedicated to explore if autopolysialylation exist for fish polysialyltransferases, as described for human polySTs (Mühlenhoff *et al.*, 1996b ; Close and Colley, 1998 ; Close *et al.*, 2001) and if this potential structural modification has a role for the these enzymes like *Cma* ST8Sia IV. These investigations are presented in next chapter below.

II. AUTOPOLYSIALYLATION OF POLYSIALYLTRANSFERASES : A REQUIREMENT FOR ENZYMATIC ACTIVITY OF *Cma* ST8Sia IV ?

Discovered in the 90's, autopolysialylation is an enzymatic reaction reported for mammalian polySTs which are able to catalyze the transfer of Neu5Ac to generate polySia on their N-glycans (Mühlenhoff *et al.*, 1996b ; Close and Colley, 1998 ; Close *et al.*, 2001). Indeed, human polySTs ST8Sia II and ST8Sia IV have N-glycosylation sites identified by mutagenesis (Table 10) which could be autopolysialylated and subsequently affect their activity. The objective of this short study was to determine if fish polySTs like *Cma* ST8Sia IV could be autopolysialylated and if this potential autopolysialylation is required or affects its enzymatic activity.

A. PRODUCTION OF *Cma* ST8Sia IV IN DIFFERENT CELL TYPES

Firstly, we focused on *Cma* ST8Sia IV, the most active fish polyST studied in the first chapter, that we produced in different cell lines to modify its glycosylation and influence its ability to be autopolysialylated. In comparison to HEK293 cells, we chose HEK293S which are GnTII deficient cells and carry only Man₅GlcNAc₂ N-glycans and no complex N-glycans (Lin *et al.*, 2014), suggesting no possibility for autopolysialylation. We also performed cellular transfection in CHO K1 cells which have no ST6Gal enzymes and no α 2,6-sialylation as well as CHO-Lec2 cells which are SLC35A1 deficient and display only asialylated glycans.

Cma ST8Sia IV transfections led to distinct levels of secretion according cell types (Fig. 66), particularly with lower expression in CHO cell lines than in HEK293 and in HEK293S cells.

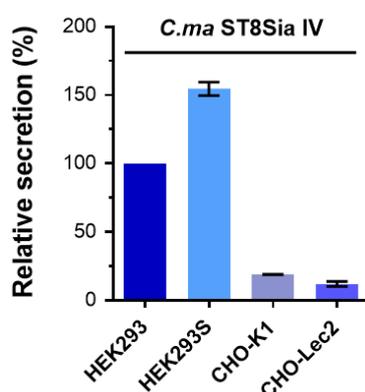


Figure 66: Production and secretion of *Cma* ST8Sia IV in different cell lines. p3×FLAG-CMV9Δ28ST8Sia IV plasmid was transiently transfected using lipofectamine on HEK293, HEK293S, CHO-K1 and CHO-Lec2 cells. Media culture were collected 72 hours post-transfection, loaded on SDS-PAGE and analyzed onto WB. Bands intensity was measured with ImageJ and normalized towards HEK293 production used as standard. Data were reported onto GraphPad Prism 6.

B. AUTOPOLYSIALYLATION OF *CMA ST8SIA IV* PRODUCED IN DIFFERENT CELL TYPES

Using the same quantity of enzymes secreted in the medium culture of transfected cells, we explored autopolysialylation status of each *Cma ST8Sia IV* production on WB. We checked initial autopolysialylation and we performed sialylation reactions during 4 hours at 27°C using 100 µM of CMP-Neu5Ac and various volumes of *Cma ST8Sia IV* with 20 µL for HEK293 production and adjusted to secretion levels obtained in other cell types (Fig. 66). We observed that there is no or very little pre-existing autopolysialylation in each cell type compared to Mock control (Fig. 67). Interestingly after sialylation reaction, we observed strong autopolysialylation for production in HEK293 and in CHO-K1 cells which is completely abolished after PNGase-F treatment. In contrast, no autopolysialylation was detected for production in CHO-Lec2 (Fig. 67), consistent with their asialylated N-glycans profile. Low levels of polysialylation are detected for the HEK293S production and which remain present after PNGase-F, indicating that *in vitro* autopolysialylation could also be generated on other pre-existing sialylated structures than N-glycans.

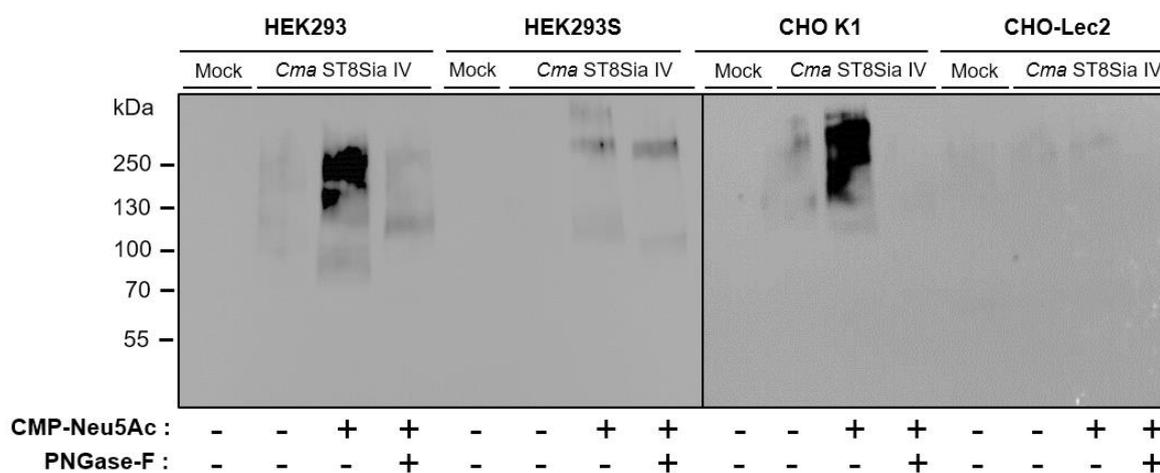


Figure 67: Autopolysialylation status of each *Cma ST8Sia IV* production. Initial autopolySia was checked for production in HEK293, HEK293S, CHO-K1 and CHO-Lec2 before any sialylation reaction and compared with a Mock control. Then, sialylation reactions were carried out at 27°C for 4 h with ST8Sia IV (20 µL for HEK293 production and adjusted according Fig. 66) and 100 µM of CMP-Neu5Ac. Part of polysialylated reaction mixtures was denatured in SDS buffer, Triton X-100 (10%) and treated overnight at 37°C using 5 ng of PNGase-F in acetate buffer (pH 6.0, 50 mM). After loading on SDS-PAGE and transfer onto nitrocellulose membranes, blots were incubated with the anti-polySia mAb735 (1:2000) and detected with femtoECL at 30 sec. Positions of the molecular weight control (kDa) are indicated on the left.

Although needing to be replicated with O-glycanase and endoN treatments, these preliminary results suggest that autopolysialylation mainly occurs on sialylated N-glycans of *Cma* ST8Sia IV but might also be generated on sialylated O-glycans as *Cma* ST8Sia IV contains 12 potential O-glycosylation sites as predicted using NetOGlyc 4.0 analysis (Appendix 1).

C. ENZYMATIC ACTIVITY OF *CMA* ST8Sia IV PRODUCED IN DIFFERENT CELL TYPES

Finally, we checked enzymatic activity of each *Cma* ST8Sia IV in MPSA. Using 100 μ M of CMP-SiaNAI and the same enzyme quantities, we observed that each ST8Sia IV are significantly active although at different levels (Fig. 68). Indeed, *Cma* ST8Sia IV produced in HEK293 and HEK293S exhibit the highest activity and this activity is slightly reduced for the production in CHO-Lec2. To be noted, only few levels of SiaNAI transferred were detected for *Cma* ST8Sia IV produced in CHO-K1 cells (Fig. 68) whereas this enzyme is produced in relatively same levels than the one produced in CHO-Lec2 (Fig. 66) and is able to polysialylate itself with CMP-Neu5Ac like the one produced in HEK293 cells (Fig. 67).

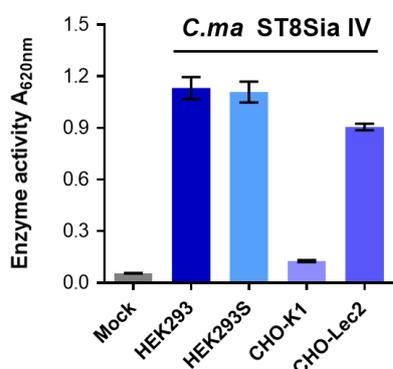


Figure 68: Enzymatic activity of each *Cma* ST8Sia IV production using CMP-SiaNAI in MPSA. Sialylation reactions were performed 4 h at 27°C using 400 ng of ALCAM with 100 μ M of CMP-SiaNAI and the same quantities of *Cma* ST8Sia IV adjusted from Fig. 66. Mock is representative to each culture medium control collected from HEK293, HEK293S, CHO-K1 and CHO-Lec2 cells.

To summarize, this approach based on different glycosylation and production of a fish polyST corroborate that complete N-glycosylation is not required for enzymatic activity of *Cma* ST8Sia IV, notably to catalyze the transfer of SiaNAI onto another acceptor substrate. Moreover, autopolysialylation appears not essential for catalytic activity of *Cma* ST8Sia IV although it might modulate substrate recognition and processivity of the enzyme as described for human polySTs (Mühlenhoff *et al.*, 2001).

III. ENZYMATIC SPECIFICITIES OF SALMONID POLYSIALYLTRANSFERASES: DIFFERENCES AT A STRUCTURAL LEVEL ?

Bioinformatic tools and particularly the rise of AlphaFold2 (Appendix 1) during my thesis allowed us to generate models which helped us to decipher structure-function relationships of polySTs and to explain their enzymatic specificities at molecular and structural levels.

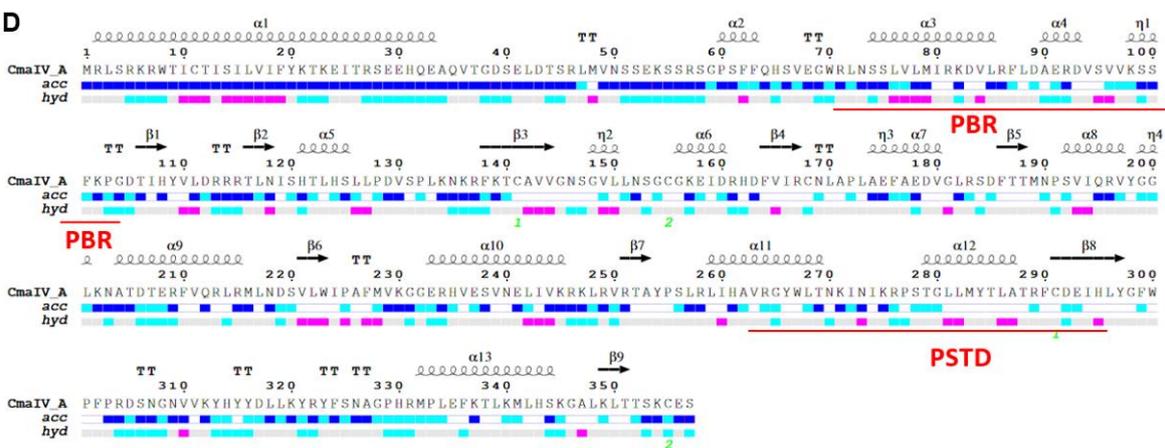
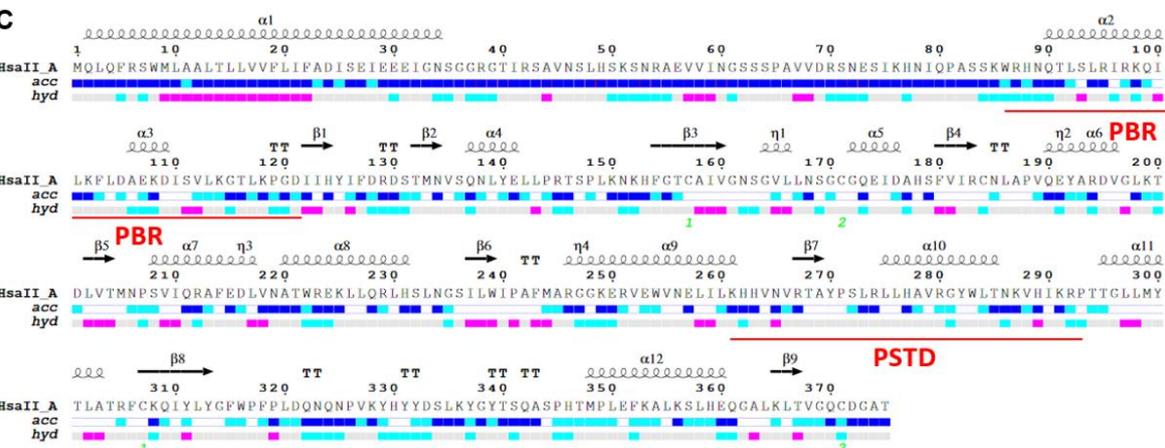
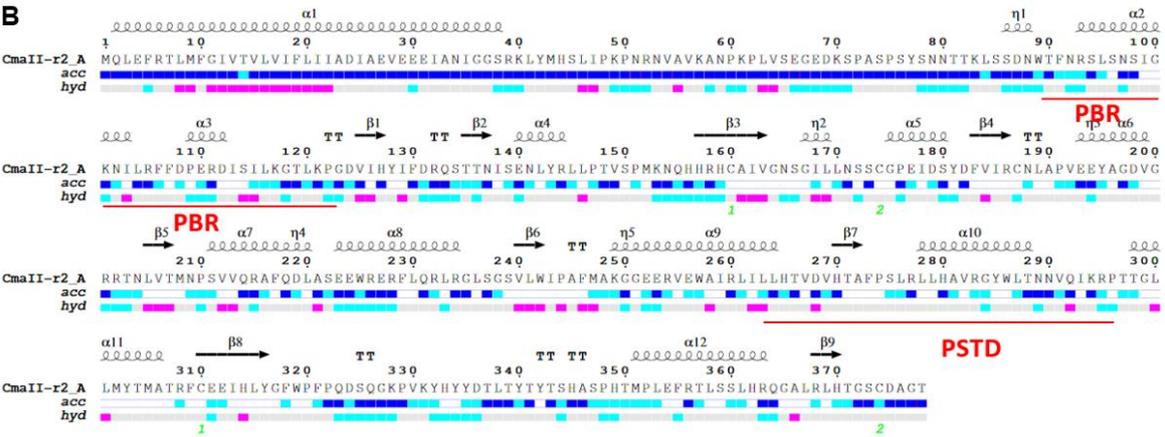
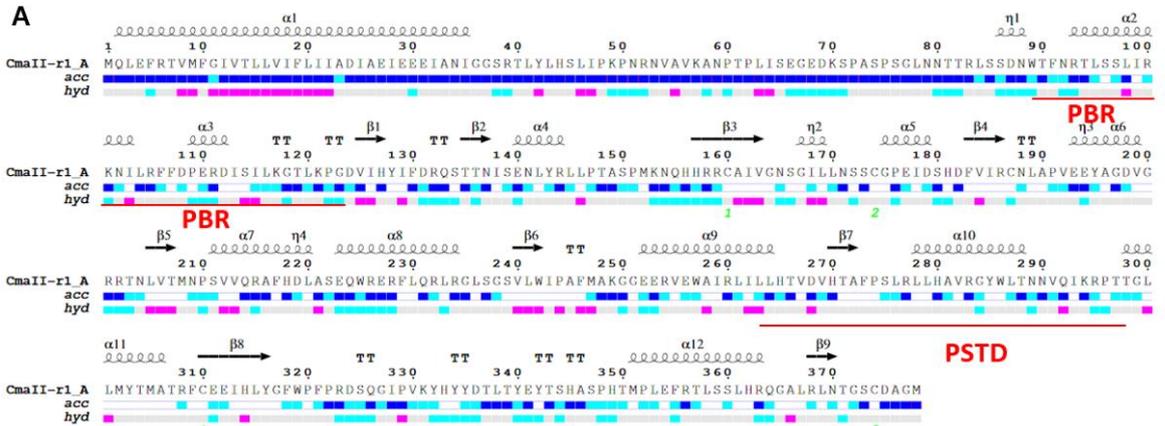
A. MODELING OF MONOMERIC POLYSTs AND COMPARISON OF 3D STRUCTURES

Introduced by Dr G. Brysbaert (UGSF, Lille) for the use of AlphaFold, we submitted the two human and the three salmonid polySTs full length sequences (Appendix 3) in AlphaFold-monomer v.2.2.0 and we obtained structural models based on MSA and previous crystallographic data as reported for ST8Sia III and ST6Gal I (Volkers *et al.*, 2015 ; Kuhn *et al.*, 2013). Indeed, we generated models of high consistent predictions for the five polySTs with top DockQ scores of 84-85.

Firstly, we analyzed these models on ENDScript tool (Appendix 1) to have a comprehensive view of primary to quaternary human and *Cma* polySTs structures (Fig. 69) and on ESPript tool (Appendix 1) to compare sequence similarities and secondary structure information resulting from sequences alignment (Fig 70).

Sequence organization analysis on ENDScript confirmed a GT-A like conformation in $\alpha/\beta/\alpha$ folds for each polyST. Indeed, *Cma* ST8Sia II-r1, *Cma* ST8Sia II-r2 and *Hsa* ST8Sia II are composed by 12 α -helices and 9 β -strands with notably their PBR motif which encompasses 2 α -helices and PSTD by 1 β -strand and 1 α -helix (Fig. 69A,B,C). *Cma* and *Hsa* ST8Sia IV are organized with 13 and 14 α -helices respectively and 9 β -strands with particularly their PBR region which includes 2 α -helices and PSTD by 1 β -strand and 2 α -helices (Fig. 69D,E).

As we mentioned in Venuto *et al.*, 2020a, alignments between each ST8Sia II/II-r showed 11 aa substitutions in PBR and 7 in PSTD whereas alignments between each ST8Sia IV showed 5 aa substitutions in PBR and 4 in PSTD. These aa substitutions in PBR and PSTD could be involved in donor substrate recognition although other aa substitutions identified with ESPript in the catalytic domain could modify interactions with substrates through aa properties (Table 19).



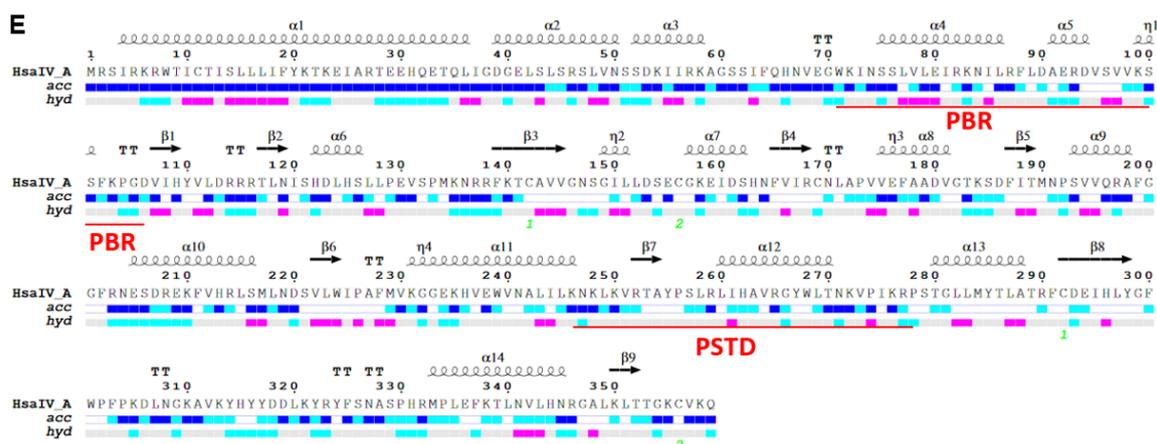


Figure 69: ENDScript analyses for structural sequence organization of the 5 polySTs. Representation of *Cma* ST8Sia II-r1 (A), *Cma* ST8Sia II-r2 (B), *Hsa* ST8Sia II (C), *Cma* ST8Sia IV (D) and *Hsa* ST8Sia IV (E) protein structures generated with AlphaFold2. PBR and PSTD motifs are underlined in red. Relative accessibility (acc) and hydrophobicity (hyd) of aa are colored from low to high values (highly accessible for dark blue and hydrophobic to hydrophilic from blue to pink). Pairs of cysteins are annotated in green.

Table 19: Summary of ESPrpt analyses between human and coregone polySTs sequences. Significant changes of aa properties in the catalytic domain are reported between ST8Sia II-r and ST8Sia IV sequences.

Motif	Protein structure	<i>Hsa</i> ST8Sia II	<i>Cma</i> ST8Sia II-r1	<i>Cma</i> ST8Sia II-r2
PBR	α2-helix	Arg95	Leu98	Ser98
-	α4-helix	Glu141	Arg144	Arg144
-	Loop	Arg145 Lys152	Thr148 Gln155	Thr148 Gln155
-	β3-strand	Gly155 Thr156	Arg158 Arg159	Arg158 His159
SM L	Loop	Ser179	Asp182	Asp182
	α6-helix	Arg194	Gly197	Gly197
	Loop	Leu198	Arg201	Arg201
-	α7-helix	Glu215	His218	Gln218
-	α8-helix	His231	Arg234	Arg234
-	α9-helix	Lys249 Glu257	Glu252 Arg260	Glu252 Arg260
PSTD	α9-helix	Lys261	Leu264	Leu264
PSTD	β7-strand	Arg267	His270	His270
		Lys287	Asn290	Asn290
SM S	β8-strand	Lys308	Glu312	Glu312
-	Loop	Leu320	Arg323	Gln323
Family motif d	Loop “lid domain”	Lys336	Thr339	Thr339
		Gly338	Glu341	Thr341

SM VS	α 12-helix	Gln342	His345	His345
-	α 12-helix	Lys356	Ser359	Ser359
		Glu360	Arg363	Arg363
Catalytic domain	Protein structure	<i>Hsa</i> ST8Sia IV	<i>Cma</i> ST8Sia IV	
PBR	α 3-helix	Glu80	Met79	
		Asn84	Asp83	
	α 5-helix	Asp123	Thr122	
	α 7-helix	Glu178	Ala179	
	α 9-helix	Glu205	Ala204	
		Ser206	Thr205	
		Arg208	Thr207	
		His213	Gln212	
		Ser216	Arg215	
	α 10-helix	Trp239	Ser238	
PSTD	α 10-helix	Asn247	Arg246	
Family motif d	Loop “lid domain”	Lys310	Asn309	
-	Loop	Asp319	Leu318	

Thanks to the expertise of Dr Elin Teppa, we precised the 3D structure of polySTs. More precisely, catalytic domain of each polyST display very high scores of confidence, notably in the four SMs, as represented in blue for *Cma* ST8Sia IV (Fig. 70A). In contrast, N^{ter} cytoplasmic tail, transmembrane domain and stem region appear less reliable with high Predicted Aligned Error (PAE) and low predicted local distance difference test (pLDDT) scores.

Comparing *Hsa* and *Cma* ST8Sia IV, we observed that their catalytic domains are highly superposed with Root Mean Square Deviation (RMSD) for alpha carbon of 0.394 Å. The two pairs of Cys involved in disulfide bonds for *Hsa* ST8Sia IV are conserved for *Cma* ST8Sia IV (i.e. Cys²⁹¹-Cys¹⁴¹ and Cys¹⁵⁵-Cys³⁵⁵) (Fig. 70B), suggesting an important role for protein stability. Moreover, we also detected some disordered regions in the catalytic domain, particularly a flexible loop of 12 residues (LLKYRYFSNAGP) for *Cma* ST8Sia IV corresponding to the family motif “d” between SM III and SM VS which could be involved in enzymatic specificities. Otherwise, preliminary analysis of specificity-determining positions (SDPs) suggests 7 SDPs scattered in the 3D structure, mostly located in flexible regions (Fig. 70C) that might be involved in substrates interactions.

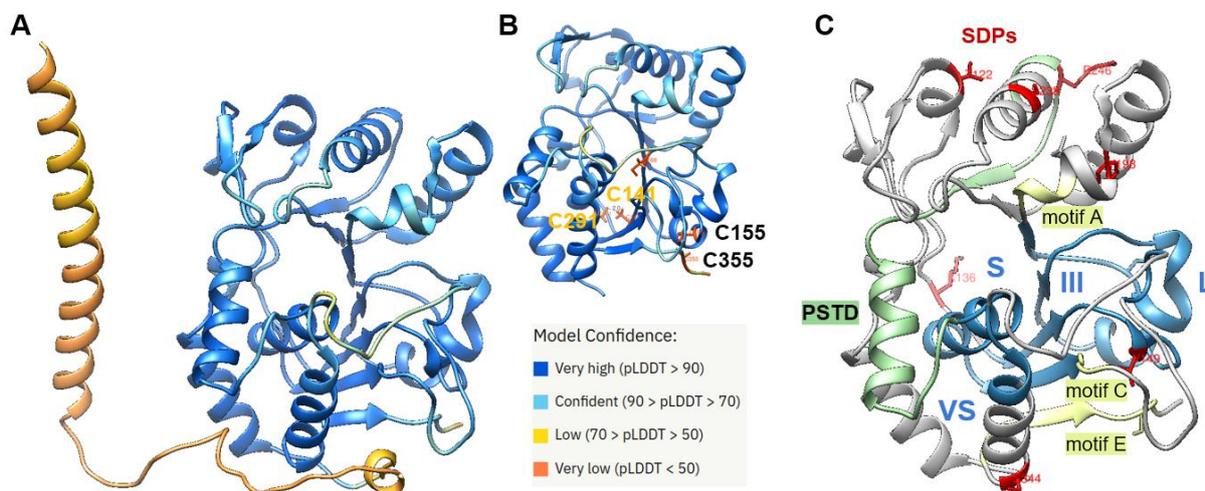


Figure 70: 3D structural modeling of *Cma* ST8Sia IV. (A) Cartoon representation of the full length 3D structure of *Cma* ST8Sia IV colored from low (orange) to very high (blue) confidence based on pLDDT score. (B) As described in *Hsa* ST8Sia IV, two pairs of Cys (i.e. Cys²⁹¹-Cys¹⁴¹ and Cys¹⁵⁵-Cys³⁵⁵) are involved in disulfide bonds for *Cma* ST8Sia IV, correctly oriented and at a distance of 2.0 Å. (C) Representation of characteristic motifs in the catalytic domain of polySTs. SMs S, L, III and VS are represented in blue. Family motif A, C and E are colored in beige. PSTD motif is represented in green and Specificity Determining Position (SDPs) are reported in red.

Since in GT-A fold variant 2, structure STs catalysis results from flexibility and movement of disorder loops (Audry *et al.*, 2011), we could hypothesized that these specific aa changes could correctly orient the nucleophilic group and confer substrate specificities for *Cma* polySTs (Harduin-Lepers, 2010 ; Harduin-Lepers, 2013 ; Moremen and Haltiwanger, 2019).

B. HOMO- AND HETERODIMERIZATION OF POLYST FOR ENZYMATIC ACTIVITY

i. Co-incubation for enzymatic assay in MPSA

Because GTs could act in cooperation to increase their activity through complex stability and conformation (Khoder-Agha *et al.*, 2019), we have incubated both polySTs and measured their enzymatic activity on ALCAM with CMP-SiaNAI (Fig. 71). Co-incubations of *Cma* ST8Sia IV with *Cma* ST8Sia II-r1 or *Cma* ST8Sia II-r2 had no impact on enzymatic activity with CMP-SiaNAI (Fig. 71A, 71B), whereas a slight increase could be detected incubating each human polyST, suggesting they could cooperate to increase enzymatic activity (Fig. 71C). No enzymatic activity were detected co-incubating *Cma* ST8Sia II-r1 and *Cma* ST8Sia II-r2 on MPSA (Fig. 71D).

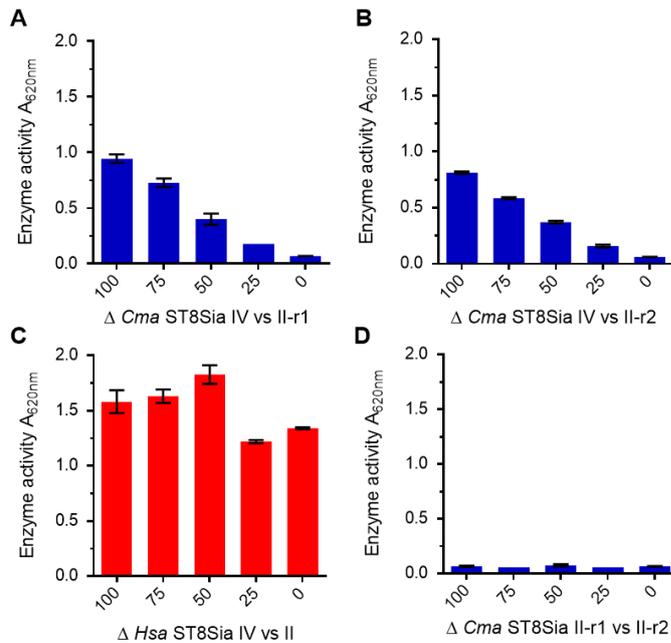


Figure 71: Potential polySTs cooperation after co-incubation in MPSA. After coating of 400 ng ALCAM, sialylation were done 4 hours at 27°C using the 3 coregone and the 2 human polySTs at various level. (A) With decreasing quantities of *Cma* ST8Sia IV and increasing quantities of *Cma* ST8Sia II-r1 (100% corresponding to 20 and 55 μL respectively). (B) With each *Cma* ST8Sia IV and II-r2. (C) With each human polyST (100% corresponding to 150 ng). (D) With each *Cma* ST8Sia II-r1 and II-r2.

However, we were not able to visualize interaction with anti-FLAG on WB after co-incubation of 3x-FLAG tagged truncated polySTs (data not shown), suggesting that soluble forms couldn't generate dimeric interactions. It could be interesting to perform co-transfection of full length enzymes and see if domains like stem region or TMD could be involved in dimerization. Moreover, it could be interesting to investigate potential cooperations using natural donor substrates like CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn to precise their enzymatic specificities.

ii. Homodimerization of polySTs

On the other hand, we investigated if potential interactions between polySTs, as homo- and heterodimerization, could influence the orientation of catalytic and recognition sites of polySTs and explain their enzymatic specificities. Recently developed to predict homodimeric and heterodimeric protein complexes with high accuracy (Evans *et al.*, 2022 ; Bryant *et al.*, 2022), AlphaFold-multimer was tested for potential homodimerization of the three *Cma* polySTs and the two human polySTs and for potential heterodimerization between them.

Among each full length sequences of polySTs, one consistent homodimer was generated for *Hsa* ST8Sia IV with an itp+ptm score = 0.82. Indeed, aa between R²⁵⁹ to T²⁷⁰ in the PSTD motif seem involved in closer interaction as shown with electrostatic potential surfaces representation (Fig. 73). Interestingly, no dimerization was detected for full length *Cma* ST8Sia IV whereas,

among 50 predictions, one significant homodimer model was obtained for a truncated sequence of *Cma* ST8Sia IV, preserving only high catalytic domain ($\Delta 122$) (Table 20). No homodimers with a significant iptm+ptm score were obtained for *Hsa* ST8Sia II, *Cma* ST8Sia II-r1 and *Cma* ST8Sia II-r2 after 50 models generated.

Table 20: iptm+ptm score prediction for homodimerization of ST8Sia IV sequences. Truncated sequences chosen are $\Delta 122$ for *Cma* ST8Sia IV and $\Delta 63$ for *Hsa* ST8Sia IV.

ST8Sia IV	Truncated sequences		Whole sequences		
	Specie	<i>Hsa</i>	<i>Cma</i>	<i>Hsa</i>	<i>Cma</i>
Itpm+ptm highest score		0.88	0.92	0.82	0.26

Consequently, we analyzed the interface of interaction for homodimer of *Cma* ST8Sia IV and we were able to detect 9 hydrogen bonds at the protein-protein interface (Fig. 72). Moreover, in comparison to *Hsa* ST8Sia III structure (PDB ID: 5BO9), the protein-protein interaction interface for *Cma* ST8Sia IV homodimer is completely different. The two H-bonds across the interface Asn¹⁴²-Asn¹⁴³ and Glu²⁶⁶-Ser¹³⁹ and the ionic interaction Arg²⁵⁸-Asp²⁶² (Fig. 16) described for the protein-protein interaction of ST8Sia III are not conserved in *Cma* ST8Sia IV explaining that the analogous interaction is not possible (Fig. 72). To be noted, this interaction might limit accessibility to catalytic His²⁶⁹ residues (Fig. 72).

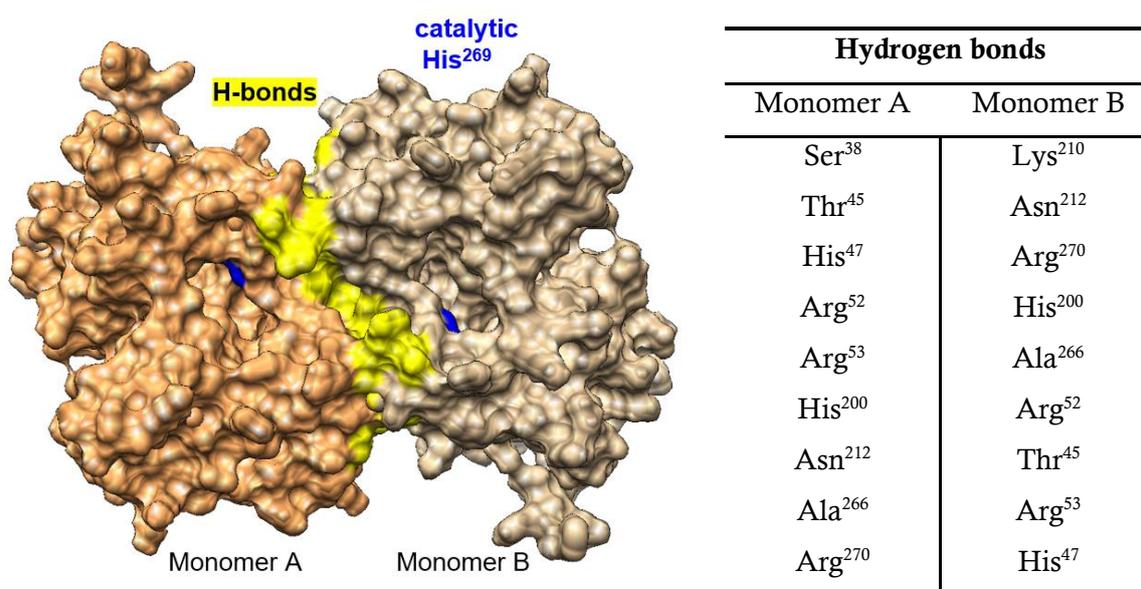


Figure 72: Surface representation for homodimeric interaction of *Cma* ST8Sia IV. Resulting from the top score in Table 20. Residues involved in the 9 hydrogen bonds (yellow) are reported at the interface of each monomer (colored in orange and beige). Catalytic His is colored in blue.

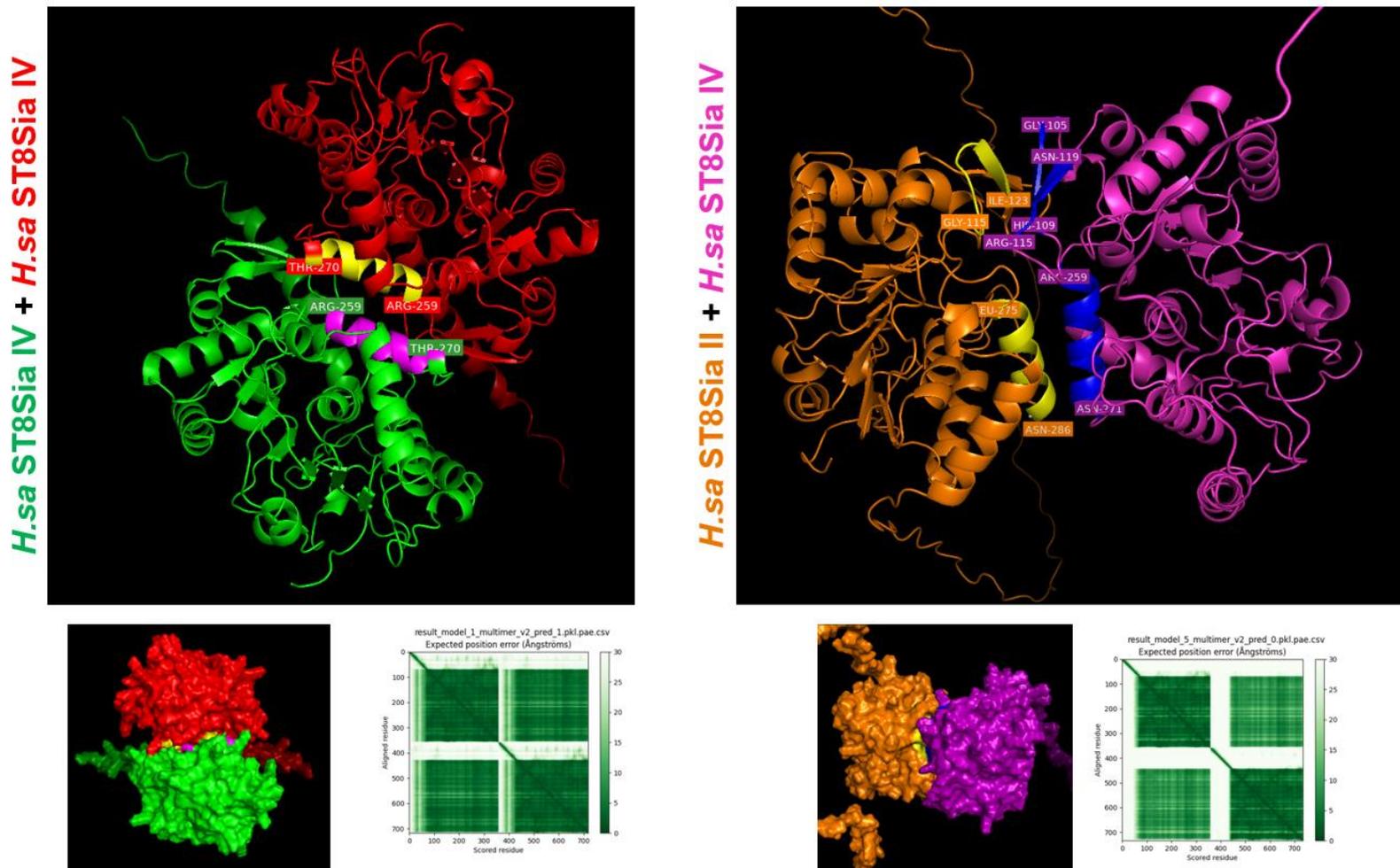
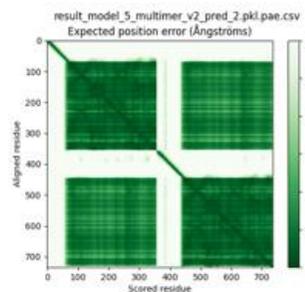
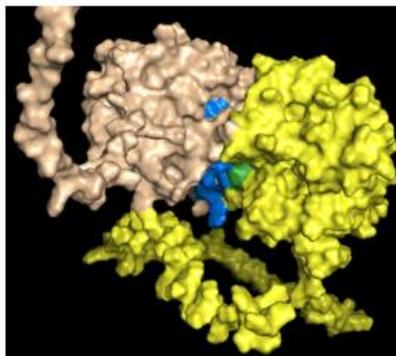
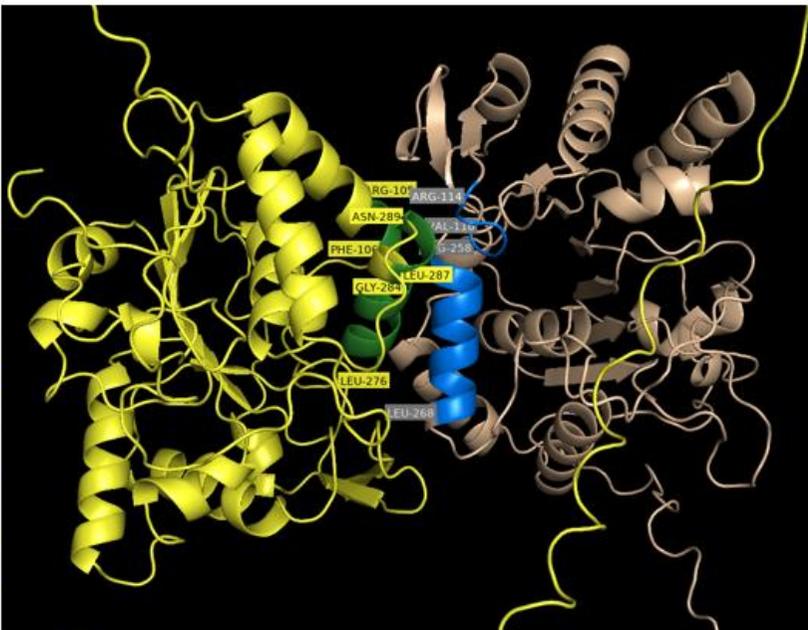


Figure 73: AlphaFold-multimer modeling for homo- and heterodimerization between human polySTs. Models were generated from AlphaFold-multimer for each polyST. PAE showed consistent models for homodimer with ST8Sia IV (left) and heterodimer (right). Each monomer are colored differently and, as shown in potential surface interactions, aa in PSTD and PBR regions which could be involved in interaction are highlighted in another color.

C.ma ST8Sia II-r1 + C.ma ST8Sia IV



C.ma ST8Sia II-r2 + C.ma ST8Sia IV

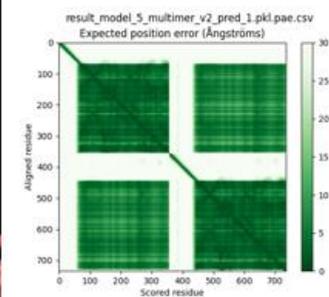
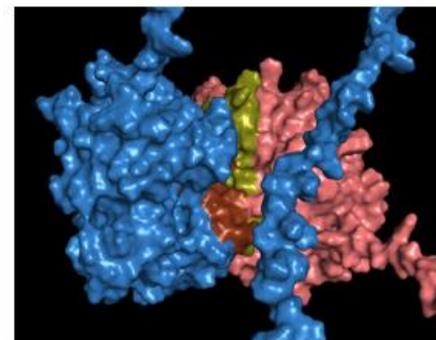
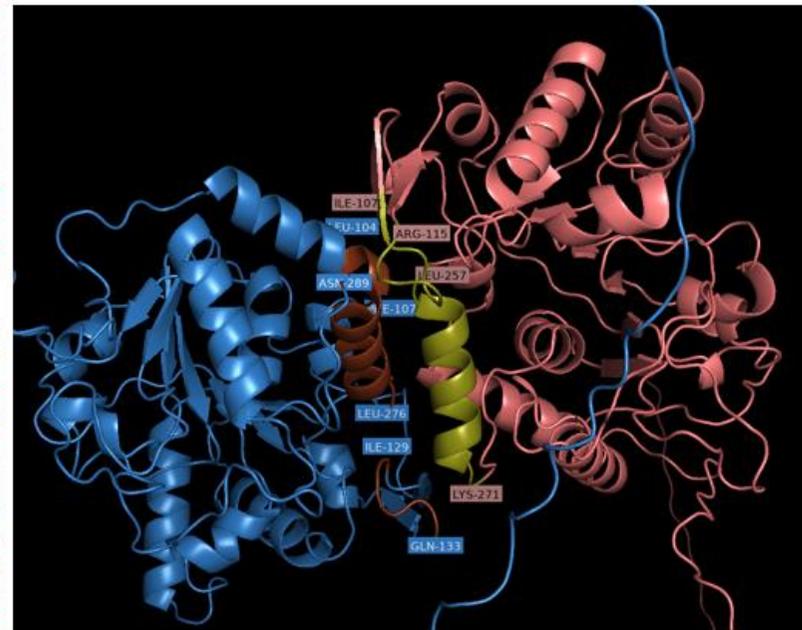


Figure 74: AlphaFold-multimer modeling for heterodimers between *C.ma* polySTs. Structural models were generated from AlphaFold-multimer for each polyST. PAE showed consistent models for *C.ma* heterodimers. Each monomer are colored differently and, as shown in potential surface interactions, aa in PSTD and PBR regions which could be involved in interaction are highlighted in another color.

In the future, it will be interesting to investigate if SDPs are substituted in *Cma* ST8Sia IV in its active site and if these substitutions could impact properties and interactions in homodimer model.

iii. Heterodimerization of polySTs

Furthermore, heterodimers were generated for human ST8Sia II and ST8Sia IV, the model with an highest iptm+ptm score was 0.66, relatively reliable. In this model, ST8Sia IV displays aa between R²⁵⁹ to N²⁷¹ in the PSTD and G¹⁰⁵ to N¹¹⁹ near to the PBR and ST8Sia II displays aa between L²⁷⁵ and N²⁸⁶ in the PSTD and G¹¹⁵ to I¹²³ near to the PBR which seem involved involved in closer interaction as shown with electrostatic potential surfaces representation (Fig. 73). Interestingly, this interaction could influence stability or enzymatic activity as described in biochemical assays (Fig. 71).

About coregone polySTs, heterodimers with *Cma* ST8Sia II-r1 or ST8Sia II-r2 and ST8Sia IV were generated, best models showed iptm+ptm scores of 0.75 and 0.74 respectively. As for human polySTs, *Cma* ST8Sia IV displays aa between R²⁵⁸ to L²⁶⁸ in the PSTD and V¹¹⁰ to R¹¹⁴ near to the PBR and ST8Sia II-r1 or ST8Sia II-r2 between L²⁷⁶ and N²⁸⁹ in the PSTD and L¹⁰⁴ to F¹⁰⁷ near to the PBR which seem involved in closer interaction as shown on surfaces representation (Fig. 74).

To conclude, these first investigations are a starting point to understand the functional organization of polySTs in complex which could be important for their enzymatic specificities. Particularly, it could be interesting to explore the role of these flexible loops for the recognition and binding of acceptor and donor substrates. Although needing to be investigated and completed at atomic level with molecular docking for binding substrate site analysis, biocomputational approaches offer interesting bases to decipher how polySTs could interact in complex and modulate their enzymatic specificities.

DISCUSSION

I. PRODUCTION OF POLYSTs

During my thesis, we reported the molecular cloning, production and biochemical characterization of the polySTs from *C.maraena*, a salmonid specie which have undergone a particular genome distribution with three polyST genes *st8sia2-r1*, *st8sia2-r2* and *st8sia4* (Venuto *et al.*, 2020a). From extraction of cDNA in *C.maraena* gonads and brain tissues, we engineered soluble N-terminally truncated 3xFLAG-tagged fused forms of the enzymes to optimize secretion of these recombinant proteins in the culture medium of transfected HEK293 cells as assessed by WB with an anti-FLAG antibody. To be noted, HEK293 cells remain the preferred host to produce difficult-to-express GTs such as polySTs requiring up to six partially polysialylated *N*-glycans and disulfide bonds to achieve proper folding and function (Mühlenhoff *et al.*, 2001 ; Close *et al.*, 2001; Moremen *et al.*, 2018; Jaroentomeechai *et al.*, 2022).

Our transfection results show that all these enzymes are not produced and secreted in the same proportions, which corroborates a recent study which showed that there is no correlation between the level of expression and the efficiency of secretion of human STs (Moremen *et al.*, 2018). A possible explanation relates to the plasmid construction designed, the instability index (II) and the nature of the enzymes, as shown for ST3Gal I (Vallejo-Ruiz *et al.*, 2001). Thus, we opted for constructs leading to the expression of homologous human and salmonid enzymes with a catalytic domain of equivalent size. Among the three salmonid polySTs, *Cma* Δ 28 ST8Sia IV was the most produced, in a mature glycosylated form and the most active for biochemical characterization. Concerning the two other *Cma* polySTs ST8Sia II-r1 and ST8Sia II-r2, we adapted their molecular construction, obtained better II and higher level of productions with Δ 35 form rather than Δ 28. Moreover, we noticed various levels of STs production depending on different parameters: the specie origin of the sequence (human, zebrafish or coregone) ; the construct and deletion designed in the stem region to optimize soluble secretion of active enzymes ; their level expression within retention in cell lysate compared to secretion in the culture medium ; the homologous or heterologous cell system used for transfection : human HEK293 and HEK293S, chinese hamster

ovary CHO-K1, CHO-Lec2 and Chinook salmon CHSE-214 cells ; the condition used in cell culture (\pm sodium butyrate, FBS, initial confluence, post-transfection collecting time). All these parameters appear crucial for STs engineering and further biochemical characterization.

II. CHEMO-ENZYMATIC SYNTHESIS OF CMP-SIA DONOR SUBSTRATES

Vertebrate STs are so-called “Leloir enzymes” using two donor and acceptor substrates. Here, for the first time in STs studies, we chemo-enzymatically synthesized various unnatural and natural donor substrates (CMP-SiaNA1, CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn) and checked their purity by ^{31}P NMR giving us the great advantage to minimize background and inhibitor side products due to CMP hydrolysis (Wang *et al.*, 2002). To be noted, we prepared CMP-Neu5Ac, CMP-Neu5Gc and CMP-SiaNA1 using the CSS from *Neisseria meningitidis* (Gilormini *et al.*, 2016a) whereas for CMP-Kdn synthesis, we also used rainbow trout CMP-Kdn synthetase (rtCSS) described to be more efficient than bacterial CSS to generate CMP-Kdn (Nakata *et al.*, 2001; Wu *et al.*, 2022). Indeed, 2 hours of synthesis are necessary for 100% conversion of CTP and Kdn (1:1) into CMP-Kdn with rtCSS whereas at least 5 hours are required with *N.meningitidis* CSS for a maximum of 90%, suggesting diverse affinity of CSS to use Kdn for CMP-Kdn synthesis. Moreover, we checked CMP-Sia stability by ^{31}P NMR after several months of storage at -80°C confirming there is no hydrolysis under these storage conditions.

III. BIOCHEMICAL CHARACTERIZATION OF POLYSTs

Up to now, mostly ST activities have been studied in optimized conditions of temperature during sialylation reaction, pH and salinity of cacodylate buffer and their sialylated products were monitored through a variety of strategies. As detailed in the introductory part, we can relate the detection of radiolabeled Sia from CMP- ^{14}C Neu5Ac (Nakamura *et al.*, 1991; Lepers *et al.*, 1990 ; Rohfritsch *et al.*, 2006), CMP side product quantified in colorimetric assay (Wu *et al.*, 2011), the use of lectins for their sugar affinity (Houeix and Cairns, 2019), fluorescent acceptor substrates like DMB-labelled Sia (Ehrit *et al.*, 2017 ; Guo *et al.*, 2019) or even the characterization of bacterial polySTs with *in vitro* and analytic assays (Willis *et al.*, 2008 ; Yu *et al.*, 2014 ; Guo *et al.*, 2020).

We previously developed a rapid and sensitive MPSA to distinguish human monoST activities (Noel *et al.*, 2018 ; Chang *et al.*, 2019). We took advantage of this enzymatic assay based on CuAAC click chemistry and bioorthogonal coupling between an azide-functionalized probe (azido-biotin) and an alkynyl-modified Sia reporter (SiaNAI) resulting from efficient chemo-enzymatic synthesis of CMP-SiaNAI (Gilormini *et al.*, 2016a). To be noted, our team has demonstrated that a pentynoyl-C5 modification on Sia did not impair donor substrate recognition and transfer activity of the human ST6Gal I and ST3Gal I (Noel *et al.*, 2017) as the CMP-Sia bound the catalytic pocket of human ST6Gal I with the side chain at C5 oriented outwards (Harrus *et al.*, 2020).

To be noted, MPSA has the advantage to be sensitive and does not require prior purification of the recombinant STs directly produced in the culture medium of the transfected HEK293 cells (Noel *et al.*, 2018). Firstly checking that MPSA could be used for polySTs, we reproduced similar data for optimum parameters (pH, temperature) with recombinant human polySTs as mentioned for human ST8Sia II (Guo *et al.*, 2020). Then, we extended this assay establishing the optimal conditions for the biochemical characterization of salmonid polySTs. Interestingly, we observed a higher thermal stability for *Cma* ST8Sia IV suggesting intrinsic properties of the protein acquired during evolution to adapt to variable environment (Siddiqui and Cavicchioli, 2006).

Since only a limited set of natural polysialylated proteins like NCAM, CD36, SynCAM-1, CCR7, ESL-1, NRP2 and polySTs have been reported (Table 5), we also used additional substrates like bovine fetuin, human DNase-I, BSM, orosomuroid, ALCAM/CD166, and PSGP which exhibit various sialylated N- and/or O-glycans and could serve as suitable acceptor substrates (Table 13) for the vertebrate polySTs. Notably using ALCAM, we observed a Michaelis Menten profile for each ST8Sia IV varying CMP-SiaNAI quantities used for sialylation reactions in MPSA. We noticed a significant difference of K_m between the two enzymes towards CMP-SiaNAI where the human ST8Sia IV exhibits higher affinity towards this donor substrate than the fish ST8Sia IV. Taking a step back, enzymes produced in the medium culture exhibit various components as cofactors, inhibitors or complexes which could refer closely to what exist in its cellular microenvironnement in comparison to commercial purified enzymes. Thus, K_m parameters

towards donor and acceptor substrates could be lower, affinities higher and less consistent with K_m in biological system for purified enzymes compared to enzymes produced in the medium culture.

Likewise, we obtained remarkable activities with the human and fish ST8Sia IV using the other acceptor substrates, whereas low or no activity could be detected for the human and fish ST8Sia II, respectively. These data are consistent with a previous report by Kitajima's group about very low levels of enzymatic activity for rainbow trout *O. mykiss* polySTs (Asahina *et al.*, 2006). As highlighted for human ST8Sia II *in vivo* (Vogt *et al.*, 2012), alkyn-C5 modification on Sia might impair donor substrate recognition and transfer catalysis for *Cma* ST8Sia II-r1 and ST8Sia II-r2. Since structural modeling demonstrated that ST6Gal I catalytic site binds CMP-Neu5Ac on its side chain at C5-position (Harrus *et al.*, 2020), it could be interesting to investigate the different polyST catalytic sites and their docking with CMP-SiaNA1 to explain our results on MPSA.

To summarize, these data showed that MPSA could be used as a biotechnological tool to determine polyST activities onto various glycoprotein acceptor substrates although this enzymatic assay displays noticeable limitations like the detection of polySTs processivity.

IV. POLYSIA SYNTHESIS AND ANALYSIS

i. Detection on WB using mAb735 and endoN

So far, polySia are detected mainly using mAb735, described to be one of the most powerful and affordable commercial antibody (Table 16) to recognize α 2-8-linked Neu5Ac polymers (Nagae *et al.*, 2013). Our first investigations using mAb735 after *in vitro* sialylation reactions on microplate didn't allow us to obtain significant polysialylation (data not shown), which prompted us to use higher reactives quantities and detected polySia on WB. Firstly optimizing polySia detection on WB using polysialylated cell lysates, we then explored enzymatic activity of each polyST after *in vitro* sialylation reactions. We clearly demonstrated an efficient level of polysialylation onto ALCAM/CD166 for each ST8Sia IV using CMP-Neu5Ac. To be noted, we also performed sialylation reactions with CMP-SiaNA1 but we were not able to detect any polySia with mAb735 on WB for each enzyme, suggesting either no polySiaNA1 are formed either polySiaNA1 are not detected with mAb735 (data not shown).

Furthermore, we investigated if mAb735 antibody could be applied for various polySia detection composed by Neu5Ac, Neu5Gc and Kdn residues. From structural data of mAb735 (PDB ID: 3WBD), we assessed molecular docking of mAb735 with polySia composed by Neu5Ac, Neu5Gc and Kdn residues and we confirmed our previous observations that mAb735 but could be unambiguously used to visualize other diverse polySia chains (Venuto *et al.*, 2020b). It could be interesting to assay the same molecular docking using modified Sia residues like SiaNA1.

Therefore, we confirmed that polySia of various nature could be synthesized using CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn only with fish polySTs, particularly with *Cma* ST8Sia IV and at lower levels with *Cma* ST8Sia II-r1 and *Cma* ST8Sia II-r2, whereas the human polySTs are able to use only CMP-Neu5Ac as detected on WB. Indeed, this is the first study showing that mAb735 recognizes *in vitro* enzymatically synthesized polySia composed by Neu5Gc and Kdn. Nevertheless, the differences in terms of level of various polySia detected might results either on enzymatic specificities of *Cma* polySTs towards each donor substrate either on mAb735 affinity towards polyNeu5Ac, polyNeu5Gc or polyKdn chains.

Remarkably, we also observed on WB that endoN treatment effectively remove polySia generated from these various donor substrates, highlighting that these polymers composed by Neu5Gc and Kdn residues could be recognized and cleaved by endoN, as previously suggested for Neu5G-containing polySia (Venuto *et al.*, 2020b). So far, endoN homotrimeric 3D structure (PDB ID: 3GVL) is incomplete and didn't allow us to analyze various polySia interactions by molecular docking as done with mAb735, although it should be investigated in the future.

Finally, the use of CMP-Neu5Ac, CMP-Neu5Gc or CMP-Kdn in polysialylation reactions highlighted the unique catalytic specificities for the three *Cma* polySTs towards these various Sia donors which are not observed for their human homologues. These data suggest a larger enzymatic promiscuity of the fish polySTs which are able to use various Sia donor substrates as proposed in the past for polySTs in *O. mykiss* ovaries (Angata *et al.*, 1994).

ii. Analytical methods to confirm polySia composition and size

In parallel, we confirmed the transfer of various polySia using various fluorescent derivatizations. After first trials using APTS-labeling and FACE analysis, we opted for DMB-labelling, more adapted to detect Sia residues. Analyzed on HPLC, we confirmed that *Cma* ST8Sia IV could generate polySia composed by Neu5Ac, Neu5Gc and Kdn residues although Sia quantities were limited. For further polySia analytical detections, it will be necessary to scale up and obtain more quantities after polySia reactions. Indeed, it appeared difficult to quantify Sia and polySia after enzymatic synthesis with vertebrate polySTs using DMB analytical methods. Manufacturer like Ludger in their LudgerTagDMB Analysis Kit (LT-KDMB-A1) recommend typical starting amount of 50 µg of highly sialylated glycoprotein or 100-500 µg of slightly sialylated glycoprotein. Thus, purification strategies have to be optimized to concentrate and enrich polysialylated products on column or beads for higher analytical detection.

Moreover, we could opt for endoN-affinity based strategies since iEndoN exhibit higher affinity towards polySia compared to mAb735 (Schwarzer *et al.*, 2009). Indeed, iEndoN, which conserves its lectinic domain (Yu *et al.*, 2014), has been a very useful tool to enrich diverse polySia synthesized *in vitro* and to strengthen our results of the donor specificities for *Cma* ST8Sia IV. Likewise, iEndoN-GFP was used as fluorescent lectin to detect polySia on microplates and on WB after polysialylation reaction of *Cma* ST8Sia IV with CMP-Neu5Ac, although no significant or reproducible data could be obtained due to low levels of iEndoN-GFP (< 1.5%) bound onto acceptor and high background of Mock control (data not shown).

To summarize, the next challenge will be to produce enough polysialylated glycoconjugates using these polySTs specificities and methodologies reported for extraction and purification of polysialylated glycoconjugates (size exclusion chromatography, anionic and acidic charges) or more specific methods with affinity-based and the use of mAb735 and endoN on beads prior analytical detection (Guo *et al.*, 2019). This large-scale production will contribute to apply these bioactive polysialylated products for further biotechnological applications.

V. POLYSIALYLATION REMODELING ON CELL SURFACE

As previously reported with monoSTs (Noel *et al.*, 2017; Sun *et al.*, 2016), exogenous cell surface sialylation has the advantage to use living cell systems without any crosstalk with endogenous enzymes of the Sia pathway (Sun *et al.*, 2016). However, polysialylation on cell surface requires relevant cell system with good acceptor substrates expressed like adherence molecules as ALCAM, SynCAM-1 or NCAM. After verifying that HEK293 cells express NCAM initially polysialylated on its cell surface, we removed polySia using endoN and we conducted exogenous sialylation with the fish and human ST8Sia IV. We used the natural donor substrates CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn and detected polySia with mAb735 antibody. These results are consistent with WB data confirming specificities for the fish ST8Sia IV where polySia on cell surface are detected with the three natural donor substrates, although not at the same intensity.

Furthermore, we demonstrated that these enzymes could be useful for remodel cell surface sialylated glycoconjugates using various CMP-Sia donor substrates.

Other strategies on cell surfaces were tried during my thesis. Notably, we performed exogenous sialylation using the two ST8Sia IV with CMP-SiaNAI coupled to azido-biotin and we remarked a relative labeling on HEK293 cells after biotin detection (data not shown) although further investigations need to be done to check potential polySia formation with SiaNAI residues. It could be interesting to compare these data to those using SiaNAI precursor incorporated in bacteria and the subsequent detection of capsular polySia by click-chemistry which colocalized with mAb735 detection (Rigolot *et al.*, 2023), suggesting incorporation of SiaNAI into polySia chains.

On the other hand, we used other cell types like LAN-5 (ACC 673) and human neuroblastoma cells SH-SY5Y (ACC 209) cells, known to carry polysialylated NCAM. Although needing to be improved, preliminary results tend to show that fish ST8Sia IV uses CMP-Neu5Ac and CMP-Kdn to generate polySia on LAN-5 cell surface (data not shown). Finally, other tools and strategies could be explored like the use of other antibodies as 12E3 reported to recognize polysialylated NCAM, the use of iEndoN-GFP to detect various polySia without secondary incubation required. Flow cytometry analysis could also be performed to obtain quantitative variations of polySia on

cell surface of living cells although it required optimization as preliminary results were not convincing (data not shown). CHSE-214 salmonid cells could also be a good model as they naturally express salmonid polySTs. Feeding them with Neu5Ac, Neu5Gc and Kdn residues could be an option to investigate if various polySia are detected on cell surface using FACS analysis or other analytical approaches as previously described.

VI. AUTOPOLYSIALYLATION OF *Cma* ST8Sia IV

So far, it has not been clearly demonstrated that autopolysialylation of polySTs is essential for their enzymatic activity and for the recognition of donor and acceptor substrates. The human ST8Sia II and IV have 2 N-glycosylation sites on their protein sequence which can serve as acceptor for autopolysialylation. While several studies suggest that autopolysialylation is necessary for enzymatic activity of the two human polySTs (Mühlenhoff *et al.*, 2001; Close *et al.*, 1998 ; Bhide *et al.*, 2018), my results obtained producing salmonid polySTs in different cell types seem discordant. If any *Cma* ST8Sia IV versions produced in HEK293, HEK293S, CHO-K1 and CHO-Lec2 cells were initially autopolysialylated, we observed significant differences incubating them with CMP-Neu5Ac. Indeed, enzymes produced in HEK293 and CHO-K1 are able to be autopolysialylated, as the commercial human enzymes (data not shown) whereas low or no autopolysialylation were detected for the productions in HEK293S and CHO-Lec2 respectively. Interestingly, the version produced in CHO Lec2, non able to be autopolysialylated, retains an enzymatic activity comparable to *Cma* ST8Sia IV produced in HEK293 in MPSA, suggesting that autopolysialylation of *Cma* ST8Sia IV is not essential for its enzymatic activity.

In the cellular context, polySTs have their TMD embedded in the Golgi membrane and autopolysialylation is reported mainly on N-glycans *in vivo* and *in vitro* (Table 10) (Mühlenhoff *et al.*, 1996 ; Close and Colley, 1998 ; Close *et al.*, 2001), however we suggested here that autopolysialylation could also occur on O-glycans of *Cma* ST8Sia IV *in vitro* as shown for the production in HEK293S. Indeed, this version which is non-autopolysialylated on its N-glycans appear highly active in MPSA, proposing that autopolysialylation on N-glycans is not essential for *Cma* ST8Sia IV enzymatic activity.

To be noted, we also produced zebrafish ST8Sia III (Table 17) in HEK293 cells and investigated its potential autopolysialylation and influence on its enzymatic activity. We showed that this enzyme is active on MPSA using CMP-SiaNA1 and ALCAM as acceptors whereas, in contrast to polySTs, zebrafish ST8Sia III is not able to be autopolysialylated with CMP-Neu5Ac, consistent with the function of oligoST reported for the human ST8Sia III (Angata *et al.*, 2000).

To summarize, we will determine if autopolysialylation of *Cma* polySTs could be an advantage or a limit for the synthesis of polySia onto an exclusive acceptor substrate of biotechnological interest, like ALCAM and DNase I. We will notably explore if *Cma* polySTs could use the various donor substrates for autopolysialylation. Likewise, it might be interesting to autopolysialylate enzymes prior sialylation reactions and compare its activity with non-autopolysialylated enzymes.

VII. MODELING AND COOPERATION OF POLYSTS

Our previous study based on 3D-modeling of the PSTD motif and MSA unveiled specific aa substitutions, involved in donor and acceptor substrates recognitions, and potential changes of function in the salmonids polyST sequences (Venuto *et al.*, 2020a). During my thesis, the rise of AlphaFold allowed us to obtain 3D structural data and a clear view of the organization of the polySTs which could explain their distinct specificities. Although structural differences could be noticed between homolog polySTs, our first results confirm that specific substitutions located in SDPs, SMs, PBR and PSTD motifs could be associated with disordered loops, could influence the flexibility of the catalytic site, the potential dimeric interactions and could explain the specificities of the *Cma* polySTs. To be noted, since we used truncated recombinant polySTs in enzymatic assays, we might lose part of structural information of these disordered regions for further interaction between enzymes. Thus, it could be interesting to use full length of *Cma* polySTs and co-transfected them in HEK293 cells to evaluate their dimeric interaction in the Golgi membrane.

On the other hand, next strategies might be focused on the development of mutants of different interests. The design of catalytically inactive H330K mutant of *Cma* ST8Sia IV will help to investigate potential cooperation with *Cma* ST8Sia II-r1 or *Cma* ST8Sia II-r2 which could explain their enzymatic function. Likewise, *Cma* ST8Sia IV mutants of N-glycosylation could be designed

to investigate the role of N-glycosylation and autopolysialylation for potential interaction as it might explain that only monomers of *Hsa* ST8Sia IV are detected in solution (Angata *et al.*, 2001).

Proposing SDPs responsible for *Cma* polySTs specificities, the development of mutants could be investigated to mimic and modify enzymatic specificities of human polySTs. In parallel, the molecular docking of polySTs active site with various CMP-Sia should be explored to complete this study and to validate the specificities of *Cma* polySTs by computational approaches.

Finally, the use of computational methods like AlphaFold2 will help to predict protein-protein interactions, modelize protein complexes with reasonable confidence, gain access to their structural organization and will offer the possibility to modulate the binding affinity site, although it is still challenging when membrane or transmembrane domain are not taken into account in the docking procedure (Launay *et al.*, 2023).

VIII. MID AND LONG-TERM PERSPECTIVES

The molecular tools developed and the unique polySTs specificities in fish will be useful in the future for the biosynthesis of new bioinspired materials. The production of various reliable polySia will be important to provide a new generation of polySia-based coating material which will have with various affinities to interact and potentialize antimicrobial binding partners. Indeed, my thesis is part of an international collaborative project which aims to graft different polySia onto natural or unnatural acceptors to generate new bioconjugates with high therapeutic potential (Fig. 75). Polysialylated beads or acceptors will be used to interact with biological factors like antimicrobial peptides (AMPs) to potentialize its antimicrobial activity, particularly modulating sensitivity of binding partners and preventing from proteolytic cleavage. Likewise, this polysialylated system could also be applied for histone derived-peptides in NETosis process to target bacterial membrane according to their higher acidic PLs composition compared to mammalian membrane. Indeed, affinity between polySia and its antimicrobial interaction partners is high enough to form complexes which will drive the release of AMPs to attack and perforate the cell surface of microbes.

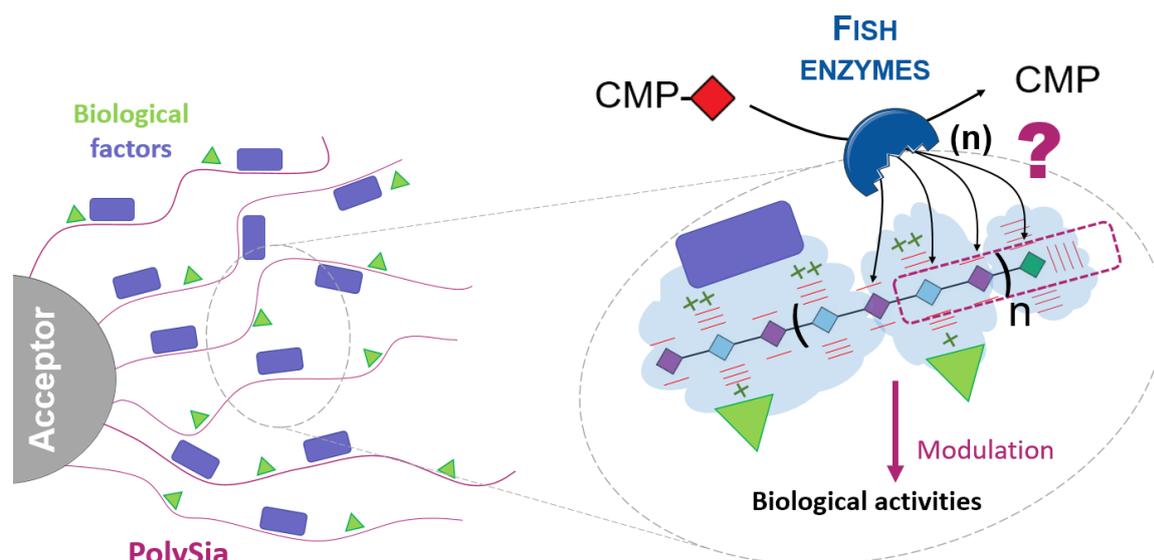


Figure 75: Contribution of this work for biotechnological applications. The enzymatic characterization of salmonid polySTs and highlights of their unique donor specificities is part of an international project in order to synthesize various polySia chains onto different acceptors (glycoproteins, beads) and modulate interactions with biological factors in immunological process.

Indeed, marine environment offers a tremendous biodiversity, a real potential for natural biopolymers product drug discovery and original polysaccharides (Gallo *et al.*, 2012) presenting a great chemical diversity with high interesting physico-chemical properties. The development of these polysialylated biomolecules combinations will represent affordable tools to biomedical investigators to develop the next generation of marine antibiotics that could be used in aquaculture preserving animals, plant, water and thus human health. In the long term, these natural tools have the potential to replace pesticides and chemical products used against pathogens, improving quality of marine products, animal welfare and having a huge impact in the socio-economic field.

Otherwise, the use of these recombinant enzymes offers very interesting prospects for the specific glycosylation of therapeutic proteins such as IgG IV for example, whose activity is known to be modulated by α 2-6-sialylation (von Gunten *et al.*, 2014) or for remodeling the surface of cells and modifying their destiny in the body (Sackstein *et al.*, 2012). Moreover, polysialylated acceptors could be much more important than we think, as suggested recently after large approaches of polySia glycoconjugates purification and analysis from tissues (Boyancé thesis, 2021). These observations offer good perspectives to use polySTs specificities as natural bioenzymes to remodel cell surface glycosylation and influence cells interaction in *in vivo* context.

IX. CONCLUSION

In summary, my thesis was the opportunity to implement a whole set of approaches ranging from the production of recombinant proteins, the optimization of enzymatic assays in order to visualize polysialylation synthesized *in vitro*, detect it by analytical technics and on cell surfaces using specific tools. These data enabled us to highlight unique enzymatic activities and specificities for the salmonid polySTs acquired during vertebrate evolution with functional divergence, distinct from those described in humans. In contrast to human polySTs, salmonid polySTs have kinetic properties and specific characteristics to recognize and catalyze the transfer of various donor substrates in order to synthesize diverse polySia chains in *in vitro* assays and on cell surface. Altogether, these approaches provide a conceptual framework to understand sialylation machinery evolution and explain why closely related species differ in their sialome. Understanding how these enzymes evolved will help addressing challenges like engineering of STs with unique specificities and will contribute to the synthesis of new bioinspired materials in the future.

To conclude, this work has been valued during many **oral presentations** and **posters**, has result in many **awards** and was accepted on September 2023 for a first author publication in Nature Scientific Reports entitled **Salmonid polysialyltransferases to generate a variety of sialic acid polymers** as summarized in this model below (Fig. 76).

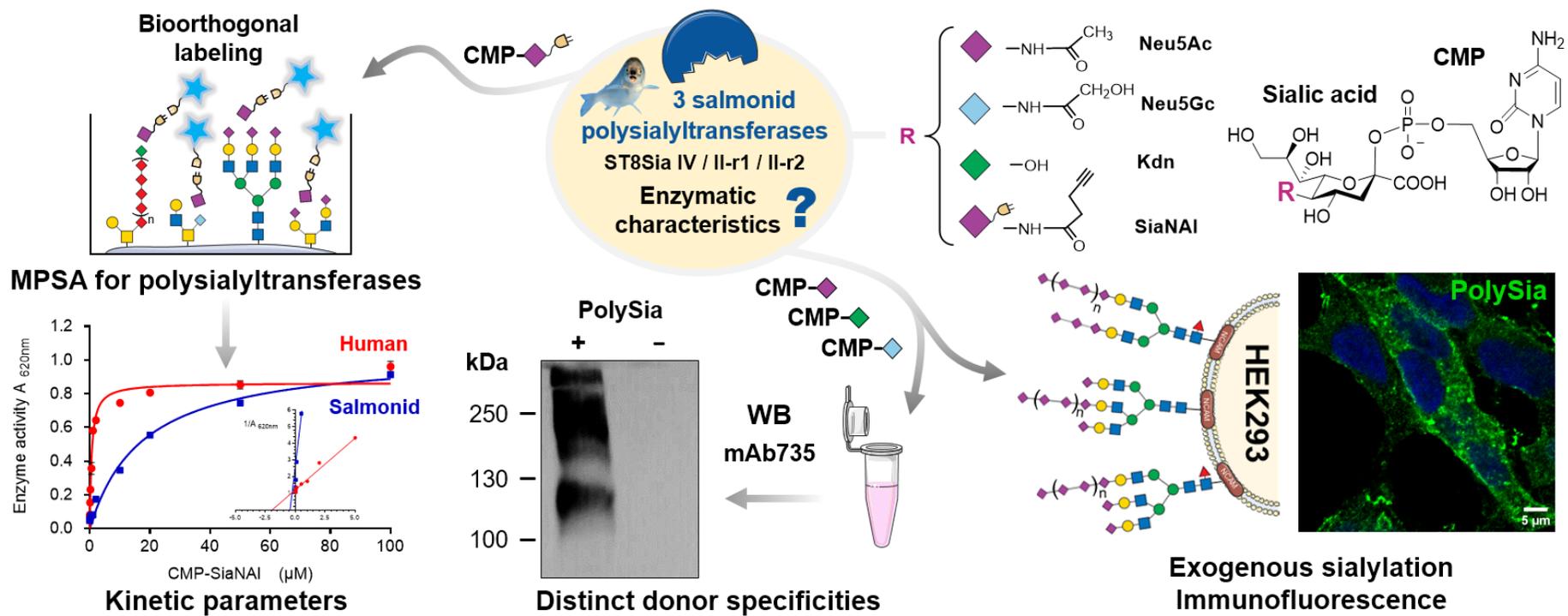


Figure 76: Model that resume main project experiments based on salmonid polySTs enzymatic specificities to explain the polySia biodiversity occurring in salmonid. Work accepted for publication in Nature Scientific Reports.

CONTRIBUTIVE AND COLLABORATIVE RESEARCH PROJECTS

My thesis was also the opportunity to study and compare other vertebrate ST activities, specificities and characteristics depending on their molecular evolution. As previously investigated during my master thesis, I obtained new interesting results for the zebrafish monosialyltransferase ST6Gal I, never studied before, in comparison to its mostly studied human homolog ST6Gal I. Notably, I highlighted distinct optimum conditions and kinetic parameters of the two ST6Gal I using MPSA and exogenous sialylation on cell surface and flow cytometry analyses. These data have been valued with a **poster presentation** in international congress and will contribute to confirm neofunctionalization of ST6Gal I which appears during Vertebrate evolution (Fig. 18).

Moreover, I continued to investigate the enzymatic activity of the zebrafish ST8Sia VIII, which undergone pseudogenization during Vertebrate evolution (Fig. 18) and have potential distinct enzymatic activity compared to its human paralog ST8Sia VI. Indeed, zebrafish ST8Sia VIII seems not active on MPSA as I showed on our **scientific publication** (Chang *et al.*, 2019) whereas this enzyme is effectively active for exogenous sialylation on HEK293 cells, although its preferential acceptor has to be identified by glycomic analyses. This work has notably been valued during an **oral presentation** in international congress.

In parallel, I explored different aspects in glycosylation and GTs activity and contributed at different levels with many collaborators. I was notably involved in a **team project** to study B4GALNT2 glycosyltransferase and influence of its N-glycosylation on its function (Cogez *et al.*, 2023) and in an **intern collaboration** with the “Regulation of early development” team of our lab to write a review focused on diversity of sialylation involved in fertilization processes (Fliniaux *et al.*, 2022). Finally, a new **international collaboration** with Dr B. Richichi (UNIFI, IT) allowed me to initiate studies of *in vitro* selectivity of potential inhibitor compound of STs (Fig. 31) and I obtained interesting preliminary results on MPSA using human monoSTs. These data will be useful to precise selectivity of inhibitor and for drug designs in the future, as our collaborator have recently done for specific fucosyltransferase inhibitors (Martin *et al.*, 2021).

MATERIALS AND METHODS

I. MATERIALS AND ANIMALS

Fetuin, orosomucoid, bovine submandibular mucin (BSM), TMB (3, 3', 5, 5'-tetramethylbenzidine), azide-PEG3-biotin, monoclonal Anti-FLAG® BioM2 antibody (1 mg/mL), sialic acid aldolase from *Escherichia coli* K-12, CMP-sialic acid synthetase from *Neisseria meningitidis* (CSS), pyrophosphatase (PPase), N-acetylneuraminic acid (Neu5Ac), N-glycolylneuraminic acid (Neu5Gc) and 2-keto-3-deoxy-nonulosonic acid (Kdn), APTS (8-aminopyrene-1,3,6-trisulfonic acid trisodium salt), 1 M sodium cyanoborohydride in THF were purchased from Sigma Aldrich. Rainbow trout CSS (rtCSS) used for CMP-Kdn synthesis was obtained from a previous collaborative published work (Wu *et al.*, 2022). Sodium cytidine-5'-triphosphate (CTP) was obtained from TCI chemicals. Horseradish peroxidase (HRP) conjugated anti-biotin antibody (monoclonal fraction of mouse IgG; 0.8 mg/mL) was purchased from Jackson ImmunoResearch. 2-[4-({bis[(1-tert-butyl-1H-1,2,3-triazol-4-yl)methyl]amino}methyl)-1H-1,2,3-triazol-1-yl]acetic acid (BTAA) was synthesized as described previously (Lopez Aguilar *et al.*, 2017). The human recombinant enzymes $\Delta 23ST8Sia$ II and $\Delta 39ST8Sia$ IV were obtained from R&D Systems (Rennes, France). PSGP-L and PSGP-H was collected from a previous collaborative published work in our laboratory (Chang *et al.*, 2009). Deoxyribonuclease I Human Recombinant (DNase I) was purchased from Boster Bio (Pleasanton, CA, USA). PageRuler Plus Prestained Protein Ladder was from Thermo Scientific. Polysialic acid monoclonal antibody (mAb735, Rabbit IgG κ , 1 mg/mL) was obtained from Enzo Life Sciences. Sodium cacodylate buffer was from Prolabo. Endoneuraminidase-N (endoN) was a gift from Dr Martina Mühlenhoff (MHH, DEU). Activated-Leukocyte Cellular Adherence Molecule (ALCAM) 3xFLAG tagged recombinant plasmid was a gift from Joachim Bentrop (KIT, DEU). Whitefish (*Coregonus maraena*) were provided by the Institute of Fisheries of the Mecklenburg Western Pomerania Research Center for Agriculture and Fisheries (Born, Germany). Animals were handled and sacrificed following the standard guidelines described previously (Venuto *et al.*, 2020a). Brain and kidney tissues from whitefish were snap frozen in liquid nitrogen and kept at -80°C for RNA isolation. The biotinylated

lectins SNA (*Sambucus nigra*), MAL I, MAL II (*Maackia amurensis*) and RCA I (*Ricinus communis*) come from Vector Laboratories. CHSE-214 (CVCL_2780) salmonid embryonic cells were cultivated within the team of Dr A. Rebl in the fish genetics unit of FBN institute (DEU).

II. BIOINFORMATICS ANALYSES

A. *IN SILICO IDENTIFICATION OF POLYSIALYLTRANSFERASES SEQUENCES*

The three genes *st8Sia2 r-1*, *st8sia2 r-2* and *st8sia4* identified previously (Venuto *et al.*, 2020b) in *C. maraena* transcriptome and deposited at the NCBI Sequence Read Archive (Bioproject ID: PRJNA302355) were compared to both human *ST8SIA2* (HGNC: 10870) and *ST8SIA4* (HGNC: 10871) genes through multiple sequence alignments using ClustalW at PRABI-Gerland (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_clustalw.html). The amino acid sequence analysis was performed using the software of Expert Protein Analysis System (ExPASy; Swiss Institute of Bioinformatics, Switzerland; website (<https://www.expasy.org/> accessed on: 04 November 2022)). Hydropathy analyses and determination of potential *N*-glycosylation sites were performed using the servers TMHMM - 2.0 : Prediction of Transmembrane Regions (<https://services.healthtech.dtu.dk/service.php?TMHMM-2.0> accessed on: 04 November 2022) and the NetNGlyc 1.0 Internet program (<https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0> , accessed on: 04 November 2022) of ExPASy. Sequences were also submitted to Compute pI/MW (ExPASy) analysis to determine its theoretical molecular weight. Multiple sequence alignments (MSA) were performed using clustalW at the PRABI website (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_clustalw.html, accessed on: 04 November 2022).

B. *MODELING USING ALPHAFOLD 2*

The three coregone polySTs and the two human polySTs aa sequences (Appendix 3) were submitted to AlphaFold database to generate PDB models. Our locally installed version of the

AlphaFold v2.2.0 software allows building models of monomers as well as multimers. The options used were following for monomer and multimer :

AlphaFold-monomer	AlphaFold-multimer
name='X1' (X1 corresponding to aa sequence of monomer submitted)	name='multimerX1vsX2' (X1 and X2 corresponding to each aa sequence of monomer submitted)
userpath="/home/\$USER/alphafold_runs"	userpath="/home/\$USER/alphafold_runs"
model_preset='monomer_ptm'	model_preset='multimer'
max_template_date='2022-03-16'	max_template_date='2022-03-16'
db_preset='full_dbs'	db_preset='full_dbs'
run_relax='true'	run_relax='true'
enable_gpu_relax='true'	enable_gpu_relax='true'
use_precomputed_msas='false'	num_multimer_predictions_per_model='3' (5 predictions per monomer : 15 models generated)
	use_precomputed_msas='false'

All submissions and predictions are available on our internal AlphaFold server for bigdata storage.

Five and fifteen models were generated for each monomer and multimer respectively, resulting from MSA with other protein sequences in the AlphaFold database sharing homology or conserved regions for modeling. For monomer, each prediction for the 5 polySTs studied have iptm+ptm scores between 80.4 and 85.8, which correspond to very good prediction on a [0 ;100] scale (Guo *et al.*, 2022). All top models exhibited very similar PAE matrices and pLDDT scores with a high confidence in the catalytic domain whereas the stem region is less reliable, probably with more disorder or flexibility. Secondary structures represented on AlphaFold monomer are in accordance with ENDScript analyses of the 5 polySTs for sequences organization (Fig. 69). For multimer, iptm+ptm scores are based on a [0.00 ;1.00] scale (Evans *et al.*, 2022) developed for homodimeric and heterodimeric protein complexes and interaction (Bryant *et al.*, 2022). Interactions between each coregone polySTs and each human polySTs were submitted to AlphaFold-multimer and showed more heterogenous results. Only one predicted model is of high confidence for homodimer, with an iptm+ptm top score of 0.827 for *Hsa* ST8Sia IV vs *Hsa* ST8Sia IV. Heterodimers have more good iptm+ptm scores with *Hsa* ST8Sia II vs *Hsa* ST8Sia IV, *Cma* ST8Sia II r-1 vs *Cma* ST8Sia IV and *Cma* ST8Sia II r-2 vs *Cma* ST8Sia IV, with top scores of 0.659, 0.746 and 0.744 respectively.

C. MOLECULAR DOCKING FOR POLYSIA TOOLS

To investigate to what extent analogous interactions could be found between the antibody and polySia chains made of eight Neu5Gc or Kdn, we took as reference direct and indirect interactions described for mAb735 complexed with polyNeu5Ac in ProteinDataBank (PDB) 3WBD and modified the Neu5Ac to NeuGc/Kdn molecules in the PDB 3WBD using the biocomputing software UCSF Chimera. The complex was minimized using the AMBER force field, and charges were computed using ANTECHAMBER.

III. MOLECULAR BIOLOGY

A. RNA EXTRACTION, cDNA SYNTHESIS AND PCR AMPLIFICATION OF POLYSTS

Total RNA was extracted from homogenized brain or kidney tissue using the Nucleospin RNA Plus (Macherey-Nagel, Düren, Germany) and quantified with the NanoDrop ®ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). Total RNA was reverse-transcribed using the Maxima First strand cDNA synthesis for RT-qPCR kit according to the manufacturer protocol (Thermo Scientific). To obtain a full-length cDNA, an initial RT-PCR was performed with each kidney and brain cDNA preparation, Q5 high fidelity DNA polymerase (New England Biolabs), sense and antisense oligonucleotide primers containing *Hind*III and *Kpn*I restriction sites respectively. PCR reactions were run for 2 min at 98°C followed by 35 cycles (98°C for 30 sec, 65°C for 30 sec, 72°C for 15 sec) and an extension step of 2 min at 72°C.

The following primers were used for *C. maraena* genes: for *cma st8sia4* 5'-CCCAAGCTTATGCTCTCACGGAAACGC-3' (sense) and 5'-CGGGGTACCTTAAGATTCGCACTTCGAAGTCG-3' (antisense), for *cma st8sia2 r-1* 5'-CCCAAGCTTATGCAGTTAGAATTCCGAACGCTG-3' (sense) and 5'-CGGGGTACCTTACATTCCTGCATCACAAGAGCC-3' (antisense) and for *cma st8sia2 r-2* 5'-CCCAAGCTTATGCAGTTAGAATTCCGAACATTG-3' (sense) and 5'-CGGGGTACCTTACGTTCCCTGGATCACAAGAGCC-3' (antisense). The resulting 1074 bp amplified *st8sia4* cDNA was subcloned into the PJET vector of Clone JET PCR cloning 1.2 kit

(Thermo Scientific), cut out with *Hind*III and *Kpn*I and inserted into the *Hind*III and *Kpn*I sites of the plasmid p3×FLAG-CMV10 expression vector. The resulting 1075 bp amplified *st8sia2 r-1* cDNA was inserted directly into p3×FLAG-CMV10. For *st8sia2 r-2*, no cDNA could be amplified. Consequently, a synthetic vector based on its full-length sequence was constructed using the pCDNA3.1 plasmid at GeneArt Instant Designer (ID project: 2020AAEFFD, Thermo Fisher Scientific). A cDNA encoding a truncated form of the *C. maraena (cma) st8sia2 r-1*, *st8sia2 r-2* and *st8sia4* lacking their first 28, 35 or 50 aa residues of the open reading frame were amplified by PCR using the previous construct p3×FLAG-CMV10 for *st8sia4* and *st8sia2 r-1* and pCDNA3.1 for *st8sia2 r-2* as templates DNA with the sense and antisense primers described in Appendix 3.

Touchdown reactions were run for 2 min at 94°C followed by 10 cycles (94°C for 30 sec, 60°C (minus 1°C per cycle) for 30 sec, 72°C for 70 sec) and then 20 cycles (94°C for 30 sec, 55°C for 30 sec, 72°C for 70 sec) and an extension step of 5 min at 72°C. The resulting PCR amplified cDNAs were inserted into the *Hind*III and *Kpn*I sites of the p3×FLAG-CMV9 expression vector. These last constructs encoded a FLAG-tagged protein with a signal peptide sequence, the FLAG octapeptide and either the three Δ 28ST8Sia II-r1, Δ 35ST8Sia II-r1 and Δ 50ST8Sia II-r1, either the three Δ 28ST8Sia II-r2, Δ 35ST8Sia II-r2 and Δ 50ST8Sia II-r2 or Δ 28ST8Sia IV sequence deleted of their cytoplasmic tail and transmembrane domain.

B. PRODUCTION AND PURIFICATION OF PLASMIDS

Plasmid constructions were carried out team in the vector p3×FLAG-CMV9 (Sigma Aldrich) to allow the expression of a truncated and labeled form of the enzyme in the culture medium of transfected cells. *E. coli* TOP10 bacteria transformed with these various plasmids and stored in glycerol at -20° C were cultured overnight at 37° C in an LB/Ampicillin medium (50 µg/mL). After centrifugation, the bacterial pellets were lysed and the plasmid DNA of the bacteria was extracted with the MidiPrep plasmid DNA kit (ThermoScientific). The concentration and purity of the plasmids were measured on the NanoDrop 2000 spectrophotometer (ThermoScientific).

IV. CELLULAR BIOLOGY

C. CELL CULTURE

HEK293 (ATCC CRL-1573), CHO-K1 (ATCC CCL-61) and CHO-Lec2 cells were grown in DMEM medium (DMEM, Dutscher) supplemented with fetal calf serum (10% FCS, Biowest) at 37°C, 5% CO₂. Confluent cells (~70%) were transiently transfected with either 2 or 10 µg of purified expression constructs in 6 well plates or in 100 mm petri dishes with Lipofectamine 2000 (Thermo Scientific) in UltraMEM medium (Lonza) according to the manufacturer's instructions. An empty p3×FLAG-CMV9 plasmid was used as a control (Mock). Transiently transfected cells and cell culture media were collected 24 h to 72 h after transfection. The cells are lysed in 100 µL of RIPA buffer (10 mM Tris HCl; 150 mM NaCl; Triton X-100 1%, pH 6.4), 2 µL of 50X protease inhibitors by 10 aspirations and discharges with a syringe and a needle 26Gx½ (Terumo). Recombinant enzymes produced in DMEM-FCS culture media were used without further purification as a crude enzyme source for enzymatic assays. The p3×FLAG-ALCAM plasmid was transiently transfected in CHO-K1 cells and UltraMEM medium was collected 48 h after transfection.

D. EXOGENOUS SIALYLATION ON CELLS: IF AND FACS

HEK293 cells were grown in DMEM containing 4.5 g/L of glucose and supplemented with 10% of FCS (Biowest) under humid atmosphere, at 37°C and 5% CO₂. HEK cells were grown on glass coverslips during 24 h, then washed three times with PBS and fixed in 4% paraformaldehyde (PFA) for 30 min at room temperature (RT). After three washes with PBS, cells were treated with endoN at 100 ng/mL for 1h at 37°C. Exceed of EndoN was eliminated by five washes of 5 min with PBS under gently agitation. Sialylation reactions were performed at 27°C for 4 h with the chosen enzymatic source and 500 µM of CMP-Sia donor substrate (CMP-Neu5Ac, CMP-Neu5Gc and CMP-Kdn) in 100 µL final volume. The rest of the experiment consists of immunofluorescence (IF) to visualise *de novo* sialylation. Briefly, after three washes with PBS, the cells were incubated with blocking buffer (2% BSA in PBS) for 1 h at RT and then incubated overnight at 4°C in humid chamber with the primary antibodies diluted in blocking buffer: 2 µg/mL anti-polySia mAb735 (Enzo Lifes Sciences) and 5 µg/mL anti-CD56 (NCAM, clone 123C3) (Invitrogen, ThermoFisher

Scientific). After three washes with PBS, cells were incubated 1 h at RT in the dark with AlexaFluor™ conjugated secondary antibodies: respectively goat anti-rabbit 488 and goat anti-mouse 568 (Invitrogen, ThermoFisher Scientific) diluted at 1: 600 (3.33 µg/mL) in blocking buffer. Finally, cells were washed three times with PBS, the nuclei were stained with DAPI (5 µg/mL in PBS) and the coverslips were mounted on glass slides with Mowiol®. The fluorescence was detected using the inverted Zeiss LSM780 confocal microscope with a 40 × oil immersion and data were collected with the ZEN 2010 software (Zeiss, Oberkochen, Germany). The images were analyzed with ImageJ software.

IF analyses with CMP-SiaNAI were done using HEK293, CHO-K1 and CHO-Lec2 cells. Briefly, 50 000 cells were grown in DMEM/SVF medium (10% FCS) in 24 well plates under humid atmosphere, at 37°C, 5% CO₂ and placed on glass coverslips the day before for IF. Prior sialylation reaction, two mM of CMP-SiaNAI and azido-biotin were coupled for 1 h at 37°C using the click-chemistry solution described above. Then, cells were incubated with 100 µM of CMP-SiaNAI-Azido-biotin and 100 µL of enzymatic source for 1 h at 37° C. After exogenous sialylation reaction, cells were analyzed either by flow cytometry (FACS) or by confocal microscopy.

For the FACS analyses, cells were washed on PBS, detached using PBS-EDTA (5 mM) and deposited in 96-well plates. After centrifugation (1200 rpm, 5 min), streptavidin coupled with fluorescein isothiocyanate (streptavidin-FITC) (DyLight 488, Vector-Laboratories) diluted in PBS-BSA 1% (1: 500) was incubated on cells 1 hour at 4°C in the dark. After three washes with PBS, fluorescence was quantified for 10,000 cells on BD Accuri C6 Flow Cytometer (Becton Dickinson). Gates were determined for each cell type in order to analyze the fluorescence on a homogeneous cell population using FlowJo software.

For the confocal microscopy analyses, cells were washed on PBS and fixed using PFA 4% for 20 min, RT. After three washes, cells were incubated with blocking buffer PBS-BSA 1% for 1 hour at RT. Then, cells were incubated for 1 h at RT with streptavidin-FITC diluted in PBS-BSA 1% (1:500). After three washes, DAPI diluted in PBS (1:200) was incubated on cells for 15 min at RT. Finally, after 3 washes with PBS, the coverslips were mounted on slides and analyzed as described above using the inverted Zeiss LSM780 confocal microscope.

V. BIOCHEMISTRY AND ANALYTIC APPROACHES

A. WESTERN BLOT

The presence of FLAG-tagged recombinant proteins was visualized by WB. 30 μ L of transfected cell culture media or 30 μ g of cell lysates were boiled for 5 min at 95°C in 4X Laemmli buffer (235 mM Tris-HCl pH 6.8, 8% SDS, 40% glycerol, 10% β -mercaptoethanol, 0.01% bromophenol blue), separated on an 8% SDS-PAGE gel for 1 hour at 95V (1X Tris-Glycine buffer (Euromedex)/ Methanol: 20%) then transferred to an Amersham Protran nitrocellulose membrane (GE Healthcare Life Sciences). Verification of the transfer is done by staining the membrane with Rouge-Ponceau. These were saturated in blocking buffer for 1 h at room temperature and then incubated with the anti-FLAG® antibody (1/1000; Sigma-Aldrich) overnight at 4°C. After 3 washes with 0.05% TBST, the membrane was incubated with an anti-mouse antibody coupled to horseradish peroxidase (HRP) (1/10000; GE Healthcare) for 1 hour and then washed 5 times with 0.05% TBST. Chemiluminescence revelation was performed using the ECL Pico substrate (ThermoScientific) with a chemiluminescence camera (Vilber Lourmat) on Fusion software.

WB was used to visualize and quantify tagged recombinant proteins produced by transiently transfected HEK293 or CHO-K1 cells. For that purpose, 20 μ L of culture media of transfected cells were boiled for 5 min at 95°C in 4 \times Laemmli buffer (235 mM Tris-HCl pH 6.8, 8% SDS, 40% glycerol, 10% β -mercaptoethanol, 0.01% bromophenol blue), and resolved by 8% SDS-PAGE. Proteins were transferred onto an Amersham Protran nitrocellulose membrane (GE Healthcare Life Sciences) for 75 min at 200 mA and checked with Ponceau red staining (5% acetic acid, 0.1% Ponceau). Membranes were washed then blocked using 5% non-fat milk in TBS-T (TBS with 0.05% Tween 20). Detection of recombinant proteins was achieved with the primary mouse anti-FLAG® M2 antibody (1 μ g/ml; Sigma-Aldrich) in TBS-T 0.05% overnight at 4°C. After 3 washes with TBS-T 0.05%, the membrane was incubated with an anti-mouse antibody coupled to horseradish peroxidase (HRP) (0.1 μ g/mL; InVitrogen) for 1 h then was washed 5 times with TBS-T 0.05%. Blots were developed using enhanced chemiluminescence (ECL West Pico Plus, Thermo

Scientific). The images were acquired using a CCD camera (Fusion Solo, Vilber Lourmat) the Fusion software. Densitometry analysis were done with ImageJ and GraphPad Prism 6.

After sialylation reaction and endoN treatment, polysialylated CD166/ALCAM samples were heated for 5 min at 60°C in 4 × Laemmli buffer, separated on an 6% SDS-PAGE gel for 90 min at 95 V (Tris-Glycine buffer 1X (Euromedex)/Methanol: 20%) then transferred either onto an Amersham Protran nitrocellulose membrane (GE Healthcare Life Sciences). Membranes were saturated in 5% non-fat milk in TBS-T 0.05% for 1 h at RT and then incubated with the anti-polysialic acid mAb735 antibody (1 µg/mL) or EndoN (6.7 µg/mL) overnight at 4°C. After 3 washes with TBS-T 0.05%, the membrane was incubated with secondary anti-rabbit antibody coupled to horseradish peroxidase (HRP) (0.1 µg/mL; InVitrogen) for 1 h then was washed 5 times with TBS-T 0.05%. Detection was achieved by chemiluminescence as reported above.

B. PLATE LECTIN RECOGNITION ASSAY

Different quantities of acceptors are taken up in 100 µL of sodium bicarbonate (20 mM, pH 9.6) and adsorbed in the bottom of 96-well plates (F8 MaxiSorp Loose Nunc-Immuno Module ThermoScientific) at 4°C overnight. After 3 washes in 150 µL of PBST-0.05% (Tween saline phosphate buffer), saturation is carried out for 1 h at RT using 100 µL of bovine serum albumin (BSA) oxidized at 0.05% in bicarbonate buffer . After washing, the wells are incubated with a biotinylated SNA, MAL I, MAL II or RCA I lectin at 20 µg/mL in PBS. After 1 hour of incubation at 37°C, washings are carried out with PBST-0.05% and the wells are incubated with 100 µL of anti-biotin antibody coupled to HRP diluted in PBST-0.05% (1/25000) 1 h at 37°C. After washing, 100 µl of TMB are added and incubated for 20 min at room temperature in the dark. Finally, the absorbance is measured with a spectrophotometer at 620 nm (SpectroStar Nano, BMG Labtech).

C. ENZYMATIC ACTIVITY TESTS

Enzymatic assays were performed using the recently described MicroPlate Sialyltransferase Assay (MPSA) (Noel *et al.*, 2018). In brief, 400 ng of glycoprotein acceptors in 100 µL of sodium bicarbonate buffer (20 mM, pH 9.6) were adsorbed into the bottom of 96-well plates (F8 MaxiSorp Loose Nunc-Immuno Module ThermoScientific) at 4°C overnight. After 3 washes with 150 µL of

phosphate buffered saline containing 0.05% Tween 20 (PBS-T 0.05%), saturation was carried out for 1 h at room temperature (RT) with 100 μ L of oxidized BSA at 0.05% diluted in bicarbonate buffer. The sialylation transfer reaction was carried out for one to several hours at 27°C with the chosen enzymatic source and 100 μ M of CMP-SiaNAI in 100 mM cacodylate buffer (MnCl₂ 10 mM, Triton CF-54 0.2%, pH 6.2) in a final volume of 100 μ L. After sialylation, the wells were washed with PBS-T 0.05%, then the copper-catalysed azide-alkyne cycloaddition (CuAAC) labelling reaction was performed by adding 100 μ L of a solution containing 300 μ M CuSO₄, 600 μ M BTAA, 2.5 mM sodium ascorbate and 250 μ M azide-PEG3-biotin in PBS (Kolb *et al.*, 2001 ; Rostovtsev *et al.*, 2002 ; Tornøe *et al.*, 2002). After 1 h incubation at 37°C, the reaction was stopped by washing three times with PBS-T 0.05%, then 100 μ L of an HRP-conjugated anti-biotin antibody (32 ng/mL) was added for 1 h at 37°C. After washing, 100 μ L of TMB were added and incubated for 20 min at RT in the dark. Finally, the absorbance is measured at 620 nm with a spectrophotometer (SpectroStar Nano, BMG Labtech). The data were analyzed on GraphPad Prism 6 using the statistical ANOVA test between samples.

ST assays were also performed at 27°C for 4 h in 100 mM cacodylate buffer with 100 μ M CMP-Sia (*i.e.* CMP-Neu5Ac, CMP-Neu5Gc or CMP-Kdn), 4 μ g of the ALCAM acceptor substrate and 230 μ L of enzymatic source in a total volume of 1 mL. The enzymatic source with empty vector p3 \times FLAG-CMV9 (Mock) and recombinant human Δ 39ST8Sia IV or Δ 23ST8Sia II were used as controls. Samples were cooled in ice, dialyzed overnight on 10 kDa dialysis membrane in ammonium bicarbonate buffer 50 mM and lyophilized. Samples were resuspended in 18 μ L of RIPA buffer (Tris HCl 10 mM; NaCl 150 mM; Triton X-100 1%; pH 6.4) and treated or not with EndoN (6.7 μ g/mL) 1 h, 37°C. ALCAM polysialylation was visualized by WB. Total samples were boiled for 5 min at 65°C, resolved on a 6% SDS-PAGE for WB as previously described.

D. ANALYTICAL APPROACHES TO STUDY POLYMERS OF SIALIC ACIDS

i. Chain length analysis by HPLC

Fish ST8Sia IV was used to polysialylate CD166/ALCAM. The polysialylated fraction was isolated using affinity precipitation. Therefore, inactive endoN was coupled to tosyl-activated

magnetic DynabeadsM-280 (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions (Zlatina *et al.*, 2018 ; Kuhnle *et al.*, 2019). Inactive endoN binds to polySia, but is not able to degrade the polymer. The samples were dialyzed against TBS for 2 hours using Spectra/Por® Biotech CE Tubing (MWCO: 50 kDa; Repligen, Rancho Dominguez, CA). Subsequently, the dialyzed samples were incubated with the beads for 30 min. After washing, polySia was eluted using 100 mM triethylamine, 150 mM NaCl and dried in a vacuum concentrator. To control the affinity precipitation, 10 % of the eluate was analyzed by WB against polySia. 90 % of the samples were used for HPLC analysis.

The polySia chain length determination of polysialylated CD166/ALCAM was performed using the 1,2-diamino-4,5-methylenedioxybenzene (DMB)-HPLC strategy (Inoue and Inoue, 2003 ; Kuhnle *et al.*, 2019 ; Sato *et al.*, 1998). Briefly, the isolated polySia-CD166/ALCAM samples were dissolved in 80 μ L DMB reagent (9 mM sodium hydrosulfite, 0.5 M mercaptoethanol, 20 mM trifluoroacetic acid (TFA), 0.61 mg/mL DMB) and incubated at 11°C overnight. The reaction was stopped with 20 μ L of 1 M NaOH for 1 h. The fluorescently labeled polySia chains were separated by anion exchange chromatography using a DNAPac PA-100 column (4 \times 250 mm; 13 μ m; Dionex). The eluents, MilliQ water (E1) and 2 M ammonium acetate (E2) were used with a flow rate of 1 mL/min following the gradient: 0 min = 0 % E2, 5 min = 0 % E2, 15 min = 13 % E2, 30 min = 21 % E2, 55 min = 33 % E2, 100 min = 43 % E2, 101 min = 100 % E2, 110 min = 100 % E2, 111 min = 0 % E2, 145 min = 0 % E2. Fluorescent signals were detected using an extinction wavelength of 372 nm and an emission wavelength of 456 nm.

ii. Chain length analysis by FACE

APTS-FACE method was developed to analyze polySia. MonoSia and colominic acid standards were prepared and injected for capillary electrophoresis calibration. 100 μ g of Neu5Ac, Neu5Gc, Kdn, 0,5 μ mol of Sia mix or 100 μ g of colominic acid were derivatized overnight at 42°C with 2 μ L of APTS 200 mM in acetic acid 15% and 2 μ L of sodium cyanoborohydride 1M in THF.

About samples, 2 μ g of ALCAM were sialylated with either CMP-Neu5Ac, CMP-Neu5Gc or CMP-Kdn (100 μ M) and 20 μ L of *Cma* ST8Sia IV or 100 ng of *Hsa* ST8Sia IV. After 4 hours of

sialylation at 27°C, samples were dialyzed in NH₄CO₃, lyophilized, fixed on 3xFLAG beads overnight at 4°C then treated with PNGase F for 2 hours at 37°C. Finally samples were APTS-derivatized as for standards prior injection. To avoid lactonization, samples were treated in a sodium hydroxide buffer 2 hours at RT, then diluted 50 to 100 times in water before injections. 10 nL of samples were injected during 10 sec. Electrophoresis was performed on reverse polarity with a Beckman Coulter PA800 plus instrument (Sciex Separations, Les Ulis, France) equipped with a laser induced fluorescence detector. Separation was performed in a 50 µm I.D., 375 µm O.D. bare fused silica capillary of 60.2 cm in length (Sciex Separations, Les Ulis, France). Sample injection was carried out at 0.5 psi during 10 s and separation was achieved at 30 kV in carbohydrate separation gel buffer (Sciex Separations, Les Ulis, France), pure or diluted 1/3 in ultrapure water.

iii. Quantitative analysis of polySia composition by HPLC

To quantitatively analyze the transfer of different Sia onto ALCAM by *C. ma* ST8Sia IV, we established an endoN approach using Microcon® centrifugal filter units (MRCPRT010, Merck Millipore Ltd.; NMWL: 10 kDa) prepared according to the manufacturer's instructions. Polysialylation reaction was performed overnight at 37 °C in cacodylate buffer containing 100 µL of fish ST8Sia IV culture medium, 13.3 µL of CD166/ALCAM and 100 µM of CMP-Sia in a total volume of 250 µl. To remove remaining free CMP-Sia, samples were centrifuged with 14,000 × g at 4 °C until almost all the liquid was filtered. The samples were subsequently washed 3 times with 250 µl 50 mM NH₄HCO₃ followed by centrifugal-filtration step. Volume was adjusted to 110 µL with 50 mM NH₄HCO₃ buffer. EndoN was added (1.34 µg/mL) to degrade polySia. In parallel, untreated samples (without endoN) of each CMP-Sia preparation were used as negative controls. Aliquots (10 µL) of all samples were collected for WB after 1 h at 37 °C. The remaining samples were centrifuged as described above and washed three times. For quantitative Sia analysis, the flow-through (containing degraded sialic acid chains) was dried in a vacuum concentrator and dissolved in 0.2 N TFA. Hydrolysis was performed at 80°C for 4 h. Hydrolyzed samples were dried and dissolved in 80 µL DMB reagent. After 2 h at 55°C, reaction was stopped with 20 µL 0.2 N NaOH. Resulting fluorescently labeled Sia were injected into a HPLC system (Nexera, Shimadzu)

and separated using a Superspher® 100 RP-18 end-capped column (250 mm × 40 mm, Merck-Hitachi, Darmstadt, Germany). A gradient with the eluents 92 % MilliQ water, 4 % ACN, 4 % methanol, 0,1 % TFA (E1) and 10 % MilliQ water, 45 % acetonitrile (ACN), 45 % MeOH, 0,1 % TFA (E2) was applied with a flowrate of 0.25 ml/min as follows: 0 min = 0 % E2, 2 min = 0 % E2, 25 min = 2 % E2, 35 min = 5 % E2, 40 min = 50 % E2, 45 min = 100 % E2, 50 min = 100 % E2, 51 min = 0 % E2, 60 min = 0 % E2. Fluorescent signals were detected using an extinction wavelength of 372 nm and an emission wavelength of 456 nm. Sia standards were used to obtain a calibration line for quantification.

VI. CHEMICAL SYNTHESIS

A. CHEMO-ENZYMATIC SYNTHESIS OF CMP-ACTIVATED SIALIC ACIDS

SiaNA1 was synthesized as previously described (Gilormini *et al.*, 2016). In brief, *N*-mannosamine was converted to *N*-4-pentynoylmannosamine by coupling with succinimidyl 4-pentynoate, which was then reacted with sodium pyruvate in the presence of sialic acid aldolase from *E. coli* K12 in pH 7.5 phosphate buffer and purified by anion exchange chromatography (Dowex 1×8) and gel filtration chromatography (P2 resin) to yield pure SiaNA1 (86% overall yield over 2 steps). The syntheses of cytidine-5'-monophospho-*N*-sialic acids (CMP-Sia) CMP-Neu5Ac, CMP-Neu5Gc, CMP-Kdn and CMP-SiaNA1 were carried out in equimolarity of CTP and each Sia (1:1) with 0.3 U CSS and 0.5 U PPase in 100 mM Tris-HCl, 20 mM MgCl₂ buffer (pH 8.5) at 37°C. Reactions were carried out in a 5 mm NMR tube and monitored by ³¹P NMR spectroscopy in Brüker Avance II 400 MHz spectrometer. Upon completion, pure formed CMP-Sia was immediately stored at -80°C until use (Gilormini *et al.*, 2016a, Noel *et al.*, 2018).

B. PREPARATION OF THE DIFFERENT ACCEPTOR SUBSTRATES

Sialylated glycoproteins (fetuin, orosomuroid, bovine submandibular mucin) were desialylated with 0.1 M trifluoroacetic acid for 2 h at 80°C. Released Sia residues were removed using MWCO 3500 Spectra/Por3 or MWCO 1000 Spectra/Por6 dialysis membrane (18 mm width, Spectrum Laboratories) for 24 hours. The contents were then transferred to a glass tube and freeze-dried.

REFERENCES

- Aasheim HC, Aas-Eng DA, Deggerdal A, Blomhoff HK, Funderud S, Smeland EB. Cell-specific expression of human beta-galactoside alpha 2,6-sialyltransferase transcripts differing in the 5' untranslated region. *Eur J Biochem.* 1993;213(1):467-75.
- Abeln M, Albers I, Peters-Bernard U, Flächsig-Schulz K, Kats E, Kispert A, *et al.* Sialic acid is a critical fetal defense against maternal complement attack. *Journal of Clinical Investigation.* 2018;129(1):422-36.
- Abukar T, Rahmani S, Thompson NK, Antonescu CN, Wakarchuk WW. Development of BODIPY labelled sialic acids as sialyltransferase substrates for direct detection of terminal galactose on N- and O-linked glycans. *Carbohydrate Research.* 2021;500:108249.
- Al-Saraireh YMJ, Sutherland M, Springett BR, Freiburger F, Ribeiro Morais G, Loadman PM, *et al.* Pharmacological inhibition of polysialyltransferase ST8SiaII modulates tumour cell migration. *PLoS One.* 2013;8(8):e73366.
- Almaraz RT, Li Y. Labeling glycans on living cells by a chemoenzymatic glycoengineering approach. *Biology Open.* 2017;bio.021600.
- Almaraz RT, Tian Y, Bhattarcharya R, Tan E, Chen S-H, Dallas MR, *et al.* Metabolic Flux Increases Glycoprotein Sialylation: Implications for Cell Adhesion and Cancer Metastasis. *Mol Cell Proteomics.* 2012;11(7):M112.017558.
- Alshanski I, Shitrit A, Sukhran Y, Unverzagt C, Hurevich M, Yitzchaik S. Effect of Interfacial Properties on Impedimetric Biosensing of the Sialylation Process with a Biantennary N-Glycan-Based Monolayer. *Langmuir.* 2022;38(2):849-55.
- Alshanski I, Sukhran Y, Mervinetsky E, Unverzagt C, Yitzchaik S, Hurevich M. Electrochemical biosensing platform based on complex biantennary N-glycan for detecting enzymatic sialylation processes. *Biosensors and Bioelectronics.* 2021;172:112762.
- Angata T. Possible Influences of Endogenous and Exogenous Ligands on the Evolution of Human Siglecs. *Front Immunol.* 2018;9:2885.
- Angata K, Huckaby V, Ranscht B, Terskikh A, Marth JD, Fukuda M. Polysialic acid-directed migration and differentiation of neural precursors are essential for mouse brain development. *Mol Cell Biol.* oct 2007;27(19):6659-68.
- Angata K, Fukuda M. Polysialyltransferases: major players in polysialic acid synthesis on the neural cell adhesion molecule. *Biochimie.* 2003;85(1-2):195-206.
- Angata T, Kerr SC, Greaves DR, Varki NM, Crocker PR, Varki A. Cloning and characterization of human Siglec-11. A recently evolved signaling molecule that can interact with SHP-1 and SHP-2 and is expressed by tissue macrophages, including brain microglia. *J Biol Chem.* 2002b;277(27):24466-74.
- Angata K, Yen TY, El-Battari A, Macher BA, Fukuda M. Unique Disulfide Bond Structures Found in ST8Sia IV Polysialyltransferase Are Required for Its Activity. *Journal of Biological Chemistry.* 2001;276(18):15369-77.
- Angata K, Suzuki M, McAuliffe J, Ding Y, Hindsgaul O, Fukuda M. Differential biosynthesis of polysialic acid on neural cell adhesion molecule (NCAM) and oligosaccharide acceptors by three distinct alpha 2,8-sialyltransferases, ST8Sia IV (PST), ST8Sia II (STX), and ST8Sia III. *J Biol Chem.* 2000b;275(24):18594-601.
- Angata T, Varki A. Chemical Diversity in the Sialic Acids and Related α -Keto Acids: An Evolutionary Perspective. *Chem Rev.* 2002a;102(2):439 70.
- Angata T, Matsuda T, Kitajima K. Synthesis of neoglycoconjugates containing deaminated neuraminic acid (KDN) using rat liver alpha2,6-sialyltransferase. *Glycobiology.* mars 1998b;8(3):277-84.
- Angata K, Suzuki M, Fukuda M. Differential and Cooperative Polysialylation of the Neural Cell Adhesion Molecule by Two Polysialyltransferases, PST and STX. *Journal of Biological Chemistry.* 1998a;273(43):28524-32.
- Angata T, Kitazume S, Terada T, Kitajima K, Inoue S, Troy FA, *et al.* Identification, characterization, and developmental expression of a novel $\alpha 2 \rightarrow 8$ -KDN-transferase which terminates elongation of $\alpha 2 \rightarrow 8$ -linked oligopolysialic acid chain synthesis in trout egg polysialoglycoproteins. *Glycoconjugate J.* 1994;11(5):493-9.

- Antony P, Rose M, Heidenreich A, Knüchel R, Gaisa NT, Dahl E. Epigenetic inactivation of ST6GAL1 in human bladder cancer. *BMC Cancer*. 2014;14:901.
- Aoki T. Behaviour of a Sialo-Oligosaccharide from Glycophorin in Teleost Red Blood Cell Membranes. *IntechOpen*. 2022.
- Aoki K, Kumagai T, Ranzinger R, Bergmann C, Camus A, Tiemeyer M. Serum N-Glycome Diversity in Teleost and Chondrostean Fishes. *Front Mol Biosci*. 2021;8:778383.
- Aoki T, Chimura K, Sugiura H, Mizuno Y. Structure of a Sialo-Oligosaccharide from Glycophorin in Carp Red Blood Cell Membranes. *Membranes*. 2014;4(4):764-77.
- Arai M, Yamada K, Toyota T, Obata N, Haga S, Yoshida Y, Nakamura K, Minabe Y, Ujike H, Sora I, Ikeda K, Mori N, Yoshikawa T, Itokawa M. Association between polymorphisms in the promoter region of the sialyltransferase 8B (SIAT8B) gene and schizophrenia. *Biol Psychiatry*. 2006(59): 652–659.
- Asahina S, Sato C, Matsuno M, Matsuda T, Colley K, Kitajima K. Involvement of the α 2,8-Polysialyltransferases II/STX and IV/PST in the Biosynthesis of Polysialic Acid Chains on the O-Linked Glycoproteins in Rainbow Trout Ovary. *J. Biochem*. 2006;140(5):687-701.
- Audry M, Jeanneau C, Imberty A, Harduin-Lepers A, Delannoy P, Breton C. Current trends in the structure-activity relationships of sialyltransferases. *Glycobiology*. 2011;21(6):716-26.
- Azurmendi HF, Vionnet J, Wrightson L, Trinh LB, Shiloach J, Freedberg DI. Extracellular structure of polysialic acid explored by on cell solution NMR. *Proc Natl Acad Sci USA*. 2007;104(28):11557-61.
- Bader R, Wardwell PR. Polysialic acid: overcoming the hurdles of drug delivery. *Therapeutic Delivery*. 2014;5(3):235-7.
- Baenziger JU, Fiete D. Structure of the complex oligosaccharides of fetuin. *J Biol Chem*. 1979;254(3):789-95.
- Baos SC, Phillips DB, Wildling L, McMaster TJ, Berry M. Distribution of sialic acids on mucins and gels: a defense mechanism. *Biophys J*. 2012;102(1):176-84.
- Bardor M, Nguyen DH, Diaz S, Varki A. Mechanism of uptake and incorporation of the non-human sialic acid N-glycolylneuraminic acid into human cells. *J Biol Chem*. 2005;280(6):4228-37.
- Barry GT, Goebel WF. Colominic Acid, a Substance of Bacterial Origin related to Sialic Acid. *Nature*. 1957;179(4552):206-206.
- Bassens C, Rigaux P, Hendrick V, Cherlet M, Sato K, Kotarsky K, *et al*. Sodium Butyrate Enhancement of Protein Expression in Hela Cells: Follow-Up of a Chimeric Reporter Gene Behaviour for Process Development. *ESACT Proceedings*. 2005;2 :625-7.
- Basu S, Wallner B. DockQ: A Quality Measure for Protein-Protein Docking Models. Levy YK, editor. *PLoS ONE*. 2016;11(8):e0161879.
- Benkoulouche M, Fauré R, Remaud-Siméon M, Moulis C, André I. Harnessing glycoenzyme engineering for synthesis of bioactive oligosaccharides. *Interface Focus*. 2019;9(2):20180069.
- Bennett EP, Mandel U, Clausen H, Gerken TA, Fritz TA, Tabak LA. Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology*. 2012;22(6):736-56.
- Bentrop J, Marx M, Schattschneider S, Rivera-Milla E, Bastmeyer M. Molecular evolution and expression of zebrafish St8SiaIII, an alpha-2,8-sialyltransferase involved in myotome development. *Dev Dyn*. 2008;237(3):808-18.
- Berninsone P, Eckhardt M, Gerardy-Schahn R, Hirschberg CB. Functional Expression of the Murine Golgi CMP-Sialic Acid Transporter in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*. 1997;272(19):12616-9.
- Betenbaugh MJ, Tomiya N, Narang S, Hsu JT, Lee YC. Biosynthesis of human-type N-glycans in heterologous systems. *Current Opinion in Structural Biology*. 2004;14(5):601-6.
- Bhide GP, Zapater JL, Colley KJ. Autopolysialylation of polysialyltransferases is required for polysialylation and polysialic acid chain elongation on select glycoprotein substrates. *Journal of Biological Chemistry*. 2018;293(2):701-16.
- Bhide GP. Biophysical and Biochemical Determinants of Protein-Specific Polysialylation. Doctoral thesis. University of Illinois, Chicago (USA). 2017.

- Bhide GP, Prehna G, Ramirez BE, Colley KJ. The Polybasic Region of the Polysialyltransferase ST8Sia-IV Binds Directly to the Neural Cell Adhesion Molecule, NCAM. *Biochemistry*. 2017b;56(10):1504-17.
- Bhide GP, Colley KJ. Sialylation of N-glycans: mechanism, cellular compartmentalization and function. *Histochem Cell Biol*. 2017a;147(2):149-74.
- Bhide GP, Fernandes NRJ, Colley KJ. Sequence Requirements for Neuropilin-2 Recognition by ST8SiaIV and Polysialylation of Its O-Glycans. *Journal of Biological Chemistry*. 2016;291(18):9444-57.
- Blixt O, Allin K, Bohorov O, Liu X, Andersson-Sand H, Hoffmann J, *et al*. Glycan microarrays for screening sialyltransferase specificities. *Glycoconj J*. 2008;25(1):59-68.
- Bobowski M, Vincent A, Steenackers A, Colomb F, Van Seuning I, Julien S, *et al*. Estradiol represses the G(D3) synthase gene ST8SIA1 expression in human breast cancer cells by preventing NFκB binding to ST8SIA1 promoter. *PLoS One*. 2013;8(4):e62559.
- Boccuti L, Aoki K, Flanagan-Stee H, Chen CF, Fan X, Bartel F, *et al*. A mutation in a ganglioside biosynthetic enzyme, ST3GAL5, results in salt & pepper syndrome, a neurocutaneous disorder with altered glycolipid and glycoprotein glycosylation. *Human Molecular Genetics*. 2014;23(2):418-33.
- Bornhöft KF, Viergutz T, Kühnle A, Galuska SP. Nanoparticles Equipped with α2,8-Linked Sialic Acid Chains Inhibit the Release of Neutrophil Extracellular Traps. *Nanomaterials (Basel)*. 2019;9(4):610.
- Bowles WHD, Gloster TM. Sialidase and Sialyltransferase Inhibitors: Targeting Pathogenicity and Disease. *Front Mol Biosci*. 2021;8:705133.
- Boyancé A. Rôle de l'acide polysialique dans la réponse immunitaire innée : un focus sur les macrophages, les monocytes et les éosinophiles. University of Toulouse 3 (FR). 2021.
- Braasch I, Gehrke AR, Smith JJ, Kawasaki K, Manousaki T, Pasquier J, *et al*. The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. *Nat Genet*. 2016;48(4):427-37.
- Breton C, Šnajdrová L, Jeanneau C, Koča J, Imberty A. Structures and mechanisms of glycosyltransferases. *Glycobiology*. 2005;16(2):29R-37R.
- Bryant P, Pozzati G, Elofsson A. Improved prediction of protein-protein interactions using AlphaFold2. *Nat Commun*. 2022;13(1):1265.
- Bukowski RM, Tandler C, Cutler D, Rose E, Laughlin MM, Statkevich P. Treating cancer with PEG Intron: Pharmacokinetic profile and dosing guidelines for an improved interferon-alpha-2b formulation. *Cancer*. 2002;95(2):389-96.
- Büll C, Heise T, Adema GJ, Boltje TJ. Sialic Acid Mimetics to Target the Sialic Acid–Siglec Axis. *Trends in Biochemical Sciences*. 2016;41(6):519-31.
- Büll C, Boltje TJ, Wassink M, de Graaf AMA, van Delft FL, den Brok MH, *et al*. Targeting aberrant sialylation in cancer cells using a fluorinated sialic acid analog impairs adhesion, migration, and *in vivo* tumor growth. *Mol Cancer Ther*. 2013;12(10):1935-46.
- Bussink AP, van Swieten PF, Ghauharali K, Scheij S, van Eijk M, Wennekes T, *et al*. N-Azidoacetylmannosamine-mediated chemical tagging of gangliosides. *Journal of Lipid Research*. 2007;48(6):1417-21.
- Caliceti P, Veronese FM. Pharmacokinetic and biodistribution properties of poly(ethylene glycol)-protein conjugates. *Adv Drug Deliv Rev*. 2003;55(10):1261-77.
- Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res*. 2009;37:D233-238.
- Carey D, Sommers L, Hirschberg C. CMP-N-Acetylneuraminic acid: Isolation from and penetration into mouse liver microsomes. *Cell*. 1980;19(3):597-605.
- Cavdarli S, Schröter L, Albers M, Baumann AM, Vicogne D, Le Doussal JM, *et al*. Role of Sialyl-O-Acetyltransferase CASD1 on GD2 Ganglioside O-Acetylation in Breast Cancer Cells. *Cells*. 2021;10(6):1468.
- Cavdarli S, Delannoy P, Groux-Degroote S. O-acetylated Gangliosides as Targets for Cancer Immunotherapy. *Cells*. 2020;9(3):741.

- Cavdarli S, Dewald JH, Yamakawa N, Guérardel Y, Terme M, Le Doussal JM, *et al.* Identification of 9-O-acetyl-N-acetylneuraminic acid (Neu5,9Ac2) as main O-acetylated sialic acid species of GD2 in breast cancer cells. *Glycoconj J.* 2019b;36(1):79-90.
- Cavdarli S, Groux-Degroote S, Delannoy P. Gangliosides: The Double-Edge Sword of Neuro-Ectodermal Derived Tumors. *Biomolecules.* 2019a;9(8):311.
- Chan PHW, Lairson LL, Lee HJ, Wakarchuk WW, Strynadka NCJ, Withers SG, *et al.* NMR Spectroscopic Characterization of the Sialyltransferase CstII from *Campylobacter jejuni*: Histidine 188 Is the General Base. *Biochemistry.* 2009;48(47):11220-30.
- Chang A, Singh S, Phillips GN, Thorson JS. Glycosyltransferase structural biology and its role in the design of catalysts for glycosylation. *Current Opinion in Biotechnology.* 2011;22(6):800-8.
- Chang L-Y, Teppa E, Noel M, Gilormini P-A, Decloquement M, Lion C, *et al.* Novel Zebrafish Mono- α 2,8-sialyltransferase (ST8Sia VIII): An Evolutionary Perspective of α 2,8-Sialylation. *International Journal of Molecular Sciences.* 2019;20(3):622.
- Chang L-Y, Harduin-Lepers A, Kitajima K, Sato C, Huang C-J, Khoo K-H, *et al.* Developmental regulation of oligosialylation in zebrafish. *Glycoconj J.* 2009;26(3):247-61.
- Chang WW, Yu CY, Lin TW, Wang PH, Tsai YC. Soyasaponin I decreases the expression of α 2,3-linked sialic acid on the cell surface and suppresses the metastatic potential of B16F10 melanoma cells. *Biochem Biophys Res Commun.* 2006;341(2):614-9.
- Charter NW, Mahal LK, Koshland DE, Bertozzi CR. Differential Effects of Unnatural Sialic Acids on the Polysialylation of the Neural Cell Adhesion Molecule and Neuronal Behavior. *Journal of Biological Chemistry.* 2002;277(11):9255-61.
- Charter NW, Mahal LK, Koshland DE, Bertozzi CR. Biosynthetic incorporation of unnatural sialic acids into polysialic acid on neural cells. *Glycobiology.* 2000;10(10):1049-56.
- Chen SC, Huang CH, Lai SJ, Yang CS, Hsiao TH, Lin CH, *et al.* Mechanism and inhibition of human UDP-GlcNAc 2-epimerase, the key enzyme in sialic acid biosynthesis. *Sci Rep.* 2016;6:23274.
- Chen G, Zhang D, Fuchs TA, Manwani D, Wagner DD, Frenette PS. Heme-induced neutrophil extracellular traps contribute to the pathogenesis of sickle cell disease. *Blood.* 2014;123(24):3818-27.
- Chen JY, Tang YA, Huang SM, Juan HF, Wu LW, Sun YC, *et al.* A novel sialyltransferase inhibitor suppresses FAK/paxillin signaling and cancer angiogenesis and metastasis pathways. *Cancer Res.* 2011;71(2):473-83.
- Chen C. Generation and Characterization of Recombinantly Polysialylated Antibody. Doctoral thesis. University of Imperial College London (EN). 2011.
- Chen C, Colley KJ. Minimal structural and glycosylation requirements for ST6Gal I activity and trafficking. *Glycobiology.* 2000;10(5):531-83.
- Cheng MC, Lin CH, Lin HJ, Yu YP, Wu SH. Hydrolysis, lactonization, and identification of α 2 \rightarrow 8/ α 2 \rightarrow 9 alternatively linked tri-, tetra-, and polysialic acids. *Glycobiology.* 2004;14(2):147-55.
- Chiu CPC, Watts AG, Lairson LL, Gilbert M, Lim D, Wakarchuk WW, *et al.* Structural analysis of the sialyltransferase CstII from *Campylobacter jejuni* in complex with a substrate analog. *Nat Struct Mol Biol.* 2004;11(2):163-70.
- Cho JW, Troy FA. Polysialic acid engineering: synthesis of polysialylated neoglycosphingolipids by using the polysialyltransferase from neuroinvasive *Escherichia coli* K1. *Proc Natl Acad Sci USA.* 1994;91(24):11427-31.
- Chou HH, Takematsu H, Diaz S, Iber J, Nickerson E, Wright KL, *et al.* A mutation in human CMP-sialic acid hydroxylase occurred after the *Homo-Pan* divergence. *Proc Natl Acad Sci USA.* 1998;95(20):11751-6.
- Chuang W, Lagenaur CF. Central nervous system antigen P84 can serve as a substrate for neurite outgrowth. *Developmental Biology.* 1990;137(2):219-32.
- Close BE, Mendiratta SS, Geiger KM, Broom LJ, Ho LL, Colley KJ. The minimal structural domains required for neural cell adhesion molecule polysialylation by PST/ST8Sia IV and STX/ST8Sia II. *J Biol Chem.* 2003;278(33):30796-805.

- Close BE, Wilkinson JM, Bohrer TJ, Goodwin CP, Broom LJ, Colley KJ. The polysialyltransferase ST8Sia II/STX: posttranslational processing and role of autopolysialylation in the polysialylation of neural cell adhesion molecule. *Glycobiology*. 2001;11(11):997-1008.
- Close BE, Tao K, Colley KJ. Polysialyltransferase-1 autopolysialylation is not requisite for polysialylation of neural cell adhesion molecule. *J Biol Chem*. 2000;275(6):4484-91.
- Close BE, Colley KJ. *In Vivo* Autopolysialylation and Localization of the Polysialyltransferases PST and STX. *Journal of Biological Chemistry*. 1998;273(51):34586-93.
- Cogez V, Vicogne D, Schulz C, Portier L, Venturi G, de Ruyck J, Decloquement M, Lensink M, Brysbaert G, Dall'Olio F, Groux-Degroote S, Harduin-Lepers A. N-glycan on the non-consensus N-X-C glycosylation site impacts activity, stability and localization of the Sd^a synthase B4GALNT2. *IJMS*. 2023.
- Cohen M, Varki A. The Sialome—Far More Than the Sum of Its Parts. *OMICS: A Journal of Integrative Biology*. 2010;14(4):455-64.
- Colomb F, Giron LB, Trbojevic-Akmacic I, Lauc G, Abdel-Mohsen M. Breaking the Glyco-Code of HIV Persistence and Immunopathogenesis. *Curr HIV/AIDS Rep*. 2019;16(2):151-68.
- Colley KJ, Kitajima K, Sato C. Polysialic acid: Biosynthesis, novel functions and applications. *Critical Reviews in Biochemistry and Molecular Biology*. 2014;49(6):498-532.
- Constantinou A, Chen C, Deonarain MP. Modulating the pharmacokinetics of therapeutic antibodies. *Biotechnol Lett*. 2010;32(5):609-22.
- Corfield T. Bacterial sialidases--roles in pathogenicity and nutrition. *Glycobiology*. 1992;2(6):509-21.
- Couceiro JNSS, Paulson JC, Baum LG. *Influenza* virus strains selectively recognize sialyloligosaccharides on human respiratory epithelium; the role of the host cell in selection of hemagglutinin receptor specificity. *Virus Research*. 1993;29(2):155-65.
- Coutinho PM, Henrissat B. Carbohydrate-active enzymes: an integrated database approach. In: Gilbert HJ, Davies G, Henrissat H, Svensson B, editors. *Recent Advances in Carbohydrate Bioengineering*. Cambridge: The Royal Society of Chemistry; 1999. pp. 3–12.
- Cox EC, Thornlow DN, Jones MA, Fuller JL, Merritt JH, Paszek MJ, *et al*. Antibody-Mediated Endocytosis of Polysialic Acid Enables Intracellular Delivery and Cytotoxicity of a Glycan-Directed Antibody-Drug Conjugate. *Cancer Res*. 2019;79(8):1810-21.
- Cui T, Man Y, Wang F, Bi S, Lin L, Xie R. Glycoenzyme Tool Development: Principles, Screening Methods, and Recent Advances. *Chin J Chem*. 2022;40(6):746-58.
- Curreli S, Arany Z, Gerardy-Schahn R, Mann D, Stamatou NM. Polysialylated Neuropilin-2 Is Expressed on the Surface of Human Dendritic Cells and Modulates Dendritic Cell-T Lymphocyte Interactions. *Journal of Biological Chemistry*. 2007;282(42):30346-56.
- Dall'Olio F, Malagolini N, Chiricolo M, Trinchera M, Harduin-Lepers A. The expanding roles of the Sda/Cad carbohydrate antigen and its cognate glycosyltransferase B4GALNT2. *Biochimica et Biophysica Acta (BBA) - General Subjects*. 2014;1840(1):443-53.
- Dall'Olio F, Chiricolo M, D'Errico A, Gruppioni E, Altimari A, Fiorentino M, *et al*. Expression of beta-galactoside alpha2,6 sialyltransferase and of alpha2,6-sialylated glycoconjugates in normal human liver, hepatocarcinoma, and cirrhosis. *Glycobiology*. 2004;14(1):39-49.
- Daniotti JL, Lardone RD, Vilcaes AA. Dysregulated Expression of Glycolipids in Tumor Cells: From Negative Modulator of Anti-tumor Immunity to Promising Targets for Developing Therapeutic Agents. *Front Oncol*. 2016;5.
- Daniotti JL, Vilcaes AA, Torres Demichelis V, Ruggiero FM, Rodriguez-Walker M. Glycosylation of Glycolipids in Cancer: Basis for Development of Novel Therapeutic Approaches. *Front Oncol*. 2013;3.
- Datta AK. Comparative sequence analysis in the sialyltransferase protein family: analysis of motifs. *Curr Drug Targets*. 2009;10(6):483-98.
- Datta AK, Chammas R, Paulson JC. Conserved cysteines in the sialyltransferase sialylmotifs form an essential disulfide bond. *J Biol Chem*. 4 mai 2001;276(18):15200-7.

- Datta AK, Sinha A, Paulson JC. Mutation of the sialyltransferase S-sialylmotif alters the kinetics of the donor and acceptor substrates. *J Biol Chem.* 1998;273(16):9608-14.
- Datta AK, Paulson J. Sialylmotifs of sialyltransferases. *Indian J Biochem Biophys.* 1997;34(1-2):157-65.
- Datta AK, Paulson JC. The sialyltransferase « sialylmotif » participates in binding the donor substrate CMP-NeuAc. *J Biol Chem.* 1995;270(4):1497-500.
- Davies LRL, Varki A. Why Is N-Glycolylneuraminic Acid Rare in the Vertebrate Brain? In: *SialoGlyco Chemistry and Biology I. Top Curr Chem.* 2015; 366: 31-54.
- Davies LRL, Pearce OMT, Tessier MB, Assar S, Smutova V, Pajunen M, *et al.* Metabolism of Vertebrate Amino Sugars with N-Glycolyl Groups. *Journal of Biological Chemistry.* 2012;287(34):28917-31.
- de Jong H, Moure MJ, Hartman JEM, Bosman GP, Ong JY, Bardoel BW, *et al.* Selective Exoenzymatic Labeling of Lipooligosaccharides of *Neisseria gonorrhoeae* with α 2,6-Sialoside Analogues. *ChemBioChem.* 2022 ; 23(19).
- Delgado C, Bu L, Zhang J, Liu FY, Sall J, Liang FX, *et al.* Neural cell adhesion molecule is required for ventricular conduction system development. *Development.* 2021;148(11):dev199431.
- Diekmann H, Stuermer CAO. Zebrafish neuropilin-a and -b, orthologs of ALCAM, are involved in retinal ganglion cell differentiation and retinal axon pathfinding. *J Comp Neurol.* 2009;513(1):38-50.
- Dobie C, Skropeta D. Insights into the role of sialylation in cancer progression and metastasis. *Br J Cancer.* 2021;124(1):76-90.
- Dobie C, Montgomery AP, Szabo R, Skropeta D, Yu H. Computer-aided design of human sialyltransferase inhibitors of hST8Sia III. *J Mol Recognit* 2018;31(2).
- Dodd J, Morton SB, Karageorgos D, Yamamoto M, Jessell TM. Spatial regulation of axonal glycoprotein expression on subsets of embryonic spinal neurons. *Neuron.* 1988;1(2):105-16.
- Donadio S, Dubois C, Fichant G, Roybon L, Guillemot JC, Breton C, *et al.* Recognition of cell surface acceptors by two human alpha-2,6-sialyltransferases produced in CHO cells. *Biochimie.* 2003;85(3-4):311-21.
- Dorsett KA, Marciel MP, Hwang J, Ankenbauer KE, Bhalerao N, Bellis SL. Regulation of ST6GAL1 sialyltransferase expression in cancer cells. *Glycobiology.* 2021;31(5):530-9.
- Drake PM, Nathan JK, Stock CM, Chang PV, Muench MO, Nakata D, *et al.* Polysialic acid, a glycan with highly restricted expression, is found on human and murine leukocytes and modulates immune responses. *J Immunol.* 2008;181(10):6850-8.
- Drivsholm L, Vangsted A, Pallesen T, Hansen M, Dombernowsky P, Hirsch F, *et al.* Fucosyl-GM1 in small-cell lung cancer. A comparison with the tumour marker neuron-specific enolase. *Ann Oncol.* 1994;5(7):623-6.
- Drula E, Garron ML, Dogan S, Lombard V, Henrissat B, Terrapon N. The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res.* 2022;50(D1):D571-7.
- Dubois C, Manuguerra JC, Hauttecoeur B, Maze J. Monoclonal antibody A2B5, which detects cell surface antigens, binds to ganglioside GT3 (II3 (NeuAc)3LacCer) and to its 9-O-acetylated derivative. *J Biol Chem.* 1990;265(5):2797-803.
- Eckhardt M, Bukalo O, Chazal G, Wang L, Goridis C, Schachner M, *et al.* Mice Deficient in the Polysialyltransferase ST8SiaIV/PST-1 Allow Discrimination of the Roles of Neural Cell Adhesion Molecule Protein and Polysialic Acid in Neural Development and Synaptic Plasticity. *J Neurosci.* 2000;20(14):5234-44.
- Eckhardt M, Mühlenhoff M, Bethe A, Gerardy-Schahn R. Expression cloning of the Golgi CMP-sialic acid transporter. *Proc Natl Acad Sci USA.* 1996;93(15):7572-6.
- Edvardson S, Baumann AM, Mühlenhoff M, Stephan O, Kuss AW, Shaag A, *et al.* West syndrome caused by ST3Gal-III deficiency. *Epilepsia.* 2013;54(2):e24-27.
- Ehrit J, Keys TG, Sutherland M, Wolf S, Meier C, Falconer RA, *et al.* Exploring and Exploiting Acceptor Preferences of the Human Polysialyltransferases as a Basis for an Inhibitor Screen. *ChemBioChem.* 2017;18(13):1332-7.
- El Maarouf A, Petridis AK, Rutishauser U. Use of polysialic acid in repair of the central nervous system. *Proc Natl Acad Sci U S A.* 2006;103(45):16989-94.
- El Yazidi-Belkoura I, Lefebvre T. O-GlcNAcylation and Cancer. *MDPI*, pp.112, 2022, 978-30365-3082-6.

- Elkashaf SM, Allison SJ, Sadiq M, Basheer HA, Ribeiro Morais G, Loadman PM, *et al.* Polysialic acid sustains cancer cell survival and migratory capacity in a hypoxic environment. *Sci Rep.* 2016;6:33026.
- Ertunc N, Phitak T, Wu D, Fujita H, Hane M, Sato C, *et al.* Sulfation of sialic acid is ubiquitous and essential for vertebrate development. *Sci Rep.* 2022;12(1):12496.
- Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, *et al.* Protein complex prediction with AlphaFold-Multimer. *DeepMind BioRxiv.* 2022.
- Eyre TA, Ducluzeau F, Sneddon TP, Povey S, Bruford EA, Lush MJ. The HUGO Gene Nomenclature Database, 2006 updates. *Nucleic Acids Research.* 2006;34(90001):D319-21.
- Falconer RA, Errington RJ, Shnyder SD, Smith PJ, Patterson LH. Polysialyltransferase: a new target in metastatic cancer. *Curr Cancer Drug Targets.* 2012;12(8):925-39.
- Felder M, Kapur A, Gonzalez-Bosquet J, Horibata S, Heintz J, Albrecht R, *et al.* MUC16 (CA125): tumor biomarker to cancer therapy, a work in progress. *Mol Cancer.* 2014;13(1):129.
- Feng Y, Wang L, Wu YL, Liu HH, Ma F. [Roles of sialic acids in sperm maturation and capacitation and sperm-egg recognition]. *Zhonghua Nan Ke Xue.* 2016;22(10):944-8.
- Fenteany FH, Colley KJ. Multiple signals are required for alpha2,6-sialyltransferase (ST6Gal I) oligomerization and Golgi localization. *J Biol Chem.* 2005;280(7):5423-9.
- Fernandes AI, Gregoriadis G. The effect of polysialylation on the immunogenicity and antigenicity of asparaginase: implication in its pharmacokinetics. *Int J Pharm.* 2001;217(1-2):215-24.
- Ferragut F, Cagnoni AJ, Colombo LL, Sánchez Terrero C, Wolfenstein-Todel C, Troncoso MF, *et al.* Dual knockdown of Galectin-8 and its glycosylated ligand, the activated leukocyte cell adhesion molecule (ALCAM/CD166), synergistically delays *in vivo* breast cancer growth. *Biochimica et Biophysica Acta - Molecular Cell Research.* 2019;1866(8):1338-52.
- Ferreira I, Carrascal M, Mineiro A, Bugalho A, Borralho P, Silva Z, *et al.* Carcinoembryonic antigen is a sialyl Lewis x/a carrier and an E-selectin ligand in non-small cell lung cancer. *Int J Oncol.* 2019.
- Finne J, Mäkelä PH. Cleavage of the polysialosyl units of brain glycoproteins by a bacteriophage endosialidase. Involvement of a long oligosaccharide segment in molecular interactions of polysialic acid. *Journal of Biological Chemistry.* 1985;260(2):1265-70.
- Finne J, Finne U, Deagostini-Bazin H, Goridis C. Occurrence of alpha 2-8 linked polysialosyl units in a neural cell adhesion molecule. *Biochem Biophys Res Commun.* 1983;112(2):482-7.
- Fleischer B. Mechanism of glycosylation in the Golgi apparatus. *J Histochem Cytochem.* 1983;31(8):1033-40.
- Fliniaux I, Marchand G, Molinaro C, Decloquement M, Martoriati A, Marin M, *et al.* Diversity of sialic acids and sialoglycoproteins in gametes and at fertilization. *Front Cell Dev Biol.* 2022;10:982931.
- Flitsch SL. 7.21 Enzymatic Carbohydrate Synthesis. In: *Comprehensive Chirality.* Elsevier. 2012 p. 454-64.
- Flynn RA, Pedram K, Malaker SA, Batista PJ, Smith BAH, Johnson AG, *et al.* Small RNAs are modified with N-glycans and displayed on the surface of living cells. *Cell.* 2021;184(12):3109-3124.e22.
- Foley DA, Swartzentruber KG, Colley KJ. Identification of Sequences in the Polysialyltransferases ST8Sia II and ST8Sia IV That Are Required for the Protein-specific Polysialylation of the Neural Cell Adhesion Molecule, NCAM. *Journal of Biological Chemistry.* 2009;284(23):15505-16.
- Freeze HH, Elbein AD. *Glycosylation Precursors. Essentials of Glycobiology.* 2nd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009.
- Frosch M, Görden I, Boulnois GJ, Timmis KN, Bitter-Suermann D. NZB mouse system for production of monoclonal antibodies to weak bacterial antigens: isolation of an IgG antibody to the polysaccharide capsules of *Escherichia coli* K1 and group B meningococci. *Proc Natl Acad Sci USA.* 1985;82(4):1194-8.
- Fujitani N, Furukawa J, Araki K, Fujioka T, Takegawa Y, Piao J, *et al.* Total cellular glycomics allows characterizing cells and streamlining the discovery process for cellular biomarkers. *Proc Natl Acad Sci USA.* 2013;110(6):2105-10.
- Fujitani N, Takegawa Y, Ishibashi Y, Araki K, Furukawa J, Mitsutake S, *et al.* Qualitative and Quantitative Cellular Glycomics of Glycosphingolipids Based on Rhodococcal Endoglycosylceramidase-assisted Glycan Cleavage,

- Glycoblotting-assisted Sample Preparation, and Matrix-assisted Laser Desorption Ionization Tandem Time-of-flight Mass Spectrometry Analysis. *J Biol Chem.* 2011;286(48):41669-79.
- Fukumoto S, Miyazaki H, Goto G, Urano T, Furukawa K, Furukawa K. Expression cloning of mouse cDNA of CMP-NeuAc:Lactosylceramide alpha2,3-sialyltransferase, an enzyme that initiates the synthesis of gangliosides. *J Biol Chem.* 1999;274(14):9271-6.
- Gallo A, Costantini M. Glycobiology of Reproductive Processes in Marine Animals: The State of the Art. *Marine Drugs.* 2012;10(12):2861-92.
- Galuska CE, Dambon JA, Kühnle A, Bornhöfft KF, Prem G, Zlatina K, *et al.* Artificial Polysialic Acid Chains as Sialidase-Resistant Molecular-Anchors to Accumulate Particles on Neutrophil Extracellular Traps. *Front Immunol.* 2017c;8:1229.
- Galuska SP, Galuska CE, Tharmalingam T, Zlatina K, Prem G, Husejnov FCO, *et al.* *In vitro* generation of polysialylated cervical mucins by bacterial polysialyltransferases to counteract cytotoxicity of extracellular histones. *FEBS J.* 2017b;284(11):1688-99.
- Galuska C, Lütteke T, Galuska S. Is Polysialylated NCAM Not Only a Regulator during Brain Development But also during the Formation of Other Organs? *Biology.* 2017a;6(4):27.
- Galuska SP. Advances in Sialic Acid and Polysialic Acid Detection Methodologies. In: *Sialobiology: Structure, Biosynthesis and Function*, Bentham Science Publishers; 2013, 448-75.
- Galuska SP, Geyer H, Weinhold B, Kontou M, Rohrich RC, Bernard U, Gerardy-Schahn R, Reutter W, Munster-Kuhnel A, Geyer R. Quantification of nucleotide-activated sialic acids by a combination of reduction and fluorescent labeling. *Anal Chem.* 2010b;82(11):4591-4598.
- Galuska SP, Rollenhagen M, Kaup M, Eggers K, Oltmann-Norden I, Schiff M, *et al.* Synaptic cell adhesion molecule SynCAM1 is a target for polysialylation in postnatal mouse brain. *Proc Natl Acad Sci USA.* 2010a;107(22):10250-5.
- Galuska SP, Geyer R, Gerardy-Schahn R, Mühlenhoff M, Geyer H. Enzyme-dependent Variations in the Polysialylation of the Neural Cell Adhesion Molecule (NCAM) *in Vivo*. *Journal of Biological Chemistry.* 2008;283(1):17-28.
- Galuska SP, Oltmann-Norden I, Geyer H, Weinhold B, Kuchelmeister K, Hildebrandt H, *et al.* Polysialic acid profiles of mice expressing variant allelic combinations of the polysialyltransferases ST8SiaII and ST8SiaIV. *J Biol Chem.* 2006;281(42):31605-15.
- Garnham R, Scott E, Livermore KE, Munkley J. ST6GAL1: A key player in cancer. *Oncol Lett.* 2019;18(2):983-9.
- Geisler C, Jarvis DL. Letter to the Glyco-Forum: Effective glycoanalysis with *Maackia amurensis* lectins requires a clear understanding of their binding specificities. *Glycobiology.* 2011;21(8):988-93.
- Geissner A, Baumann L, Morley TJ, Wong AKO, Sim L, Rich JR, *et al.* 7-Fluorosialyl Glycosides Are Hydrolysis Resistant but Readily Assembled by Sialyltransferases Providing Easy Access to More Metabolically Stable Glycoproteins. *ACS Cent Sci.* 2021;7(2):345-54.
- Geremia RA, Harduin-Lepers A, Delannoy P. Identification of two novel conserved amino acid residues in eukaryotic sialyltransferases: implications for their mechanism of action. *Glycobiology.* 1997;7(2):v-vii.
- Gilormini PA, Batt AR, Pratt MR, Biot C. Asking more from metabolic oligosaccharide engineering. *Chem Sci.* 2018;9(39):7585-95.
- Gilormini P-A, Lion C, Noel M, Krzewinski-Recchi M-A, Harduin-Lepers A, Guérardel Y, *et al.* Improved workflow for the efficient preparation of ready to use CMP-activated sialic acids. *Glycobiology.* 2016a;26(11):1151-6.
- Gilormini P-A, Lion C, Vicogne D, Levade T, Potelle S, Mariller C, *et al.* Sequential Bioorthogonal Dual Strategy: ManNAI and SiaNAI as distinct tools to unravel sialic acids metabolic pathways. *Chem Commun (Camb).* 2016b;52(11):2318-21.
- Giordanengo V, Bannwarth S, Laffont C, Van Miegem V, Harduin-Lepers A, Delannoy P, *et al.* Cloning and expression of cDNA for a human Gal(beta1-3)GalNAc alpha2,3-sialyltransferase from the CEM T-cell line. *Eur J Biochem.* 1997;247(2):558-66.
- Gomord V, Fitchette AC, Menu-Bouaouiche L, Saint-Jore-Dupas C, Plasson C, Michaud D, *et al.* Plant-specific glycosylation patterns in the context of therapeutic protein production: PMP-specific glycosylation patterns. *Plant Biotechnology Journal.* 2010;8(5):564-87.

- Goni M, Javaregowda PK, Chachadi V, Virupakshaiiah D. Sialic acids: An Avenue to Target Cancer Progression, Metastasis, and Resistance to Therapy. *Forum of Clinical Oncology*. 2021;12(3):40-8.
- Gordon-Lipkin E, Cohen JS, Srivastava S, Soares BP, Levey E, Fatemi A. ST3GAL5-Related Disorders: A Deficiency in Ganglioside Metabolism and a Genetic Cause of Intellectual Disability and Choreoathetosis. *J Child Neurol*. 2018;33(13):825-31.
- Gregoriadis G, Jain S, Papaioannou I, Laing P. Improving the therapeutic efficacy of peptides and proteins: A role for polysialic acids. *International Journal of Pharmaceutics*. 2005;300(1-2):125-30.
- Gregoriadis G, McCormack B, Wang Z, Lively R. Polysialic acids: potential in drug delivery. *FEBS Letters*. 1993;315(3):271-6.
- Grewal RK, Shaikh AR, Gorle S, Kaur M, Videira PA, Cavallo L, *et al*. Structural Insights in Mammalian Sialyltransferases and Fucosyltransferases: We Have Come a Long Way, but It Is Still a Long Way Down. *Molecules*. 2021;26(17):5203.
- Gross HJ, Sticher U, Brossmer R. A highly sensitive fluorometric assay for sialyltransferase activity using CMP-9-fluoresceinyl-NeuAc as donor. *Analytical Biochemistry*. 1990;186(1):127-34.
- Gross HJ, Bunsch A, Paulson JC, Brossmer R. Activation and transfer of novel synthetic 9-substituted sialic acids. *Eur J Biochem*. 1987;168(3):595-602.
- Groux-Degroote S, Vicogne D, Coge V, Schulz C, Harduin-Lepers A. B4GALNT2 Controls Sd^a and SLe^x Antigen Biosynthesis in Healthy and Cancer Human Colon. *ChemBioChem*. 2021;22(24):3381-90.
- Groux-Degroote S, Schulz C, Coge V, Noël M, Portier L, Vicogne D, *et al*. The extended cytoplasmic tail of the human B4GALNT2 is critical for its Golgi targeting and post-Golgi sorting. *FEBS J*. 2018;285(18):3442-63.
- Groux-Degroote S, Wavelet C, Krzewinski-Recchi MA, Portier L, Mortuaire M, Mihalache A, *et al*. B4GALNT2 gene expression controls the biosynthesis of Sda and sialyl Lewis X antigens in healthy and cancer human gastrointestinal tract. *The International Journal of Biochemistry & Cell Biology*. 2014;53:442-9.
- Grundmann U, Nerlich C, Rein T, Zettlmeissl G. Complete cDNA sequence encoding human beta-galactoside alpha-2,6-sialyltransferase. *Nucleic Acids Res*. 1990;18(3):667.
- Guo HB, Perminov A, Bekele S, Kedziora G, Farajollahi S, Varaljay V, *et al*. AlphaFold2 models indicate that protein sequence determines both structure and dynamics. *Sci Rep*. 2022;12(1):10696.
- Guo J, Li W, Xue W, Ye XS. Transition State-Based Sialyltransferase Inhibitors: Mimicking Oxocarbenium Ion by Simple Amide. *J Med Chem*. 2017;60(5):2135-41.
- Guo X, Malcolm JR, Ali MM, Ribeiro Morais G, Shnyder SD, Loadman PM, *et al*. An efficient assay for identification and quantitative evaluation of potential polysialyltransferase inhibitors. *Analyst*. 2020;145(13):4512-21.
- Guo X, Elkashef SM, Loadman PM, Patterson LH, Falconer RA. Recent advances in the analysis of polysialic acid from complex biological systems. *Carbohydrate Polymers*. 2019;224:115145.
- Guerardel Y, Kitajima K, Hitochi S, Megraud F, Ferrand J. Glycan compositions, processes for preparing the same and their uses as a drug. 2015. *US Patent No. US 2005/0209408 A1*.
- Hachem NE, Humpfle L, Simon P, Kaese M, Weinhold B, Günther J, *et al*. The Loss of Polysialic Acid Impairs the Contractile Phenotype of Peritubular Smooth Muscle Cells in the Postnatal Testis. *Cells*. 2021;10(6):1347.
- Hallenbeck PC, Vimr ER, Yu F, Bassler B, Troy FA. Purification and properties of a bacteriophage-induced endo-N-acetylneuraminidase specific for poly-alpha-2,8-sialosyl carbohydrate units. *Journal of Biological Chemistry*. 1987;262(8):3553-61.
- Hane M, Chen DY, Varki A. Human-specific microglial Siglec-11 transcript variant has the potential to affect polysialic acid-mediated brain functions at a distance. *Glycobiology*. 2021;31(3):231-42.
- Hane M, Matsuoka S, Ono S, Miyata S, Kitajima K, Sato C. Protective effects of polysialic acid on proteolytic cleavage of FGF2 and proBDNF/BDNF. *Glycobiology*. 2015;25(10):1112-24.
- Hanzawa K, Suzuki N, Natsuka S. Structures and developmental alterations of N-glycans of zebrafish embryos. *Glycobiology*. 2017;27(3):228-245.
- Harduin-Lepers. The vertebrate sialylation machinery: structure-function and molecular evolution of sialyltransferases. *Glycoconjugate Journal*. 2023.

- Harduin-Lepers A. Vertebrate Sialyltransferases. In *Sialobiology: Structure, Biosynthesis and Function*, 2013, 139-187.
- Harduin-Lepers A, Krzewinski-Recchi MA, Colomb F, Foulquier F, Groux-Degroote S, Delannoy P. Sialyltransferases functions in cancers. *Front Biosci*. 2012;4(1):499-515.
- Harduin-Lepers A. Comprehensive Analysis of Sialyltransferases in Vertebrate Genomes. *Glycobiology Insights*. 2010;2:29-61.
- Harduin-Lepers A, Petit D, Mollicone R, Delannoy P, Petit J-M, Oriol R. Evolutionary history of the α 2,8-sialyltransferase (ST8Sia) gene family: Tandem duplications in early deuterostomes explain most of the diversity found in the vertebrate ST8Sia genes. *BMC Evolutionary Biology*. 2008;8(1):258.
- Harduin-Lepers A, Mollicone R, Delannoy P, Oriol R. The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. *Glycobiology*. 2005;15(8):805-17.
- Harduin-Lepers A, Vallejo-Ruiz V, Krzewinski-Recchi M-A, Samyn-Petit B, Julien S, Delannoy P. The human sialyltransferase family. *Biochimie*. 2001;83(8):727-37.
- Harrus D, Harduin-Lepers A, Glumoff T. Unliganded and CMP-Neu5Ac bound structures of human α -2,6-sialyltransferase ST6Gal I at high resolution. *Journal of Structural Biology*. 2020;212(2):107628.
- Harrus D, Kellokumpu S, Glumoff T. Crystal structures of eukaryote glycosyltransferases reveal biologically relevant enzyme homooligomers. *Cell Mol Life Sci*. 2018;75(5):833-48.
- Hartnell A, Steel J, Turley H, Jones M, Jackson DG, Crocker PR. Characterization of human sialoadhesin, a sialic acid binding receptor expressed by resident and inflammatory macrophage populations. *Blood*. 2001;97(1):288-96.
- Haselhorst T, Stummeyer K, Mühlenhoff M, Schaper W, Gerardy-Schahn R, von Itzstein M. Endosialidase NF Appears To Bind PolySia DP5 in a Helical Conformation. *ChemBioChem*. 2006;7(12):1875-7.
- Hassinen A, Rivinoja A, Kauppila A, Kellokumpu S. Golgi *N*-Glycosyltransferases Form Both Homo- and Heterodimeric Enzyme Complexes in Live Cells. *J Biol Chem*. 2010;285(23):17771-7.
- Haukedal H, Freude KK. Implications of Glycosylation in Alzheimer's Disease. *Front Neurosci*. 2021;14:625348.
- Häyrinen J, Haseley S, Talaga P, Mühlenhoff M, Finne J, Vliegthart JFG. High affinity binding of long-chain polysialic acid to antibody, and modulation by divalent cations and polyamines. *Molecular Immunology*. 2002;39(7-8):399-411.
- Häyrinen J, Bitter-Suermann D, Finne J. Interaction of meningococcal group B monoclonal antibody and its Fab fragment with α 2-8-linked sialic acid polymers: Requirement of a long oligosaccharide segment for binding. *Molecular Immunology*. 1989;26(6):523-9.
- Heise T, Langereis JD, Rossing E, de Jonge MI, Adema GJ, Büll C, *et al*. Selective Inhibition of Sialic Acid-Based Molecular Mimicry in *Haemophilus influenzae* Abrogates Serum Resistance. *Cell Chemical Biology*. 2018;25(10):1279-1285.e8.
- Higa HH, Paulson JC. Sialylation of glycoprotein oligosaccharides with N-acetyl-, N-glycolyl-, and N-O-diacetylneuraminic acids. *J Biol Chem*. 1985;260(15):8838-49.
- Higai K, Ishihara S, Matsumoto K. NF κ B-p65 dependent transcriptional regulation of glycosyltransferases in human colon adenocarcinoma HT-29 by stimulation with tumor necrosis factor alpha. *Biol Pharm Bull*. 2006;29(12):2372-7.
- Hildebrandt H, Becker C, Glüer S, Rösner H, Gerardy-Schahn R, Rahmann H. Polysialic acid on the neural cell adhesion molecule correlates with expression of polysialyltransferases and promotes neuroblastoma cell growth. *Cancer Res*. 1998;58(4):779-84.
- Hinderlich S, Weidemann W, Yardeni T, Horstkorte R, Huizing M. UDP-GlcNAc 2-Epimerase/ManNAc Kinase (GNE): A Master Regulator of Sialic Acid Synthesis. *Top Curr Chem*. 2015;366:97-137.
- Hirmo S, Kelm S, Schauer R, Nilsson B, Wadstrom T. Adhesion of *Helicobacter pylori* strains to α -2,3-linked sialic acids. *Glycoconj J*. 1996;13(6):1005-11.
- Hofer TP, van de Loosdrecht AA, Stahl-Hennig C, Cassatella MA, Ziegler-Heitbrock L. 6-Sulfo LacNAc (Slan) as a Marker for Non-classical Monocytes. *Front Immunol*. 2019;10:2052.
- Hong W, Yang J, Zou J, Bi Z, He C, Lei H, *et al*. Histones released by NETosis enhance the infectivity of SARS-CoV-2 by bridging the spike protein subunit 2 and sialic acid on host cells. *Cell Mol Immunol*. 2022;19(5):577-87.

- Horenstein BA, Bruner M. Acid-Catalyzed Solvolysis of CMP- *N*-Acetyl Neuraminate: Evidence for a Sialyl Cation with a Finite Lifetime. *J Am Chem Soc.* 1996;118(43):10371-9.
- Horstkorte R, Mühlhoff M, Reutter W, Nöhring S, Zimmermann-Kordmann M, Gerardy-Schahn R. Selective inhibition of polysialyltransferase ST8SiaII by unnatural sialic acids. *Experimental Cell Research.* 2004;298(1):268-74.
- Houeix B, Cairns MT. Engineering of CHO cells for the production of vertebrate recombinant sialyltransferases. *PeerJ.* 2019;7:e5788.
- Hsu CC, Lin TW, Chang WW, Wu CY, Lo WH, Wang PH, *et al.* Soyasaponin-I-modified invasive behavior of cancer by changing cell surface sialic acids. *Gynecol Oncol.* 2005;96(2):415-22.
- Huang W, Sun L, Wang B, Ma Y, Yao D, Han W, *et al.* Ginsenosides, potent inhibitors of sialyltransferase. *Z Naturforsch C J Biosci.* 2020;75(1-2):41-9.
- Huang RB, Cheng D, Liao SM, Lu B, Wang QY, Xie NZ, *et al.* The Intrinsic Relationship Between Structure and Function of the Sialyltransferase ST8Sia Family Members. *Current Topics in Medicinal Chemistry.* 2017;17(21):2359-2369.
- Hudgin RL, Schachter H. Porcine sugar nucleotide: glycoprotein glycosyltransferases. 3. Blood serum and liver N-acetylglucosaminyltransferase. *Can J Biochem.* 1971;49(7):847-52.
- Hugonnet M, Singh P, Haas Q, von Gunten S. The Distinct Roles of Sialyltransferases in Cancer Biology and Onco-Immunology. *Front Immunol.* 2021;12:799861.
- Hunter C, Gao Z, Chen HM, Thompson N, Wakarchuk W, Withers SG, *et al.* Attenuation of polysialic acid biosynthesis in cells by the small molecule inhibitor 8-keto-sialic acid. *ACS Chem. Biol.* 2022.
- Ikedo Y, Takahashi M. Glycosyltransferases and Glycosidases: Enzyme Mechanisms. In: *Comprehensive Glycoscience.* Elsevier; 2007. p. 115-28.
- Ikehara Y, Kojima N, Kurosawa N, Kudo T, Kono M, Nishihara S, *et al.* Cloning and expression of a human gene encoding an N-acetylgalactosamine- α 2,6-sialyltransferase (ST6GalNAc I): a candidate for synthesis of cancer-associated sialyl-Tn antigens. *Glycobiology.* 1999a;9(11):1213-24.
- Ikehara Y, Shimizu N, Kono M, Nishihara S, Nakanishi H, Kitamura T, *et al.* A novel glycosyltransferase with a polyglutamine repeat; a new candidate for GD1 α synthase (ST6GalNAc V)(1). *FEBS Lett.* 1999b;463(1-2):92-6.
- Imre T, Kremmer T, Héberger K, Molnár-Szöllösi É, Ludányi K, Pócsfalvi G, *et al.* Mass spectrometric and linear discriminant analysis of N-glycans of human serum α -1-acid glycoprotein in cancer patients and healthy individuals. *Journal of Proteomics.* 2008;71(2):186-97.
- Inamori K, Inokuchi J. Ganglioside GM3 Synthase Deficiency in Mouse Models and Human Patients. *International Journal of Molecular Sciences.* 2022;23(10):5368.
- Indelicato R, Parini R, Domenighini R, Malagolini N, Iascone M, Gasperini S, *et al.* Total loss of GM3 synthase activity by a normally processed enzyme in a novel variant and in all ST3GAL5 variants reported to cause a distinct congenital disorder of glycosylation. *Glycobiology.* 2019;29(3):229-41.
- Inoue, S. & Inoue, Y. Fish glycoproteins. *New Comprehensive Biochemistry.* Elsevier. 1997;29, 143 –161.
- Inoue S, Kanamori A, Kitajima K, Inoue Y. KDN-glycoprotein: a novel deaminated neuraminic acid-rich glycoprotein isolated from vitelline envelope of rainbow trout eggs. *Biochem Biophys Res Commun.* 1988;153(1):172-6.
- Irie A, Koyama S, Kozutsumi Y, Kawasaki T, Suzuki A. The Molecular Basis for the Absence of N-Glycolylneuraminic Acid in Humans. *Journal of Biological Chemistry.* 1998;273(25):15866-71.
- Ishii A, Ohta M, Watanabe Y, Matsuda K, Ishiyama K, Sakoe K, *et al.* Expression cloning and functional characterization of human cDNA for ganglioside GM3 synthase. *J Biol Chem.* 1998;273(48):31652-5.
- Isomura R, Kitajima K, Sato C. Structural and functional impairments of polysialic acid by a mutated polysialyltransferase found in schizophrenia. *J Biol Chem.* 2011(286): 21535–45.
- Iwasaki M, Inoue S, Troy FA. A new sialic acid analogue, 9-O-acetyl-deaminated neuraminic acid, and α -2,8-linked O-acetylated poly(N-glycolylneuraminy) chains in a novel polysialoglycoprotein from salmon eggs. *J Biol Chem.* 1990;265(5):2596-602.

- Iwasaki M, Inoue S. Structures of the carbohydrate units of polysialoglycoproteins isolated from the eggs of four species of salmonid fishes. *Glycoconjugate J.* 1985;2(3-4):209-28.
- Iwasaki M, Inoue S, Kitajima K, Nomoto H, Inoue Y. Novel oligosaccharide chains on polysialoglycoproteins isolated from rainbow trout eggs. A unique carbohydrate sequence with a sialidase-resistant sialyl group, DGalNAc β 1-4(NeuGc2-3)DGalNAc. *Biochemistry.* 1984;23(2):305-10.
- Jaffé SR, Strutton B, Levarski Z, Pandhal J, Wright PC. *Escherichia coli* as a glycoprotein production host: recent developments and challenges. *Current Opinion in Biotechnology.* 2014;30:205-10.
- Jakobsson E, Schwarzer D, Jokilampi A, Finne J. Endosialidases: Versatile Tools for the Study of Polysialic Acid. *SialoGlyco Chemistry and Biology II.* 2012;367:29-73.
- Jame-Chenarboo F, Ng HH, Macdonald D, Mahal LK. High-Throughput Analysis Reveals miRNA Upregulating α -2,6-Sialic Acid through Direct miRNA-mRNA Interactions. *ACS Cent Sci.* 2022;8(11):1527-36.
- James WM, Agnew WS. Alpha-(2-8)-polysialic acid immunoreactivity in voltage-sensitive sodium channel of eel electric organ. *Proc R Soc Lond B Biol Sci.* 1989;237(1287):233-45.
- Janas T, Janas T. Membrane oligo- and polysialic acids. *Biochim Biophys Acta.* 2011;1808(12):2923-32.
- Janesch B, Saxena H, Sim L, Wakarchuk WW. Comparison of α 2,6-sialyltransferases for sialylation of therapeutic proteins. *Glycobiology.* 2019b;29(10):735-47.
- Janesch B, Baumann L, Mark A, Thompson N, Rahmani S, Sim L, *et al.* Directed evolution of bacterial polysialyltransferases. *Glycobiology.* 2019a;29(7):588-98.
- Jarahian M, Marofi F, Maashi MS, Ghaebi M, Khezri A, Berger MR. Re-Expression of Poly/Oligo-Sialylated Adhesion Molecules on the Surface of Tumor Cells Disrupts Their Interaction with Immune-Effector Cells and Contributes to Pathophysiological Immune Escape. *Cancers.* 2021;13(20):5203.
- Jaroentomechai T, Kwon YH, Liu Y, Young O, Bhawal R, Wilson JD, *et al.* A universal glycoenzyme biosynthesis pipeline that enables efficient cell-free remodeling of glycans. *Nat Commun.* 2022;13(1):6325.
- Jeanneau C, Chazalet V, Augé C, Soumpasis DM, Harduin-Lepers A, Delannoy P, *et al.* Structure-function analysis of the human sialyltransferase ST3Gal I: role of n-glycosylation and a novel conserved sialylmotif. *J Biol Chem.* 2004;279(14):13461-8.
- Jennings HJ, Roy R, Michon F. Determinant specificities of the groups B and C polysaccharides of *Neisseria meningitidis*. *J Immunol.* 1985;134(4):2651-7.
- Jevsevar S, Kunstelj M, Porekar VG. PEGylation of therapeutic proteins. *Biotechnol J.* 2010;5(1):113-28.
- Ji S, Wang F, Chen Y, Yang C, Zhang P, Zhang X, *et al.* Developmental changes in the level of free and conjugated sialic acids, Neu5Ac, Neu5Gc and KDN in different organs of pig: a LC-MS/MS quantitative analyses. *Glycoconj J.* 2017;34(1):21-30.
- Jin C, Padra JT, Sundell K, Sundh H, Karlsson NG, Lindén SK. Atlantic Salmon Carries a Range of Novel O -Glycan Structures Differentially Localized on Skin and Intestinal Mucins. *J Proteome Res.* 2015;14(8):3239-51.
- Jokilampi A, Ollikka P, Korja M, Jakobsson E, Loimaranta V, Haataja S, *et al.* Construction of antibody mimics from a noncatalytic enzyme—detection of polysialic acid. *Journal of Immunological Methods.* 2004;295(1-2):149-60.
- Kakehi K, Kinoshita M, Kitano K, Morita M, Oda Y. Lactone formation of N-acetylneuraminic acid oligomers and polymers as examined by capillary electrophoresis. *Electrophoresis.* 2001;22(16):3466-70.
- Kanamori A, Kitajima K, Inoue Y, Inoue S, Xulei Z, Zuber C, *et al.* Monoclonal antibody specific for α 2-8-linked oligo deaminated neuraminic acid (KDN) sequences in glycoproteins. Preparation and characterization of a monoclonal antibody and its application in immunohistochemistry. *Histochemistry.* 1994;101(5):333-40.
- Kanamori A, Inoue S, Iwasaki M, Kitajima K, Kawai G, Yokoyama S, *et al.* Deaminated neuraminic acid-rich glycoprotein of rainbow trout egg vitelline envelope. Occurrence of a novel alpha-2,8-linked oligo(deaminated neuraminic acid) structure in O-linked glycan chains. *Journal of Biological Chemistry.* 1990;265(35):21811-9.
- Karlsson NG, Thomsson KA. Salivary MUC7 is a major carrier of blood group I type O-linked oligosaccharides serving as the scaffold for sialyl Lewis x. *Glycobiology.* 2009;19(3):288-300.

- Karlstetter M, Kopatz J, Aslanidis A, Shahraz A, Caramoy A, Linnartz-Gerlach B, *et al.* Polysialic acid blocks mononuclear phagocyte reactivity, inhibits complement activation, and protects from vascular damage in the retina. *EMBO Mol Med.* 2017;9(2):154-66.
- Kayser H, Zeitler R, Kannicht C, Grunow D, Nuck R, Reutter W. Biosynthesis of a nonphysiological sialic acid in different rat organs, using N-propanoyl-D-hexosamines as precursors. *Journal of Biological Chemistry.* 1992;267(24):16934-8.
- Kellokumpu S. Golgi pH, Ion and Redox Homeostasis: How Much Do They Really Matter? *Front Cell Dev Biol.* 2019;7:93.
- Khamirani HJ, Zoghi S, Faghihi F, Dastgheib SA, Hassanipour H, Bagher Tabei SM, *et al.* Phenotype of ST3GAL3 deficient patients: A case and review of the literature. *European Journal of Medical Genetics.* 2021;64(8):104250.
- Khoder-Agha F, Harrus D, Brysbaert G, Lensink MF, Harduin-Lepers A, Glumoff T, *et al.* Assembly of B4GALT1/ST6GAL1 heteromers in the Golgi membranes involves lateral interactions via highly charged surface domains. *Journal of Biological Chemistry.* 2019;294(39):14383-93.
- Kiermaier E, Mousson C, Veldkamp CT, Gerardy-Schahn R, de Vries I, Williams LG, *et al.* Polysialylation controls dendritic cell trafficking by regulating chemokine recognition. *Science.* 2016;351(6269):186-90.
- Kim DS & Hahn Y. The acquisition of novel N-glycosylation sites in conserved proteins during human evolution. *BMC Bioinformatics.* 2015;16(1):29.
- Kim J, Ryu C, Ha J, Lee J, Kim D, Ji M, *et al.* Structural and Quantitative Characterization of Mucin-Type O-Glycans and the Identification of O-Glycosylation Sites in Bovine Submaxillary Mucin. *Biomolecules.* 2020;10(4):636.
- Kim KW, Kim SW, Min KS, Kim CH, Lee YC. Genomic structure of human GM3 synthase gene (hST3Gal V) and identification of mRNA isoforms in the 5'-untranslated region. *Gene.* 2001;273(2):163-71.
- Kim YJ, Kim KS, Do S, Kim CH, Kim SK, Lee YC. Molecular cloning and expression of human alpha2,8-sialyltransferase (hST8Sia V). *Biochem Biophys Res Commun.* 1997;235(2):327-30.
- Kim YJ, Kim KS, Kim SH, Kim CH, Ko JH, Choe IS, *et al.* Molecular cloning and expression of human Gal beta 1,3GalNAc alpha 2,3-sialyltransferase (hST3Gal II). *Biochem Biophys Res Commun.* 1996;228(2):324-7.
- Kitajima K, Inoue S, Inoue Y, Troy FA. Use of a bacteriophage-derived endo-N-acetylneuraminidase and an equine antipolysialyl antibody to characterize the polysialyl residues in salmonid fish egg polysialoglycoproteins. Substrate and immunospecificity studies. *Journal of Biological Chemistry.* 1988;263(34):18269-76.
- Kitagawa H, Paulson JC. Cloning of a novel alpha 2,3-sialyltransferase that sialylates glycoprotein and glycolipid carbohydrate groups. *J Biol Chem.* 1994b;269(2):1394-401.
- Kitagawa H, Paulson JC. Differential expression of five sialyltransferase genes in human tissues. *J Biol Chem.* 1994a;269(27):17872-8.
- Kitagawa H, Paulson JC. Cloning and expression of human Gal beta 1,3(4)GlcNAc alpha 2,3-sialyltransferase. *Biochem Biophys Res Commun.* 1993;194(1):375-82.
- Kitajima K, Varki N, Sato C. Advanced Technologies in Sialic Acid and Sialoglycoconjugate Analysis. In: *SialoGlyco Chemistry and Biology II.* Cham: Springer International Publishing. 2013;367:75-103.
- Kitajima K, Inoue S, Inoue Y, Troy FA. Use of a bacteriophage-derived endo-N-acetylneuraminidase and an equine antipolysialyl antibody to characterize the polysialyl residues in salmonid fish egg polysialoglycoproteins. Substrate and immunospecificity studies. *Journal of Biological Chemistry.* 1988;263(34):18269-76.
- Kitajima K, Inoue Y, Inoue S. Polysialoglycoproteins of Salmonidae fish eggs. Complete structure of 200-kDa polysialoglycoprotein from the unfertilized eggs of rainbow trout (*Salmo gairdneri*). *Journal of Biological Chemistry.* 1986;261(12):5262-9.
- Kitazume S, Nakagawa K, Oka R, Tachida Y, Ogawa K, Luo Y, *et al.* In vivo cleavage of alpha2,6-sialyltransferase by Alzheimer beta-secretase. *J Biol Chem.* 2005;280(9):8589-95.
- Kitazume S, Tachida Y, Oka R, Kotani N, Ogawa K, Suzuki M, *et al.* Characterization of alpha 2,6-sialyltransferase cleavage by Alzheimer's beta -secretase (BACE1). *J Biol Chem.* 2003;278(17):14865-71.
- Kitazume S, Kitajima K, Inoue S, Haslam SM, Morris HR, Dell A, *et al.* The occurrence of novel 9-O-sulfated N-glycolylneuraminic acid-capped alpha2-->5-Oglycolyl-linked oligo/polyNeu5Gc chains in sea urchin egg cell

- surface glycoprotein. Identification of a new chain termination signal for polysialyltransferase. *J Biol Chem.* 1996;271(12):6694-701.
- Kitazume S, Kitajima K, Inoue S, Inoue Y, Troy F. Developmental expression of trout egg polysialoglycoproteins and the prerequisite α 2,6-, and α 2,8-Sialyl and α 2,8-polysialyltransferase activities required for their synthesis during oogenesis. *J Biol Chem.* 1994;269(14):10330-40.
- Kleene R, Schachner M. Glycans and neural cell interactions. *Nat Rev Neurosci.* 2004;5(3):195-208.
- Koga M, Gilbert M, Li J, Koike S, Takahashi M, Furukawa K, *et al.* Antecedent infections in Fisher syndrome: A common pathogenesis of molecular mimicry. *Neurology.* 2005;64(9):1605-11.
- Kojima N, Kono M, Yoshida Y, Tachida Y, Nakafuku M, Tsuji S. Biosynthesis and expression of polysialic acid on the neural cell adhesion molecule is predominantly directed by ST8Sia II/STX during in vitro neuronal differentiation. *J Biol Chem.* 1996;271(36):22058-62.
- Kojima N, Lee YC, Hamamoto T, Kurosawa N, Tsuji S. Kinetic properties and acceptor substrate preferences of two kinds of Gal beta 1,3GalNAc alpha 2,3-sialyltransferase from mouse brain. *Biochemistry.* 1994;33(19):5772-6.
- Kolb HC, Finn MG, Sharpless K B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew Chem Int Ed Engl.* 2001;40, 2004-2021.
- Koles K, Repnikova E, Pavlova G, Korochkin LI, Panin VM. Sialylation in protostomes: a perspective from Drosophila genetics and biochemistry. *Glycoconj J.* 2009;26(3):313-24.
- Koles K, Irvine KD, Panin VM. Functional characterization of Drosophila sialyltransferase. *J Biol Chem.* 2004;279(6):4346-57.
- Kolter T, Sandhoff K. Sialic acids--why always alpha-linked? *Glycobiology.* 1997;7(7):vii-ix.
- Komminoth P, Roth J, Lackie PM, Bitter-Suermann D, Heitz PU. Polysialic acid of the neural cell adhesion molecule distinguishes small cell lung carcinoma from carcinoids. *Am J Pathol.* 1991;139(2):297-304.
- Kono M, Tsuda T, Ogata S, Takashima S, Liu H, Hamamoto T, *et al.* Redefined substrate specificity of ST6GalNAc II: a second candidate sialyl-Tn synthase. *Biochem Biophys Res Commun.* 2000;272(1):94-7.
- Kono M, Ohyama Y, Lee YC, Hamamoto T, Kojima N, Tsuji S. Mouse beta-galactoside alpha 2,3-sialyltransferases: comparison of in vitro substrate specificities and tissue specific expression. *Glycobiology.* 1997;7(4):469-79.
- Kroes RA, Moskal JR. The role of DNA methylation in ST6Gal1 expression in gliomas. *Glycobiology.* 2016;26(12):1271-83.
- Krug LM, Ragupathi G, Hood C, George C, Hong F, Shen R, *et al.* Immunization with N-propionyl polysialic acid-KLH conjugate in patients with small cell lung cancer is safe and induces IgM antibodies reactive with SCLC cells and bactericidal against group B meningococci. *Cancer Immunol Immunother.* 2012;61(1):9-18.
- Krzewinski-Recchi MA, Julien S, Juliant S, Teinturier-Lelièvre M, Samyn-Petit B, Montiel MD, *et al.* Identification and functional expression of a second human beta-galactoside alpha2,6-sialyltransferase, ST6Gal II. *Eur J Biochem.* 2003;270(5):950-61.
- Kuhn B, Benz J, Greif M, Engel AM, Sobek H, Rudolph MG. The structure of human α -2,6-sialyltransferase reveals the binding mode of complex glycans. *Acta Crystallogr D Biol Crystallogr.* 2013;69(Pt 9):1826-38.
- Kühnle A, Galuska CE, Zlatina K, Galuska SP. (2019) The Bovine Antimicrobial Peptide Lactoferricin Interacts with Polysialic Acid without Loss of Its Antimicrobial Activity against Escherichia coli. *Animals (Basel).* 2019b;10(1):1.
- Kühnle A, Veelken R, Galuska CE, Saftenberger M, Verleih M, Schuppe HC, Rudloff S, Kunz C, Galuska SP. Polysialic acid interacts with lactoferrin and supports its activity to inhibit the release of neutrophil extracellular traps. *Carbohydr Polym.* 2019a; 208, 32-41.
- Kurosawa N, Takashima S, Kono M, Ikehara Y, Inoue M, Tachida Y, *et al.* Molecular cloning and genomic analysis of mouse GalNAc alpha2, 6-sialyltransferase (ST6GalNAc I). *J Biochem.* 2000;127(5):845-54.
- Kurosawa N, Inoue M, Yoshida Y, Tsuji S. Molecular cloning and genomic analysis of mouse Galbeta1, 3GalNAc-specific GalNAc alpha2,6-sialyltransferase. *J Biol Chem.* 1996;271(25):15109-16.
- Lal K, Bermeo R, Perez S. Computational tools for drawing, building and displaying carbohydrates: a visual guide. *Beilstein J Org Chem.* 2020;16:2448 68.

- Langmann T, Fauser S. [Polysialic Acid for Immunomodulation in an Animal Model for Wet Age-Related Macular Degeneration (AMD)]. *Klin Monbl Augenheilkd.* 2017;234(5):657-61.
- Laroy W, Ameloot P, Contreras R. Characterization of sialyltransferase mutants using surface plasmon resonance. *Glycobiology.* 2001;11(3):175-82.
- Lau J. Circulatory Glycan-Modifying Enzymes As Systemic Regulators In Blood Cell Development. *Blood.* 15 nov 2013;122(21):3686-3686.
- Launay R, Teppa E, Esque J, André I. Modeling Protein Complexes and Molecular Assemblies Using Computational Methods. *Methods in Molecular Biology.* 2023(2553);57-77.
- Lee JS, Yoo Y, Lim BC, Kim KJ, Song J, Choi M, *et al.* GM3 synthase deficiency due to ST3GAL5 variants in two Korean female siblings: Masquerading as Rett syndrome-like phenotype. *Am J Med Genet A.* 2016;170(8):2200-5.
- Lee MM, Nasirikenari M, Manhardt CT, Ashline DJ, Hanneman AJ, Reinhold VN, *et al.* Platelets support extracellular sialylation by supplying the sugar donor substrate. *J Biol Chem.* 2014;289(13):8742-8.
- Lee YC, Kaufmann M, Kitazume-Kawaguchi S, Kono M, Takashima S, Kurosawa N, *et al.* Molecular cloning and functional expression of two members of mouse NeuAcalpha2,3Galbeta1,3GalNAc GalNAcalpha2,6-sialyltransferase family, ST6GalNAc III and IV. *J Biol Chem.* 1999;274(17):11958-67.
- Lee YC, Kim YJ, Lee KY, Kim KS, Kim BU, Kim HN, *et al.* Cloning and expression of cDNA for a human Sia alpha 2,3Gal beta 1, 4GlcNA:alpha 2,8-sialyltransferase (hST8Sia III). *Arch Biochem Biophys.* 1998;360(1):41-6.
- Legaigneur P, Breton C, El Battari A, Guillemot JC, Auge C, Malissard M, *et al.* Exploring the acceptor substrate recognition of the human beta-galactoside alpha 2,6-sialyltransferase. *J Biol Chem.* 2001;276(24):21608-17.
- Lenman A, Liaci AM, Liu Y, Frängsmyr L, Frank M, Blaum BS, *et al.* Polysialic acid is a cellular receptor for human adenovirus 52. *Proc Natl Acad Sci USA.* 2018;115(18).
- Lepers A, Shaw L, Schneckenburger P, Cacan R, Verbert A, Schauer R. A study on the regulation of N-glycolylneuraminic acid biosynthesis and utilization in rat and mouse liver. *Eur J Biochem.* 1990;193(3):715-23.
- Lepers A, Shaw L, Cacan R, Schauer R, Montreuil J, Verbert A. Transport of CMP- N -glycolylneuraminic acid into mouse liver Golgi vesicles. *FEBS Letters.* 1989;250(2):245-50.
- Lewis AL, Desa N, Hansen EE, Knirel YA, Gordon JI, Gagneux P, *et al.* Innovations in host and microbial sialic acid biosynthesis revealed by phylogenomic prediction of nonulosonic acid structure. *Proc Natl Acad Sci USA.* 2009;106(32):13552-7.
- Li Z, Kitov PI, Kitova EN, Bui DT, Moremen KW, Wakarchuk WW, *et al.* Quantifying Carbohydrate-Active Enzyme Activity with Glycoprotein Substrates Using Electrospray Ionization Mass Spectrometry and Center-of-Mass Monitoring. *Anal Chem.* 2021;93(46):15262-70.
- Li Z, Kitov PI, Kitova EN, Mozenah F, Rodrigues E, Chapla DG, *et al.* CUPRA-ZYME: An Assay for Measuring Carbohydrate-Active Enzyme Activities, Pathways, and Substrate Specificities. *Anal Chem.* 2020;92(4):3228-36.
- Li Y, Chen X. Sialic acid metabolism and sialyltransferases: natural functions and applications. *Appl Microbiol Biotechnol.* 2012;94(4):887-905.
- Li D, Mallory T, Satomura S. AFP-L3: a new generation of tumor marker for hepatocellular carcinoma. *Clin Chim Acta.* 2001;313(1-2):15-9.
- Liao H, Winkler J, Wißfeld J, Shahraz A, Klaus C, Neumann H. Low molecular weight polysialic acid prevents lipopolysaccharide-induced inflammatory dopaminergic neurodegeneration in humanized *SIGLEC11* transgenic mice. *Glia.* 2021;69(12):2845-62.
- Liao SM, Lu B, Liu XH, Lu ZL, Liang SJ, Chen D, *et al.* Molecular Interactions of the Polysialyltransferase Domain (PSTD) in ST8Sia IV with CMP-Sialic Acid and Polysialic Acid Required for Polysialylation of the Neural Cell Adhesion Molecule Proteins: An NMR Study. *IJMS.* 2020;21(5):1590.
- Liao G, Zhou Z, Suryawanshi S, Mondal MA, Guo Z. Fully Synthetic Self-Adjuvanting α -2,9-Oligosialic Acid Based Conjugate Vaccines against Group C Meningitis. *ACS Cent Sci.* 2016;2(4):210-8.
- Lin CY, Huang Z, Wen W, Wu A, Wang C, Niu L. Enhancing Protein Expression in HEK-293 Cells by Lowering Culture Temperature. *PLoS ONE.* 2015;10(4):e0123562.

- Lin YC, Boone M, Meuris L, Lemmens I, Van Roy N, Soete A, *et al.* Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nat Commun.* 2014;5(1):4767.
- Lindhout T, Iqbal U, Willis LM, Reid AN, Li J, Liu X, *et al.* Site-specific enzymatic polysialylation of therapeutic proteins using bacterial enzymes. *Proc Natl Acad Sci U S A.* 2011;108(18):7397-402.
- Lisacek F, Tiemeyer M, Mazumder R, Aoki-Kinoshita KF. Worldwide Glycoscience Informatics Infrastructure: The GlySpace Alliance. *JACS Au.* 2022;jacsau.2c00477.
- Liu L, Prudden AR, Bosman GP, Boons G-J. Improved isolation and characterization procedure of sialylglycopeptide from egg yolk powder. *Carbohydrate Research.* 2017;452:122-8.
- Livingston BD, Paulson JC. Polymerase chain reaction cloning of a developmentally regulated member of the sialyltransferase gene family. *J Biol Chem.* 1993;268(16):11504-7.
- Livingston BD, Jacobs JL, Glick MC, Troy FA. Extended polysialic acid chains (n greater than 55) in glycoproteins from human neuroblastoma cells. *J Biol Chem.* 1988;263(19):9443-8.
- Lo NW, Lau JT. Transcription of the beta-galactoside alpha 2,6-sialyltransferase gene in B lymphocytes is directed by a separate and distinct promoter. *Glycobiology.* 1996;6(3):271-9.
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucl Acids Res.* 2014;42(D1):D490-5.
- Lopez Aguilar A, Briard JG, Yang L, Ovryn B, Macauley MS, Wu P. Tools for Studying Glycans: Recent Advances in Chemoenzymatic Glycan Labeling. *ACS Chem Biol.* 2017;12(3):611-21.
- Luley-Goedl C, Schmoelzer K, Thomann M, Malik S, Greif M, Ribitsch D, *et al.* Two N-terminally truncated variants of human β -galactoside α 2,6 sialyltransferase I with distinct properties for *in vitro* protein glycosylation. *Glycobiology.* 2016;26(10):1097-106.
- Ma H, Zhou H, Song X, Shi S, Zhang J, Jia L. Modification of sialylation is associated with multidrug resistance in human acute myeloid leukemia. *Oncogene.* 2015;34(6):726-40.
- Ma J, Simonovic M, Qian R, Colley KJ. Sialyltransferase isoforms are phosphorylated in the cis-medial Golgi on serine and threonine residues in their luminal sequences. *J Biol Chem.* 1999;274(12):8046-52.
- Ma J, Colley KJ. A disulfide-bonded dimer of the Golgi beta-galactoside alpha2,6-sialyltransferase is catalytically inactive yet still retains the ability to bind galactose. *J Biol Chem.* 1996;271(13):7758-66.
- Macauley MS, Arlian BM, Rillahan CD, Pang PC, Bortell N, Marcondes MCG, *et al.* Systemic Blockade of Sialylation in Mice with a Global Inhibitor of Sialyltransferases. *Journal of Biological Chemistry.* 2014;289(51):35149-58.
- Macqueen DJ, Johnston IA. A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proc Biol Sci.* 2014;281(1778):20132881.
- Mahal LK, Yarema KJ, Bertozzi CR. Engineering Chemical Reactivity on Cell Surfaces Through Oligosaccharide Biosynthesis. *Science.* 1997;276(5315):1125-8.
- Mandrell RE, Zollinger WD. Measurement of antibodies to meningococcal group B polysaccharide: low avidity binding and equilibrium binding constants. *J Immunol.* 1982;129(5):2172-8.
- Manhardt CT, Punch PR, Dougher CWL, Lau JTY. Extrinsic sialylation is dynamically regulated by systemic triggers *in vivo*. *Journal of Biological Chemistry.* 2017;292(33):13514-20.
- Manzi AE, Higa HH, Diaz S, Varki A. Intramolecular self-cleavage of polysialic acid. *Journal of Biological Chemistry.* 1994;269(38):23617-24.
- Mao D, Zhang C, Kenry, Liu J, Wang X, Li B, *et al.* Bio-orthogonal click reaction-enabled highly specific *in situ* cellularization of tissue engineering scaffolds. *Biomaterials.* 2020;230:119615.
- Marcos NT, Pinho S, Grandela C, Cruz A, Samyn-Petit B, Harduin-Lepers A, *et al.* Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen. *Cancer Res.* 2004;64(19):7050-7.
- Markram K, Gerardy-Schahn R, Sandi C. Selective learning and memory impairments in mice deficient for polysialylated NCAM in adulthood. *Neuroscience.* 2007;144(3):788-96.
- Martersteck CM, Kedersha NL, Drapp DA, Tsui TG, Colley KJ. Unique α 2, 8-polysialylated glycoproteins in breast cancer and leukemia cells. *Glycobiology.* 1996;6(3):289-301.

- Martin KC, Tricoli J, Corzana F, García-García A, Ceballos-Laita L, Hicks T, *et al.* Fucosyltransferase-specific inhibition via next generation of fucose mimetics. *Chem Commun.* 2021;57(9):1145-8.
- Martin NT, Wrede C, Niemann J, Brooks J, Schwarzer D, Kühnel F, *et al.* Targeting polysialic acid-abundant cancers using oncolytic adenoviruses with fibers fused to active bacteriophage borne endosialidase. *Biomaterials.* 2018;158:86-94.
- Martina JA, Daniotti JL, Maccioni HJ. Influence of N-glycosylation and N-glycan trimming on the activity and intracellular traffic of GD3 synthase. *J Biol Chem.* 1998;273(6):3725-31.
- Martinez J, Steenbergen S, Vimr E. Derived structure of the putative sialic acid transporter from *Escherichia coli* predicts a novel sugar permease domain. *J Bacteriol.* 1995;177(20):6005-10.
- Martinez-Duncker I, Salinas-Marin R, Martinez-Duncker C. Towards *In Vivo* Imaging of Cancer Sialylation. *International Journal of Molecular Imaging.* 2010;2011:1-10.
- Martinez-Duncker I, Dupré T, Piller V, Piller F, Candelier JJ, Trichet C, *et al.* Genetic complementation reveals a novel human congenital disorder of glycosylation of type II, due to inactivation of the Golgi CMP-sialic acid transporter. *Blood.* 2005;105(7):2671-6.
- Masand SN, Chen J, Perron IJ, Hammerling BC, Loers G, Schachner M, *et al.* The effect of glycomimetic functionalized collagen on peripheral nerve repair. *Biomaterials.* 2012;33(33):8353-62.
- Mbua NE, Li X, Flanagan-Steet HR, Meng L, Aoki K, Moremen KW, *et al.* Selective Exo-Enzymatic Labeling of N-Glycans on the Surface of Living Cells by Recombinant ST6Gal I. *Angew Chem Int Ed.* 2013;52(49):13012-5.
- McAuley EZ, Scimone A, Tiwari Y, Agahi G, Mowry BJ, Holliday EG, Donald JA, Weickert CS, Mitchell PB, Schofield PR, Fullerton JM. Identification of sialyltransferase 8B as a generalized susceptibility gene for psychotic and mood disorders on chromosome 15q25–26. *PLoS One.* 2012(7): e38172.
- Mendiratta SS, Sekulic N, Lavie A, Colley KJ. Specific amino acids in the first fibronectin type III repeat of the neural cell adhesion molecule play a role in its recognition and polysialylation by the polysialyltransferase ST8Sia IV/PST. *J Biol Chem.* 2005;280(37):32340-8.
- Mehta AY, Cummings RD. GlycoGlyph: a glycan visualizing, drawing and naming application. *Bioinformatics.* 2020;36(11):3613-4.
- Meng H, Jain S, Lockshin C, Shaligram U, Martinez J, Genkin D, *et al.* Clinical Application of Polysialylated Deoxyribonuclease and Erythropoietin. *DDF.* 2019;12(3):212-22.
- Meng L, Forouhar F, Thieker D, Gao Z, Ramiah A, Moniz H, *et al.* Enzymatic Basis for N-Glycan Sialylation. *Journal of Biological Chemistry.* 2013;288(48):34680-98.
- Mereiter S, Magalhães A, Adamczyk B, Jin C, Almeida A, Drici L, *et al.* Glycomic analysis of gastric carcinoma cells discloses glycans as modulators of RON receptor tyrosine kinase activation in cancer. *Biochim Biophys Acta.* 2016;1860(8):1795-808.
- Mestrom, Przypis, Kowalczykiewicz, Pollender, Kumpf, Marsden, *et al.* Leloir Glycosyltransferases in Applied Biocatalysis: A Multidisciplinary Approach. *IJMS.* 2019;20(21):5263.
- Michalides R, Kwa B, Springall D, van Zandwijk N, Koopman J, Hilkens J, *et al.* NCAM and lung cancer. *Int J Cancer Suppl.* 1994;8:34-7.
- Michon F, Brisson JR, Jennings HJ. Conformational differences between linear $\alpha(2-8)$ -linked homosialooligosaccharides and the epitope of the group B meningococcal polysaccharide. *Biochemistry.* 1987;26(25):8399-405.
- Mikolajczyk K, Kaczmarek R, Czerwinski M. How glycosylation affects glycosylation: the role of N-glycans in glycosyltransferase activity. *Glycobiology.* 2020;30(12):941-69.
- Mindler K, Ostertag E, Stehle T. The polyfunctional polysialic acid: A structural view. *Carbohydrate Research.* 2021;507:108376.
- Miyata S, Sato C, Kumita H, Toriyama M, Vacquier VD, Kitajima K. Flagelliasialin: a novel sulfated 2,9-linked polysialic acid glycoprotein of sea urchin sperm flagella. *Glycobiology.* 2006;16(12):1229-41.
- Miyata S, Sato C, Kitamura S, Toriyama M, Kitajima K. A major flagellum sialoglycoprotein in sea urchin sperm contains a novel polysialic acid, an $\alpha(2,9)$ -linked poly-N-acetylneuraminic acid chain, capped by an 8-O-sulfated sialic acid residue. *Glycobiology.* 2004;14(9):827-40.

- Moebius JM, Widera D, Schmitz J, Kaltschmidt C, Piechaczek C. Impact of polysialylated CD56 on natural killer cell cytotoxicity. *BMC Immunol.* 2007;8:13.
- Mondal N, Silva M, Castano AP, Maus MV, Sackstein R. Glycoengineering of chimeric antigen receptor (CAR) T-cells to enforce E-selectin binding. *Journal of Biological Chemistry.* 2019;294(48):18465-74.
- Montgomery AP, Skropeta D, Yu H. Transition state-based ST6Gal I inhibitors: Mimicking the phosphodiester linkage with a triazole or carbamate through an enthalpy-entropy compensation. *Sci Rep.* 2017a;7(1):14428.
- Montgomery AP, Xiao K, Wang X, Skropeta D, Yu H. Computational Glycobiology: Mechanistic Studies of Carbohydrate-Active Enzymes and Implication for Inhibitor Design. In: *Advances in Protein Chemistry and Structural Biology.* Elsevier; 2017b. p. 25-76.
- Monti E, Bonten E, D'Azzo A, Bresciani R, Venerando B, Borsani G, *et al.* Sialidases in Vertebrates. In: *Advances in Carbohydrate Chemistry and Biochemistry.* Elsevier; 2010. p. 403-79.
- Monti E, Preti A, Nesti C, Ballabio A, Borsani G. Expression of a novel human sialidase encoded by the NEU2 gene. *Glycobiology.* 1999;9(12):1313-21.
- Moons SJ, Rossing E, Janssen MACH, Heise T, Büll C, Adema GJ, *et al.* Structure–Activity Relationship of Metabolic Sialic Acid Inhibitors and Labeling Reagents. *ACS Chem Biol.* 2022;17(3):590-7.
- Moons SJ, Adema GJ, Derks MT, Boltje TJ, Büll C. Sialic acid glycoengineering using N-acetylmannosamine and sialic acid analogs. *Glycobiology.* 2019;29(6):433–445.
- Moremen KW, Haltiwanger RS. Emerging structural insights into glycosyltransferase-mediated synthesis of glycans. *Nat Chem Biol.* 2019;15(9):853-64.
- Moremen KW, Ramiah A, Stuart M, Steel J, Meng L, Forouhar F, *et al.* Expression system for structural and functional studies of human glycosylation enzymes. *Nature Chemical Biology.* 2018;14(2):156-62.
- Mori A, Yang Y, Takahashi Y, Hane M, Kitajima K, Sato C. Combinational Analyses with Multiple Methods Reveal the Existence of Several Forms of Polysialylated Neural Cell Adhesion Molecule in Mouse Developing Brains. *Int J Mol Sci.* 2020;21(16):5892.
- Mori A, Hane M, Niimi Y, Kitajima K, Sato C. Different properties of polysialic acids synthesized by the polysialyltransferases ST8SIA2 and ST8SIA4. *Glycobiology.* 2017;27(9):834-46.
- Morley TJ, Willis LM, Whitfield C, Wakarchuk WW, Withers SG. A New Sialidase Mechanism. *Journal of Biological Chemistry.* 2009;284(26):17404-10.
- Mühlenhoff M, Rollenhagen M, Werneburg S, Gerardy-Schahn R, Hildebrandt H. Polysialic Acid: Versatile Modification of NCAM, SynCAM1 and Neuropilin-2. *Neurochem Res.* 2013;38(6):1134-43.
- Mühlenhoff M, Stummeyer K, Grove M, Sauerborn M, Gerardy-Schahn R. Proteolytic Processing and Oligomerization of Bacteriophage-derived Endosialidases. *Journal of Biological Chemistry.* 2003;278(15):12634-44.
- Mühlenhoff M, Manegold A, Windfuhr M, Gotza B, Gerardy-Schahn R. The Impact of N-Glycosylation on the Functions of Polysialyltransferases. *Journal of Biological Chemistry.* 2001;276(36):34066-73.
- Mühlenhoff M, Eckhardt M, Bethe A, Frosch M, Gerardy-Schahn R. Autocatalytic polysialylation of polysialyltransferase-1. *The EMBO Journal.* 1996b;15(24):6943-50.
- Mühlenhoff M, Eckhardt M, Bethe A, Frosch M, Gerardy-Schahn R. Polysialylation of NCAM by a single enzyme. *Current Biology.* 1996a;6(9):1188-91.
- Munkley J. Aberrant Sialylation in Cancer: Therapeutic Opportunities. *Cancers.* 2022;14(17):4248.
- Munkley J, Scott E. Targeting Aberrant Sialylation to Treat Cancer. *Medicines.* 2019;6(4):102.
- Munkley J. The Role of Sialyl-Tn in Cancer. *Int J Mol Sci.* 2016;17(3):275.
- Münster-Kühnel AK, Tiralongo J, Krapp S, Weinhold B, Ritz-Sedlacek V, Jacob U, Gerardy-Schahn R. Structure and function of vertebrate CMP-sialic acid synthetases. *Glycobiology.* 2004;14(10):43R-51R.
- Nadano D, Iwasaki M, Endo S, Kitajima K, Inoue S, Inoue Y. A naturally occurring deaminated neuraminic acid, 3-deoxy-D-glycero-D-galacto-nonulosonic acid (KDN). Its unique occurrence at the nonreducing ends of oligosialyl chains in polysialoglycoprotein of rainbow trout eggs. *Journal of Biological Chemistry.* 1986;261(25):11550-7.

- Nagae M, Ikeda A, Hane M, Hanashima S, Kitajima K, Sato C, *et al.* Crystal Structure of Anti-polysialic Acid Antibody Single Chain Fv Fragment Complexed with Octasialic Acid. *Journal of Biological Chemistry*. 2013;288(47):33784-96.
- Nakamura M, Tsunoda A, Saito M. Radioimmune assay of sialyltransferase and N-acetylgalactosaminyltransferase activities using specific antibodies on a 96-well filtration plate of a multiscreen assay system. *Analytical Biochemistry*. 1991;198(1):154-9.
- Nakata D, Zhang L, Troy FA. Molecular basis for polysialylation: A novel polybasic polysialyltransferase domain (PSTD) of 32 amino acids unique to the α 2,8-polysialyltransferases is essential for polysialylation. *Glycoconj J*. 2006;23(5-6):423-36.
- Nakata D, Troy FA. Degree of polymerization (DP) of polysialic acid (polySia) on neural cell adhesion molecules (N-CAMS): development and application of a new strategy to accurately determine the DP of polySia chains on N-CAMS. *J Biol Chem*. 2005;280(46):38305-16.
- Nakata D, Munster AK, Gerardy-Schahn R, Aoki N, Matsuda T, Kitajima K. Molecular cloning of a unique CMP-sialic acid synthetase that effectively utilizes both deaminoneuraminic acid (KDN) and N-acetylneuraminic acid (Neu5Ac) as substrates. *Glycobiology*. 2001;11(8):685-92.
- Nakayama J, Fukuda MN, Fredette B, Ranscht B, Fukuda M. Expression cloning of a human polysialyltransferase that forms the polysialylated neural cell adhesion molecule present in embryonic brain. *Proc Natl Acad Sci U S A*. 1995;92(15):7031-5.
- Nara K, Watanabe Y, Maruyama K, Kasahara K, Nagai Y, Sanai Y. Expression cloning of a CMP-NeuAc:NeuAc alpha 2-3Gal beta 1-4Glc beta 1-1'Cer alpha 2,8-sialyltransferase (GD3 synthase) from human melanoma cells. *Proc Natl Acad Sci U S A*. 1994;91(17):7952-6.
- Nishimura S, Hane M, Niimi Y, Miyata S, Kitajima K, Sato C. Comparison of Analytical Methods to Detect Polysialic Acid. *Journal of Glycomics & Lipidomics*. 2014 ; 4:2.
- Nizet V, Esko JD. *Bacterial and Viral Infections. Essentials of Glycobiology 2nd ed.* Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009.
- Noel M. Etude des relations structure-fonction des sialyltransférases humaines : développement de nouveaux outils. Doctoral thesis. University of Lille (FR). 2018.
- Noel M, Gilormini P-A, Coge V, Lion C, Biot C, Harduin-Lepers A, *et al.* MicroPlate Sialyltransferase Assay: A Rapid and Sensitive Assay Based on an Unnatural Sialic Acid Donor and Bioorthogonal Chemistry. *Bioconjugate Chemistry*. 2018;29(10):3377-84.
- Noel M, Gilormini P-A, Coge V, Yamakawa N, Vicogne D, Lion C, *et al.* Probing the CMP-Sialic Acid Donor Specificity of Two Human β -D-Galactoside Sialyltransferases (ST3Gal I and ST6Gal I) Selectively Acting on O- and N-Glycosylproteins. *ChemBioChem*. 2017;18(13):1251-9.
- Nohara K, Ozawa H, Tai T, Saji H, Hidekazu Fujimaki. Gangliosides involved in activation of rat T lineage cells. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism*. 1997;1345(2):207-14.
- Ohta K, Sato C, Matsuda T, Toriyama M, Vacquier VD, Kitajima K. Co-localization of receptor and transducer proteins in the glycosphingolipid-enriched, low density, detergent-insoluble membrane fraction of sea urchin sperm. *Glycoconj. Journal*. 2000;17:205-14.
- Okajima T, Fukumoto S, Ito H, Kiso M, Hirabayashi Y, Urano T, *et al.* Molecular cloning of brain-specific GD1alpha synthase (ST6GalNAc V) containing CAG/Glutamine repeats. *J Biol Chem*. 1999b;274(43):30557-62.
- Okajima T, Fukumoto S, Miyazaki H, Ishida H, Kiso M, Furukawa K, *et al.* Molecular cloning of a novel alpha2,3-sialyltransferase (ST3Gal VI) that sialylates type II lactosamine structures on glycoproteins and glycolipids. *J Biol Chem*. 1999a;274(17):11479-86.
- Oltmann-Norden I, Galuska SP, Hildebrandt H, Geyer R, Gerardy-Schahn R, Geyer H, *et al.* Impact of the polysialyltransferases ST8SiaII and ST8SiaIV on polysialic acid synthesis during postnatal mouse brain development. *J Biol Chem*. 2008;283(3):1463-71.
- Ong E, Nakayama J, Angata K, Reyes L, Katsuyama T, Arai Y, *et al.* Developmental regulation of polysialic acid synthesis in mouse directed by two polysialyltransferases, PST and STX. *Glycobiology*. avr 1998;8(4):415-24.

- Ono S, Hane M, Kitajima K, Sato C. Novel regulation of fibroblast growth factor 2 (FGF2)-mediated cell growth by polysialic acid. *J Biol Chem*. 2012;287(6):3710-22.
- Papayannopoulos V. Neutrophil extracellular traps in immunity and disease. *Nat Rev Immunol*. 2018;18(2):134-47.
- Park JJ, Lee M. Increasing the α 2, 6 sialylation of glycoproteins may contribute to metastatic spread and therapeutic resistance in colorectal cancer. *Gut Liver*. 2013;7(6):629-41.
- Parodi AJ, Caramelo JJ, D'Alessio C. UDP-Glucose: Glycoprotein Glucosyltransferase 1,2 (UGGT1,2). In *Handbook of Glycosyltransferases and Related Genes*. Tokyo: Springer Japan; 2014 p.15-30.
- Patel RY, Balaji PV. Identification of linkage-specific sequence motifs in sialyltransferases. *Glycobiology*. 2006;16(2):108-16.
- Patil SA, Bshara W, Morrison C, Chandrasekaran EV, Matta KL, Neelamegham S. Overexpression of α 2,3sialyl T-antigen in breast cancer determined by miniaturized glycosyltransferase assays and confirmed using tissue microarray immunohistochemical analysis. *Glycoconj J*. 2014;31(6-7):509-21.
- Patnaik SK, Stanley P. Lectin-Resistant CHO Glycosylation Mutants. *Methods Enzymol*. 2006;416:159-82.
- Paulson JC, Rademacher C. Glycan terminator. *Nat Struct Mol Biol*. 2009;16(11):1121-2.
- Paulson JC, Colley KJ. Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation. *J Biol Chem*. 1989;264(30):17615-8.
- Paulson JC, Rearick JI, Hill RL. Enzymatic properties of beta-D-galactoside alpha2 leads to 6 sialyltransferase from bovine colostrum. *J Biol Chem*. 1977;252(7):2363-71.
- Péanne R, de Lonlay P, Foulquier F, Kornak U, Lefeber DJ, Morava E, *et al*. Congenital disorders of glycosylation (CDG): Quo vadis? *Eur J Med Genet*. 2018;61(11):643-63.
- Peng LX, Liu XH, Lu B, Liao SM, Zhou F, Huang JM, *et al*. The Inhibition of Polysialyltransferase ST8SiaIV Through Heparin Binding to Polysialyltransferase Domain (PSTD). *Medicinal Chemistry*. 2019;15(5):486-95.
- Perez SJLP, Fu CW, Li WS. Sialyltransferase Inhibitors for the Treatment of Cancer Metastasis: Current Challenges and Future Perspectives. *Molecules*. 2021;26(18):5673.
- Petit D, Teppa RE, Harduin-Lepers A. A phylogenetic view and functional annotation of the animal β 1,3-glycosyltransferases of the GT31 CAZy family. *Glycobiology*. 2021;31(3):243-59.
- Petit D, Teppa E, Cenci U, Ball S, Harduin-Lepers A. Reconstruction of the sialylation pathway in the ancestor of eukaryotes. *Sci Rep*. 2018;8(1):2946.
- Petit D, Teppa RE, Petit J-M, Harduin-Lepers A. A Practical Approach to Reconstruct Evolutionary History of Animal Sialyltransferases and Gain Insights into the Sequence-Function Relationships of Golgi-Glycosyltransferases. Inka Brockhausen (ed.), *Glycosyltransferase, Methods in Molecular Biology*. 2013;1022 :73-97.
- Petit D, Mir A-M, Petit J-M, Thisse C, Delannoy P, Oriol R, *et al*. Molecular Phylogeny and Functional Genomics of β -Galactoside α 2,6-Sialyltransferases That Explain Ubiquitous Expression of *st6gal1* Gene in Amniotes. *J Biol Chem*. 2010;285(49):38399-414.
- Pertti Aula, MD, Seppo Autio MD, Kari O. Raivio MD, Juhani Rapola MD, Carl-Johan Thodén MD, Sirkka-Liisa Koskela MD, Ikuo Yamashina. 'Salla Disease': A New Lysosomal Storage Disorder. *Arch Neurol*. 1979;36(2):88-94.
- Pratama F, Linton D, Dixon N. Genetic and process engineering strategies for enhanced recombinant N-glycoprotein production in bacteria. *Microb Cell Fact*. 2021;20(1):198.
- Pröpster JM, Yang F, Rabbani S, Ernst B, Allain FHT, Schubert M. Structural basis for sulfation-dependent self-glycan recognition by the human immune-inhibitory receptor Siglec-8. *Proc Natl Acad Sci USA*. 2016;113(29).
- Qian R, Chen C, Colley KJ. Location and mechanism of alpha 2,6-sialyltransferase dimer formation. Role of cysteine residues in enzyme dimerization, localization, activity, and processing. *J Biol Chem*. 2001;276(31):28641-9.
- Rakic B, Rao FV, Freimann K, Wakarchuk W, Strynadka NCJ, Withers SG. Structure-based mutagenic analysis of mechanism and substrate specificity in mammalian glycosyltransferases: porcine ST3Gal-I. *Glycobiology*. 2013;23(5):536-45.

- Rao FV, Rich JR, Rakić B, Buddai S, Schwartz MF, Johnson K, *et al.* Structural insight into mammalian sialyltransferases. *Nat Struct Mol Biol.* 2009;16(11):1186-8.
- Rebl A, Verleih M, Nipkow M, Altmann S, Bochert R, Goldammer T. Gradual and Acute Temperature Rise Induces Crossing Endocrine, Metabolic, and Immunological Pathways in Maraena Whitefish (*Coregonus maraena*). *Front Genet.* 2018;9:241.
- Reily C, Stewart TJ, Renfrow MB, Novak J. Glycosylation in health and disease. *Nat Rev Nephrol.* 2019;15(6):346-66.
- Rewar S, Mirdha D, Rewar P. Treatment and Prevention of Pandemic H1N1 *Influenza*. *Annals of Global Health.* 2016;81(5):645.
- Rey-Gallardo A, Escribano C, Delgado-Martín C, Rodríguez-Fernández JL, Gerardy-Schahn R, Rutishauser U, *et al.* Polysialylated neuropilin-2 enhances human dendritic cell migration through the basic C-terminal region of CCL21. *Glycobiology.* 2010;20(9):1139-46.
- Reynders E, Foulquier F, Annaert W, Matthijs G. How Golgi glycosylation meets and needs trafficking: the case of the COG complex. *Glycobiology.* 2011;21(7):853-63.
- Richard E. Synthèse par ingénierie métabolique d'oligosaccharides sialylés pour l'élaboration de glycoconjugués d'intérêt médical. Doctoral thesis. University of Grenoble (FR). 2017.
- Rigolot V, Rossez Y, Biot C, Lion C. A bioorthogonal chemistry approach to detect the K1 polysialic acid capsule in *Escherichia coli*. *RSC Chem. Biol.*, 2023, Advance Article.
- Rillahan CD, Antonopoulos A, Lefort CT, Sonon R, Azadi P, Ley K, *et al.* Global metabolic inhibitors of sialyl- and fucosyltransferases remodel the glycome. *Nat Chem Biol.* 2012;8(7):661-8.
- Rini J, Esko J, Varki A. Glycosyltransferases and Glycan-processing Enzymes. *Essentials of Glycobiology* Cold Spring Harbor Laboratory Press; 2009.
- Rivinoja A, Pujol FM, Hassinen A, Kellokumpu S. Golgi pH, its regulation and roles in human disease. *Ann Med.* 2012;44(6):542-54.
- Rivinoja A, Hassinen A, Kokkonen N, Kauppila A, Kellokumpu S. Elevated Golgi pH impairs terminal N-glycosylation by inducing mislocalization of Golgi glycosyltransferases. *J Cell Physiol.* 2009;220(1):144-54.
- Robinson NE, de Vries T, Davis RE, Stults CL, Watson SR, van den Eijnden DH, *et al.* Expression of fucosylated antigens and alpha 1,3 fucosyltransferases in human leukaemia cell lines. *Glycobiology.* 1994;4(3):317-26.
- Rodrigues E, Macauley MS. Hypersialylation in Cancer: Modulation of Inflammation and Therapeutic Opportunities. *Cancers (Basel).* 2018;10(6):207.
- Rogers GN, Herrler G, Paulson JC, Klenk HD. Influenza C virus uses 9-O-acetyl-N-acetylneuraminic acid as a high affinity receptor determinant for attachment to cells. *Journal of Biological Chemistry.* 1986;261(13):5947-51.
- Rohfritsch PF, Joosten JAF, Krzewinski-Recchi M-A, Harduin-Lepers A, Laporte B, Juliant S, *et al.* Probing the substrate specificity of four different sialyltransferases using synthetic β -D-Galp-(1 \rightarrow 4)- β -D-GlcpNAc-(1 \rightarrow 2)- α -D-Manp-(1 \rightarrow O)(CH₂)₇CH₃ analogues general activating effect of replacing *N*-acetylglucosamine by *N*-propionylglucosamine. *Biochim Biophys Acta.* 2006;1760(4):685-92.
- Rojas Á, Sánchez-Torrijos Y, Gil-Gómez A, Liu CH, Rodríguez-Rivas C, Ferrer MT, *et al.* Performance of different biomarkers for the management of hepatocellular carcinoma. *HR.* 6 2018;4(7):31.
- Rollenhagen M, Buettner FFR, Reismann M, Jirmo AC, Grove M, Behrens GMN, *et al.* Polysialic Acid on Neuropilin-2 Is Exclusively Synthesized by the Polysialyltransferase ST8Sia IV and Attached to Mucin-type O-Glycans Located between the b2 and c Domain. *Journal of Biological Chemistry.* 2013;288(32):22880-92.
- Rollenhagen M, Kuckuck S, Ulm C, Hartmann M, Galuska SP, Geyer R, *et al.* Polysialylation of the synaptic cell adhesion molecule 1 (SynCAM 1) depends exclusively on the polysialyltransferase ST8SiaII *in vivo*. *J Biol Chem.* 2012;287(42):35170-80.
- Rostovtsev VV, Green LG, Fokin VV, Sharpless KB. A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew Chem Int Ed Engl.* 2002;41(14), 2596-9.
- Rudman N, Gornik O, Lauc G. Altered N-glycosylation profiles as potential biomarkers and drug targets in diabetes. *FEBS Lett.* 2019;593(13):1598-615.

- Rutishauser U. Polysialic acid in the plasticity of the developing and adult vertebrate nervous system. *Nat Rev Neurosci*. 2008;9(1):26-35.
- Sackstein R. Engineering cellular trafficking *via* glycosyltransferase-programmed stereosubstitution: Navigating cell migration *via* GPS. *Ann N Y Acad Sci*. 2012;1253(1):193-200.
- Saiki I, Koike C, Obata A, Fujii H, Murata J, Kiso M, *et al*. Functional role of sialyl Lewis X and fibronectin-derived RGDS peptide analogue on tumor-cell arrest in lungs followed by extravasation. *Int J Cancer*. 1996;65(6):833-9.
- Samraj AN, Läubli H, Varki N, Varki A. Involvement of a non-human sialic Acid in human cancer. *Front Oncol*. 2014;4:33.
- Samyn-Petit B, Krzewinski-Recchi MA, Steelant WF, Delannoy P, Harduin-Lepers A. Molecular cloning and functional expression of human ST6GalNAc II. Molecular expression in various human cultured cells. *Biochim Biophys Acta*. 2000;1474(2):201-11.
- Sasaki K, Kurata K, Kojima N, Kurosawa N, Ohta S, Hanai N, *et al*. Expression cloning of a GM3-specific alpha-2,8-sialyltransferase (GD3 synthase). *J Biol Chem*. 1994;269(22):15950-6.
- Sato C, Kitajima K. Polysialylation and disease. *Molecular Aspects of Medicine*. 2021;79:100892.
- Sato C, Hane M. Mental disorders and an acidic glycan-from the perspective of polysialic acid (PSA/polySia) and the synthesizing enzyme, ST8SIA2. *Glycoconj J*. 2018;35(4):353-73.
- Sato C, Kitajima K. Disialic, oligosialic and polysialic acids: distribution, functions and related disease. *Journal of Biochemistry*. 2013;154(2):115-36.
- Sato C. Chain Length Diversity of Sialic Acids and Its Biological Significance. *Trends in Glycoscience and Glycotechnology*. 2004;16(91):331-44.
- Sato C, Matsuda T, Kitajima K. Neuronal Differentiation-dependent Expression of the Disialic Acid Epitope on CD166 and Its Involvement in Neurite Formation in Neuro2A Cells. *Journal of Biological Chemistry*. 2002;277(47):45299-305.
- Sato C, Fukuoka H, Ohta K, Matsuda T, Koshino R, Kobayashi K, *et al*. Frequent Occurrence of Pre-existing $\alpha 2 \rightarrow 8$ -Linked Disialic and Oligosialic Acids with Chain Lengths Up to 7 Sia Residues in Mammalian Brain Glycoproteins. *Journal of Biological Chemistry*. 2000;275(20):15422-31.
- Sato C, Kitajima K, Inoue S, Inoue Y. Identification of Oligo-N-glycolylneuraminic Acid Residues in Mammal-derived Glycoproteins by a Newly Developed Immunochemical Reagent and Biochemical Methods. *Journal of Biological Chemistry*. 1998;273(5):2575-82.
- Sato C, Kitajima K, Tazawa I, Inoue Y, Inoue S, Troy F. Structural diversity in the $\alpha 2 \rightarrow 8$ -linked polysialic acid chains in salmonid fish egg glycoproteins. Occurrence of poly(Neu5Ac), poly(Neu5Gc), poly(Neu5Ac, Neu5Gc), poly(KDN), and their partially acetylated forms. *J Biol Chem*. 1993;268(31):23675-84.
- Scache J, Rigolot V, Lion C, Mortuaire M, Lefebvre T, Biot C, *et al*. Switching azide and alkyne tags on bioorthogonal reporters in metabolic labeling of sialylatedglycoconjugates: a comparative study. *Sci Rep*. 2022;12(1):22129.
- Schäkel K, Kannagi R, Kniep B, Goto Y, Mitsuoaka C, Zwirner J, *et al*. 6-Sulfo LacNAc, a Novel Carbohydrate Modification of PSGL-1, Defines an Inflammatory Type of Human Dendritic Cells. *Immunity*. 2002;17(3):289-301.
- Schauer R, Kamerling JP. Chapter One - Exploration of the Sialic Acid World. In: Baker DC, editor. *Advances in Carbohydrate Chemistry and Biochemistry*. Academic Press; 2018. p. 1-213.
- Schauer R, Kamerling JP. Chemistry, biochemistry and biology of sialic acids. In: *New Comprehensive Biochemistry*. Elsevier; 1997. p. 243-402.
- Scheidegger EP, Sternberg LR, Roth J, Lowe JB. A human STX cDNA confers polysialic acid expression in mammalian cells. *J Biol Chem*. 1995;270(39):22685-8.
- Schjoldager KT, Narimatsu Y, Joshi HJ, Clausen H. Global view of human protein glycosylation pathways and functions. *Nat Rev Mol Cell Biol*. 2020;21(12):729-49.
- Schnaar RL. Glycobiology simplified: diverse roles of glycan recognition in inflammation. *Journal of Leukocyte Biology*. 2016;99(6):825-38.
- Schnaar RL, Gerardy-Schahn R, Hildebrandt H. Sialic Acids in the Brain: Gangliosides and Polysialic Acid in Nervous System Development, Stability, Disease, and Regeneration. *Physiol Rev*. 2014;94(2):461-518.

- Schultz MJ, Holdbrooks AT, Chakraborty A, Grizzle WE, Landen CN, Buchsbaum DJ, *et al.* The Tumor-Associated Glycosyltransferase ST6Gal-I Regulates Stem Cell Transcription Factors and Confers a Cancer Stem Cell Phenotype. *Cancer Res.* 1 juill 2016;76(13):3978-88.
- Schultz MJ, Swindall AF, Wright JW, Sztul ES, Landen CN, Bellis SL. ST6Gal-I sialyltransferase confers cisplatin resistance in ovarian tumor cells. *J Ovarian Res.* 2013;6(1):25.
- Schulz EC, Schwarzer D, Frank M, Stummeyer K, Mühlenhoff M, Dickmanns A, *et al.* Structural Basis for the Recognition and Cleavage of Polysialic Acid by the Bacteriophage K1F Tailspike Protein EndoNF. *Journal of Molecular Biology.* 2010;397(1):341-51.
- Schwarzer D, Stummeyer K, Haselhorst T, Freiberger F, Rode B, Grove M, *et al.* Proteolytic Release of the Intramolecular Chaperone Domain Confers Processivity to Endosialidase F. *Journal of Biological Chemistry.* 2009;284(14):9465-74.
- Seki T, Arai Y. The persistent expression of a highly polysialylated NCAM in the dentate gyrus of the adult rat. *Neuroscience Research.* 1991;12(4):503-13.
- Seidenfaden R, Krauter A, Schertzinger F, Gerardy-Schahn R, Hildebrandt H. Polysialic acid directs tumor cell growth by controlling heterophilic neural cell adhesion molecule interactions. *Mol Cell Biol.* 2003;23(16):5908-18.
- Senda M, Ito A, Tsuchida A, Hagiwara T, Kaneda T, Nakamura Y, *et al.* Identification and expression of a sialyltransferase responsible for the synthesis of disialylgalactosylgloboside in normal and malignant kidney cells: downregulation of ST6GalNAc VI in renal cancers. *Biochem J.* 2007;402(3):459-70.
- Séveno M, Bardor M, Paccalet T, Gomord V, Lerouge P, Faye L. Glycoprotein sialylation in plants? *Nat Biotechnol.* 2004;22(11):1351-2; 1352-1353.
- Sevigny MB, Ye J, Kitazume-Kawaguchi S, Troy FA. Developmental expression and characterization of the alpha2,8-polysialyltransferase activity in embryonic chick brain. *Glycobiology.* 1998;8(9):857-67.
- Sewell R, Bäckström M, Dalziel M, Gschmeissner S, Karlsson H, Noll T, *et al.* The ST6GalNAc-I sialyltransferase localizes throughout the Golgi and is responsible for the synthesis of the tumor-associated sialyl-Tn O-glycan in human breast cancer. *J Biol Chem.* 2006;281(6):3586-94.
- Shah MM, Fujiyama K, Flynn CR, Joshi L. Sialylated endogenous glycoconjugates in plant cells. *Nat Biotechnol.* 2003;21(12):1470-1.
- Shahraz A, Lin Y, Mbroh J, Winkler J, Liao H, Lackmann M, *et al.* Low molecular weight polysialic acid binds to properdin and reduces the activity of the alternative complement pathway. *Sci Rep.* 2022;12(1):5818.
- Shahraz A, Kopatz J, Mathy R, Kappler J, Winter D, Kapoor S, *et al.* Anti-inflammatory activity of low molecular weight polysialic acid on human macrophages. *Sci Rep.* 2015;5:16800.
- Shastri DG, Irudayanathan FJ, Williams A, Koffas M, Linhardt RJ, Nangia S, *et al.* Rational identification and characterisation of peptide ligands for targeting polysialic acid. *Sci Rep.* 2020;10(1):7697.
- Shaw L, Schauer R. The biosynthesis of N-glycolylneuraminic acid occurs by hydroxylation of the CMP-glycoside of N-acetylneuraminic acid. *Biol Chem Hoppe Seyler.* juin 1988;369(6):477-86.
- Siddiqui KS, Cavicchioli R. Cold-Adapted Enzymes. *Annu Rev Biochem.* 2006;75(1):403-33.
- Sjoberg ER, Kitagawa H, Glushka J, van Halbeek H, Paulson JC. Molecular cloning of a developmentally regulated N-acetylgalactosamine alpha2,6-sialyltransferase specific for sialylated glycoconjugates. *J Biol Chem.* 1996;271(13):7450-9.
- Skropeta D, Schwörer R, Haag T, Schmidt RR. Asymmetric synthesis and affinity of potent sialyltransferase inhibitors based on transition-state analogues. *Glycoconj J.* 2004;21(5):205-19.
- Song E, Mechref Y. Defining glycoprotein cancer biomarkers by MS in conjunction with glycoprotein enrichment. *Biomarkers in Medicine.* 2015;9(9):835-44.
- Soukhtehzari S, Berish RB, Fazli L, Watson PH, Williams KC. The different prognostic significance of polysialic acid and CD56 expression in tumor cells and lymphocytes identified in breast cancer. *NPJ Breast Cancer.* 2022;8(1):78.
- Stamatos NM, Zhang L, Jokilampi A, Finne J, Chen WH, El-Maarouf A, *et al.* Changes in polysialic acid expression on myeloid cells during differentiation and recruitment to sites of inflammation: role in phagocytosis. *Glycobiology.* 2014;24(9):864-79.

- Storms SD, Rutishauser U. A Role for Polysialic Acid in Neural Cell Adhesion Molecule Heterophilic Binding to Proteoglycans. *Journal of Biological Chemistry*. 1998;273(42):27124-9.
- Strehle EM. Sialic Acid Storage Disease and Related Disorders. *Genetic Testing*. 2003;7(2):113-21.
- Stummeyer K, Dickmanns A, Mühlenhoff M, Gerardy-Schahn R, Ficner R. Crystal structure of the polysialic acid-degrading endosialidase of bacteriophage K1F. *Nat Struct Mol Biol*. 2005;12(1):90-6.
- Sugimoto I, Futakawa S, Oka R, Ogawa K, Marth JD, Miyoshi E, *et al*. Beta-galactoside alpha2,6-sialyltransferase I cleavage by BACE1 enhances the sialylation of soluble glycoproteins. A novel regulatory mechanism for alpha2,6-sialylation. *J Biol Chem*. 2007;282(48):34896-903.
- Sumida M, Hane M, Yabe U, Shimoda Y, Pearce OMT, Kiso M, *et al*. Rapid Trimming of Cell Surface Polysialic Acid (PolySia) by Exovesicular Sialidase Triggers Release of Preexisting Surface Neurotrophin. *J Biol Chem*. 2015;290(21):13202-14.
- Sun T, Yu S-H, Zhao P, Meng L, Moremen KW, Wells L, *et al*. One-Step Selective Exoenzymatic Labeling (SEEL) Strategy for the Biotinylation and Identification of Glycoproteins of Living Cells. *J Am Chem Soc*. 2016;138(36):11575-82.
- Suzuki M, Suzuki M, Nakayama J, Suzuki A, Angata K, Chen S, *et al*. Polysialic acid facilitates tumor invasion by glioma cells. *Glycobiology*. 2005;15(9):887-94.
- Svensson EC, Conley PB, Paulson JC. Regulated expression of alpha 2,6-sialyltransferase by the liver-enriched transcription factors HNF-1, DBP, and LAP. *J Biol Chem*. 1992;267(5):3466-72.
- Swiech K, de Freitas MCC, Covas DT, Picanço-Castro V. Recombinant Glycoprotein Production in Human Cell Lines. *Methods in Molecular Biology*. 2015(1258):223-40.
- Swindall AF, Londoño-Joshi AI, Schultz MJ, Fineberg N, Buchsbaum DJ, Bellis SL. ST6Gal-I protein expression is upregulated in human epithelial tumors and correlates with stem cell markers in normal tissues and colon cancer cell lines. *Cancer Res*. 2013;73(7):2368-78.
- Szabo R, Skropeta D. Advancement of Sialyltransferase Inhibitors: Therapeutic Challenges and Opportunities: sialyltransferase inhibitors. *Med Res Rev*. 2017;37(2):219-70.
- Taguchi T, Seko A, Kitajima K, Muto Y, Inoue S, Khoo K-H, Morris H, Dell A, Inoue Y. Structural studies of a novel type of pentaantennary large glycan unit in the fertilization-associated carbohydrate-rich glycopeptide isolated from the fertilized eggs of *Oryzias latipes*. *Journal of Biological Chemistry*. 1994;269(12):8762-71.
- Taguchi T, Seko A, Kitajima K, Inoue S, Iwamatsu T, Khoo KH, Morris HR, Dell A, Inoue Y. Structural Studies of a Novel Type of Tetraantennary Sialoglycan Unit in a Carbohydrate-rich Glycopeptide Isolated from the Fertilized Eggs of Indian MedakaFish, *Oryzias melastigma**. *Journal of Biological Chemistry*. 1993;268(4):2353-62.
- Tajik A, Phillips KL, Nitz M, Willis LM. A new ELISA assay demonstrates sex differences in the concentration of serum polysialic acid. *Analytical Biochemistry*. 2020;600:113743.
- Takashima S, Ishida H, Inazu T, Ando T, Ishida H, Kiso M, *et al*. Molecular Cloning and Expression of a Sixth Type of α 2,8-Sialyltransferase (ST8Sia VI) That Sialylates O-Glycans. *J Biol Chem*. 2002;277(27):24030-8.
- Tanaka F, Otake Y, Nakagawa T, Kawano Y, Miyahara R, Li M, *et al*. Prognostic significance of polysialic acid expression in resected non-small cell lung cancer. *Cancer Res*. 2001;61(4):1666-70.
- Tang H, Partyka K, Hsueh P, Sinha JY, Kletter D, Zeh H, *et al*. Glycans Related to the CA19-9 Antigen Are Increased in Distinct Subsets of Pancreatic Cancers and Improve Diagnostic Accuracy Over CA19-9. *Cellular and Molecular Gastroenterology and Hepatology*. 2016;2(2):210-221.e15.
- Taniguchi A, Kaneta R, Morishita K, Matsumoto K. Gene structure and transcriptional regulation of human Gal beta1,4(3) GlcNAc alpha2,3-sialyltransferase VI (hST3Gal VI) gene in prostate cancer cell line. *Biochem Biophys Res Commun*. 2001;287(5):1148-56.
- Taniguchi A, Hasegawa Y, Higai K, Matsumoto K. Transcriptional regulation of human beta-galactoside alpha2, 6-sialyltransferase (hST6Gal I) gene during differentiation of the HL-60 cell line. *Glycobiology*. 2000;10(6):623-8.
- Tao R, Li C, Zheng Y, Qin W, Zhang J, Li X, Xu Y, Shi YY, Feng G, He L. Positive association between SIAT8B and schizophrenia in the Chinese Han population. *Schizophr Res*. 2007(90): 108–114.

- Taujale R, Zhou Z, Yeung W, Moremen KW, Li S, Kannan N. Mapping the glycosyltransferase fold landscape using interpretable deep learning. *Nat Commun.* 2021;12(1):5656.
- Taujale R, Venkat A, Huang LC, Zhou Z, Yeung W, Rasheed KM, *et al.* Deep evolutionary analysis reveals the design principles of fold A glycosyltransferases. *eLife.* 2020;9:e54532.
- Taylor-Papadimitriou J, Burchell J, Miles DW, Dalziel M. MUC1 and cancer. *Biochim Biophys Acta.* 1999;1455(2-3):301-13.
- Teinturier-Lelièvre M. Etude de la biosynthèse des motifs di-, oligo- et polysialylés chez les mammifères : Identification et caractérisation d'une nouvelle sialyltransférase humaine (hST8Sia VI) responsable de la biosynthèse de motifs diSia sur des O-glycosylprotéines. Doctoral thesis. University of Lille (FR). 2006.
- Teinturier-Lelièvre M, Julien S, Juliant S, Guerardel Y, Duonor-Cérutti M, Delannoy P, *et al.* Molecular cloning and expression of a human hST8Sia VI (α 2,8-sialyltransferase) responsible for the synthesis of the diSia motif on O-glycosylproteins. *Biochem J.* 2005;392(3):665-74.
- Teppa RE, Petit D, Plechakova O, Cogez V, Harduin-Lepers A. Phylogenetic-Derived Insights into the Evolution of Sialylation in Eukaryotes: Comprehensive Analysis of Vertebrate β -Galactoside α 2,3/6-Sialyltransferases (ST3Gal and ST6Gal). *Int J Mol Sci.* 2016;17(8):1286.
- Thiesler H, Kücükerden M, Gretenkort L, Röckle I, Hildebrandt H. News and Views on Polysialic Acid: From Tumor Progression and Brain Development to Psychiatric Disorders, Neurodegeneration, Myelin Repair and Immunomodulation. *Front Cell Dev Biol.* 2022;10:871757.
- Thiesler H, Beimdiek J, Hildebrandt H. Polysialic acid and Siglec-E orchestrate negative feedback regulation of microglia activation. *Cell Mol Life Sci.* 2021;78(4):1637-53.
- Tiralongo J. Introduction to Sialic Acid Structure, Occurrence, Biosynthesis and Function. Tiralongo J, Martinez-Duncker I, editors. In: *Sialobiology: Structure, Biosynthesis and Function Sialic Acid Glycoconjugates in Health and Disease.* Bentham Science Publishers. 2013 ; 3-32.
- Tortorici MA, Walls AC, Lang Y, Wang C, Li Z, Koerhuis D, *et al.* Structural basis for human coronavirus attachment to sialic acid receptors. *Nat Struct Mol Biol.* 2019;26(6):481-9.
- Tornøe CW, Christensen C, Meldal M. Peptidotriazoles on solid phase: [1,2,3]-triazoles by regioselective copper(i)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J Org Chem.* 2002;67(9), 3057-3064.
- Trinchera M, Aronica A, Dall'Olio F. Selectin Ligands Sialyl-Lewis a and Sialyl-Lewis x in Gastrointestinal Cancers. *Biology (Basel).* 2017;6(1):16.
- Trombetta ES. The contribution of N-glycans and their processing in the endoplasmic reticulum to glycoprotein biosynthesis. *Glycobiology.* 2003;13(9):77-91.
- Troy FA. Polysialylation: from bacteria to brains. *Glycobiology.* 1992;2(1):5-23.
- Tsang KY, Fantini M, Zaki A, Mavroukakis SA, Morelli MP, Annunziata CM, *et al.* Identification of the O-Glycan Epitope Targeted by the Anti-Human Carcinoma Monoclonal Antibody (mAb) NEO-201. *Cancers.* 2022;14(20):4999.
- Tsuchida A, Ogiso M, Nakamura Y, Kiso M, Furukawa K, Furukawa K. Molecular cloning and expression of human ST6GalNAc III: restricted tissue distribution and substrate specificity. *J Biochem.* 2005;138(3):237-43.
- Tsuchida A, Okajima T, Furukawa K, Ando T, Ishida H, Yoshida A, *et al.* Synthesis of disialyl Lewis a (Le(a)) structure in colon cancer cell lines by a sialyltransferase, ST6GalNAc VI, responsible for the synthesis of alpha-series gangliosides. *J Biol Chem.* 2003;278(25):22787-94.
- Ulm C, Saffarzadeh M, Mahavadi P, Müller S, Prem G, Saboor F, *et al.* Soluble polysialylated NCAM: a novel player of the innate immune system in the lung. *Cell Mol Life Sci.* 2013;70(19):3695-708.
- Vajaria BN, Patel KR, Begum R, Patel PS. Sialylation: an Avenue to Target Cancer Cells. *Pathology & Oncology Research.* 2016;22(3):443-7.
- Vallejo Ruiz V. Les acides sialiques comme récepteurs viraux et les enzymes impliquées dans leur expression. Doctoral thesis. University of Lille (FR). 2001.

- Vallejo-Ruiz V, Haque R, Mir A-M, Schwientek T, Mandel U, Cacan R, *et al.* Delineation of the minimal catalytic domain of human Gal β 1-3GalNAc α 2,3-sialyltransferase (hST3Gal I). *Biochimica et Biophysica Acta*. 2001;1549(2):161-73.
- van Diepen L, Buettner FFR, Hoffmann D, Thiesler CT, von Bohlen Und Halbach O, von Bohlen Und Halbach V, *et al.* A patient-specific induced pluripotent stem cell model for West syndrome caused by ST3GAL3 deficiency. *Eur J Hum Genet*. 2018;26(12):1773-83.
- Vanbeselaere J, Chang L-Y, Harduin-Lepers A, Fabre E, Yamakawa N, Slomianny C, *et al.* Mapping the Expressed Glycome and Glycosyltransferases of Zebrafish Liver Cells as a Relevant Model System for Glycosylation Studies. *Journal of Proteome Research*. 2012;11(4):2164-77.
- Varki A. Biological roles of glycans. *Glycobiology*. 2017;27(1):3-49.
- Varki A, Schnaar RL, Schauer R. Sialic Acids and Other Nonulosonic Acids. *Essentials of Glycobiology*. 2015b;15.
- Varki A, Cummings RD, Aebi M, Packer NH, Seeberger PH, Esko JD, *et al.* Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology*. 2015a;25(12):1323-4.
- Varki A. Evolutionary Forces Shaping the Golgi Glycosylation Machinery: Why Cell Surface Glycans Are Universal to Living Cells. *Cold Spring Harbor Perspectives in Biology*. 2011;3(6).
- Varki A, Schauer R. Sialic Acids. In *Essentials of Glycobiology*. 2nd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009.
- Varki A. Multiple changes in sialic acid biology during human evolution. *Glycoconjugate Journal*. 2009;26(3):231-45.
- Varki A, Lowe JB. Biological roles of glycans. In: *Essentials of Glycobiology*. Cold Spring Harbor (NY). 2009. p. 75–88.
- Venkatakrishnan V, Padra JT, Sundh H, Sundell K, Jin C, Langeland M, *et al.* Exploring the Arctic Charr Intestinal Glycome: Evidence of Increased *N*-Glycolylneuraminic Acid Levels and Changed Host–Pathogen Interactions in Response to Inflammation. *J Proteome Res*. 2019;18(4):1760-73.
- Venuto MT. Characterization of the polysialylation machinery in fish. Doctoral thesis. University of Rostock (DEU). 2021.
- Venuto MT, Martorell-Ribera J, Bochert R, Harduin-Lepers A, Rebl A, Galuska SP. Characterization of the Polysialylation Status in Ovaries of the Salmonid Fish *Coregonus maraena* and the Percid Fish *Sander lucioperca*. *Cells*. 2020a;9(11):2391.
- Venuto MT, Decloquement M, Martorell Ribera J, Noel M, Rebl A, Cogez V *et al.* Vertebrate Alpha2,8-Sialyltransferases (ST8Sia) : A Teleost Perspective. *IJMS*. 2020b;21(2):513.
- Villanueva-Cabello TM, Gutiérrez-Valenzuela LD, Salinas-Marín R, López-Guerrero DV, Martínez-Duncker I. Polysialic Acid in the Immune System. *Front Immunol*. 2022;12:823637.
- Vogt J, Glumm R, Schlüter L, Schmitz D, Rost BR, Streu N, *et al.* Homeostatic regulation of NCAM polysialylation is critical for correct synaptic targeting. *Cell Mol Life Sci*. 2012;69(7):1179-91.
- Volff JN. Genome evolution and biodiversity in teleost fish. *Heredity*. 2005;94(3):280-94.
- Volkers G, Worrall LJ, Kwan DH, Yu CC, Baumann L, Lameignere E, *et al.* Structure of human ST8SiaIII sialyltransferase provides insight into cell-surface polysialylation. *Nat Struct Mol Biol*. 2015;22(8):627-35.
- von Gunten S, Shoenfeld Y, Blank M, Branch DR, Vassilev T, Käsermann F, *et al.* IVIG pluripotency and the concept of Fc-sialylation: challenges to the scientist. *Nat Rev Immunol*. 2014;14(5):349-349.
- Vorobiev I, Matskevich V, Kovnir S, Orlova N, Knorre V, Jain S, *et al.* Chemical polysialylation: design of conjugated human oxyntomodulin with a prolonged anorexic effect *in vivo*. *Biochimie*. 2013;95(2):264-70.
- Wang Y, Neumann H. Alleviation of Neurotoxicity by Microglial Human Siglec-11. *Journal of Neuroscience*. 2010;30(9):3482-8.
- Wang X, Zhang LH, Ye XS. Recent development in the design of sialyltransferase inhibitors. *Med Res Rev*. 2003;23(1):32-47.

- Wang X, Vertino A, Eddy RL, Byers MG, Jani-Sait SN, Shows TB, *et al.* Chromosome mapping and organization of the human beta-galactoside alpha 2,6-sialyltransferase gene. Differential and cell-type specific usage of upstream exon sequences in B-lymphoblastoid cells. *J Biol Chem.* 1993;268(6):4355-61.
- Warren L. The distribution of sialic acids in nature. *Comparative Biochemistry and Physiology.* 1963;10(2):153-71.
- Wei A, Fan B, Zhao Y, Zhang H, Wang L, Yu X, *et al.* ST6Gal-I overexpression facilitates prostate cancer progression via the PI3K/Akt/GSK-3 β / β -catenin signaling pathway. *Oncotarget.* 2016;7(40):65374-88.
- Weide T, Herrmann L, Bockau U, Niebur N, Aldag I, Laroy W, *et al.* Secretion of functional human enzymes by *Tetrahymena thermophila*. *BMC Biotechnol.* 2006;6(1):19.
- Weinhold B, Seidenfaden R, Röckle I, Mühlenhoff M, Schertzing F, Conzelmann S, *et al.* Genetic ablation of polysialic acid causes severe neurodevelopmental defects rescued by deletion of the neural cell adhesion molecule. *J Biol Chem.* 2005;280(52):42971-7.
- Weinstein J, Lee EU, McEntee K, Lai PH, Paulson JC. Primary structure of beta-galactoside alpha 2,6-sialyltransferase. Conversion of membrane-bound enzyme to soluble forms by cleavage of the NH₂-terminal signal anchor. *J Biol Chem.* 1987;262(36):17735-43.
- Werneburg S, Buettner FFR, Erben L, Mathews M, Neumann H, Mühlenhoff M, *et al.* Polysialylation and lipopolysaccharide-induced shedding of E-selectin ligand-1 and neuropilin-2 by microglia and THP-1 macrophages: Polysialylation and Shedding of ESL-1 and NRP2. *Glia.* 2016;64(8):1314-30.
- West CM, Malzl D, Hykollari A, Wilson IBH. Glycomics, Glycoproteomics, and Glycogenomics: An Inter-Taxa Evolutionary Perspective. *Molecular & Cellular Proteomics.* 2021;20:100024.
- Willis LM, Gilbert M, Karwaski MF, Blanchard MC, Wakarchuk WW. Characterization of the -2,8-polysialyltransferase from *Neisseria meningitidis* with synthetic acceptors, and the development of a self-priming polysialyltransferase fusion enzyme. *Glycobiology.* 2008;18(2):177-86.
- Wu ZL, Person AD, Burton AJ, Singh R, Burroughs B, Fryxell D, *et al.* Direct fluorescent glycan labeling with recombinant sialyltransferases. *Glycobiology.* 2019;29(11):750-4.
- Wu ZL, Ethen CM, Prather B, Machacek M, Jiang W. Universal phosphatase-coupled glycosyltransferase assay. *Glycobiology.* 2011;21(6):727-33.
- Wu D, Gilormini PA, Toda S, Biot C, Lion C, Guérardel Y, *et al.* A novel C-domain-dependent inhibition of the rainbow trout CMP-sialic acid synthetase activity by CMP-deaminoneuraminic acid. *Biochemical and Biophysical Research Communications.* 2022;617:16-21.
- Xu T, Zhang XH. *Edwardsiella tarda*: an intriguing problem in aquaculture. *Aquaculture.* 2014;431:129-35.
- Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z, Liu X, *et al.* The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat Biotechnol.* 2011;29(8):735-41.
- Xu L, Kurusu Y, Takizawa K, Tanaka J, Matsumoto K, Taniguchi A. Transcriptional regulation of human beta-galactoside alpha2,6-sialyltransferase (hST6Gal I) gene in colon adenocarcinoma cell line. *Biochem Biophys Res Commun.* 2003;307(4):1070-4.
- Yamakawa N, Vanbeselaere J, Chang L-Y, Yu S-Y, Ducrocq L, Harduin-Lepers A, *et al.* Systems glycomics of adult zebrafish identifies organ-specific sialylation and glycosylation patterns. *Nat Commun.* 2018;9(1):4647.
- Yabe U, Sato C, Matsuda T, Kitajima K. Polysialic Acid in Human Milk. *Journal of Biological Chemistry.* 2003;278(16):13875-80.
- Yoshida Y, Kurosawa N, Kanematsu T, Kojima N, Tsuji S. Genomic structure and promoter activity of the mouse polysialic acid synthase gene (mST8Sia II). Brain-specific expression from a TATA-less GC-rich sequence. *J Biol Chem.* 1996;271(47):30167-73.
- Yu C-C, Hill T, Kwan DH, Chen H-M, Lin C-C, Wakarchuk W, *et al.* A plate-based high-throughput activity assay for polysialyltransferase from *Neisseria meningitidis*. *Analytical Biochemistry.* 2014;444:67-74.
- Yu SY, Chang LY, Cheng CW, Chou CC, Fukuda MN, Khoo KH. Priming mass spectrometry-based sulfoglycomic mapping for identification of terminal sulfated lacdiNAc glycocone. *Glycoconj J.* 2013;30(2):183-94.
- Zapater JL, Colley KJ. Sequences Prior to Conserved Catalytic Motifs of Polysialyltransferase ST8Sia IV Are Required for Substrate Recognition. *Journal of Biological Chemistry.* 2012;287(9):6441-53.

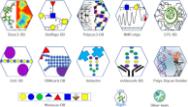
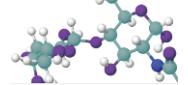
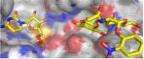
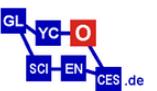
- Zhang Y, Wang R, Feng Y, Ma F. The role of sialyltransferases in gynecological malignant tumors. *Life Sciences*. 2020;263:118670.
- Zhang T, She Z, Huang Z, Li J, Luo X, Deng Y. Application of sialic acid/polysialic acid in the drug delivery systems. *Asian Journal of Pharmaceutical Sciences*. 2014;9(2):75-81.
- Zhang R, Jain S, Rowland M, Hussain N, Agarwal M, Gregoriadis G. Development and Testing of Solid Dose Formulations Containing Polysialic Acid Insulin Conjugate: Next Generation of Long-Acting Insulin. *J Diabetes Sci Technol*. 2010;4(3):532-9.
- Zhang Y, Lee YC. Acid-catalyzed Lactonization of α 2,8-Linked Oligo/Polysialic Acids Studied by High Performance Anion-exchange Chromatography. *Journal of Biological Chemistry*. 1999;274(10):6183-9.
- Zhao P, Praissman JL, Grant OC, Cai Y, Xiao T, Rosenbalm KE, *et al*. Virus-Receptor Interactions of Glycosylated SARS-CoV-2 Spike and Human ACE2 Receptor. *Cell Host & Microbe*. 2020;28(4):586-601.e6.
- Zhou GP, Huang RB. The Graphical Studies of the Major Molecular Interactions for Neural Cell Adhesion Molecule (NCAM) Polysialylation by Incorporating Wenxiang Diagram into NMR Spectroscopy. *IJMS*. 2022;23(23):15128.
- Zhou JM, Wang T, Zhang KH. AFP-L3 for the diagnosis of early hepatocellular carcinoma: A meta-analysis. *Medicine (Baltimore)*. 2021;100(43):e27673.
- Zhou GP, Huang RB, Troy II F. 3D Structural Conformation and Functional Domains of Polysialyltransferase ST8Sia IV Required for Polysialylation of Neural Cell Adhesion Molecules. *PPL*. 2015;22(2):137-48.
- Ziak M, Roth J. Expression of oligo/poly α 2,8-linked deaminoneuraminic acid and megalin during kidney development and maturation: mutually exclusive distribution with poly α 2,8-linked N-acetylneuraminic acid of N-CAM. *Histochemistry and Cell Biology*. 1999;112(2):169-78.
- Ziak M, Qu B, Zuo X, Zuber C, Kanamori A, Kitajima K, *et al*. Occurrence of poly(α 2,8-deaminoneuraminic acid) in mammalian tissues: widespread and developmentally regulated but highly selective expression on glycoproteins. *Proc Natl Acad Sci USA*. 1996;93(7):2759-63.
- Zlatina K, Galuska SP. Polysialic Acid Modulates Only the Antimicrobial Properties of Distinct Histones. *ACS Omega*. 2019;4(1):1601-10.
- Zlatina K, Saftenberger M, Kühnle A, Galuska C, Gärtner U, Rebl A, *et al*. Polysialic Acid in Human Plasma Can Compensate the Cytotoxicity of Histones. *IJMS*. 2018;19(6):1679.
- Zuber C, Lackie PM, Catterall WA, Roth J. Polysialic acid is associated with sodium channels and the neural cell adhesion molecule N-CAM in adult rat brain. *Journal of Biological Chemistry*. 1992;267(14):9965-71.
- Zuber C, Roth J. The relationship of polysialic acid and the neural cell adhesion molecule N-CAM in Wilms tumor and their subcellular distributions. *Eur J Cell Biol*. 1990;51(2):313-21.

APPENDIX

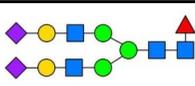
Appendix 1: Bioinformatic tools and databases to study glycans and glycosyltransferases for vertebrates. All these websites were available on December 19th, 2022.

Name	Description	Website
 Addgene	Database of plasmids useful for molecular cloning and glycosyltransferase engineering.	https://www.addgene.org/browse
 AlphaFold	Software tool based on Artificial Intelligence system to computationally predict protein structures with accuracy and speed, developed from DeepMind.	https://alphafold.ebi.ac.uk
 BLAST	Tool for Basic Local Alignment with sequences from NCBI – NIH databases to calculates the statistical significance.	https://blast.ncbi.nlm.nih.gov/Blast.cgi
 CarbArrayART	Software tool for glycan microarray experimental design, data storage, processing, presentation and reporting.	https://glycosciences.med.ic.ac.uk/carbarrayart.html
 ctdbase	Database for comparative toxicogenomics to understand how environmental exposures affect human health with chemical and glycosyltransferase gene interactions.	https://ctdbase.org/
 CASPER	Tool to determine the sequence of a poly- or oligo-saccharide from NMR chemical shifts.	http://www.casper.organ.su.se/casper/
 CAZy	Database describing the families of structurally-related catalytic and carbohydrate-binding modules (or functional domains) of enzymes that degrade, modify, or create glycosidic bonds.	http://www.cazy.org
 ChemDraw	Software to draw chemical structures like monosaccharides such as sialic acids.	https://perkinelmerinformatics.com/products/research/chemdraw
 ClustalW prabi	Tool for pairwise and multiple sequence alignment which offers additional informations with color and secondary structures predicted.	https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_clustalw.html

 DAGR	Database of anti-glycans reagents containing more than 1100 Carbohydrate-Binding Antibodies and lectins.	https://dagr.ccr.cancer.gov/
 EMBL-EBI	EMBOSS Needle and Clustal Omega tools for Pairwise and Multiple Sequence Alignment to generate alignments between two, three and more sequences.	https://www.ebi.ac.uk/Tools/psa/emboss_needle/ https://www.ebi.ac.uk/Tools/msa/clustalo/
 ENDScript	Tool for extraction and comprehensive analysis of primary to quaternary protein structure from PDB files.	https://endscript.ibcp.fr/ESript/cgi-bin/ENDscript.cgi
 Enzyme Finder	Database listing all NEB-commercially restriction enzymes for molecular cloning and glycosyltransferases engineering.	https://enzyme finder.neb.com/#!/main#nebheader
 ESript	Tool which renders sequence similarities and secondary structure information from aligned sequences for analysis.	https://esript.ibcp.fr/ESript/ESript/
 EXPASy Glycomics	Database/toolbox of glycoinformatics resources for simulating (GlycoDigest), modelling (SWISS-MODEL), predicting (GlycoMod) or visualising information (GlyConnect), relative to glycans, glycoproteins and glycan-binding proteins.	http://www.expasy.org/glycomics
 FlowJo	Software for analyzing flow cytometry data, notably after exosialylation on cells surface.	https://www.flowjo.com/
 GLYCAM	Tool to facilitate quality computer modeling of glycans gathering a 3D structure database and prediction tools.	https://glycam.org/
 Glycan Library	Database which lists around 800 glycan probes derived from diverse natural sources or chemically synthesized.	https://www.glycosciences.med.ic.ac.uk/glycanLibraryIndex.html
 Glycan Structure	Tool for glycan modeling and simulation on the target protein.	http://www.glycanstructure.org/

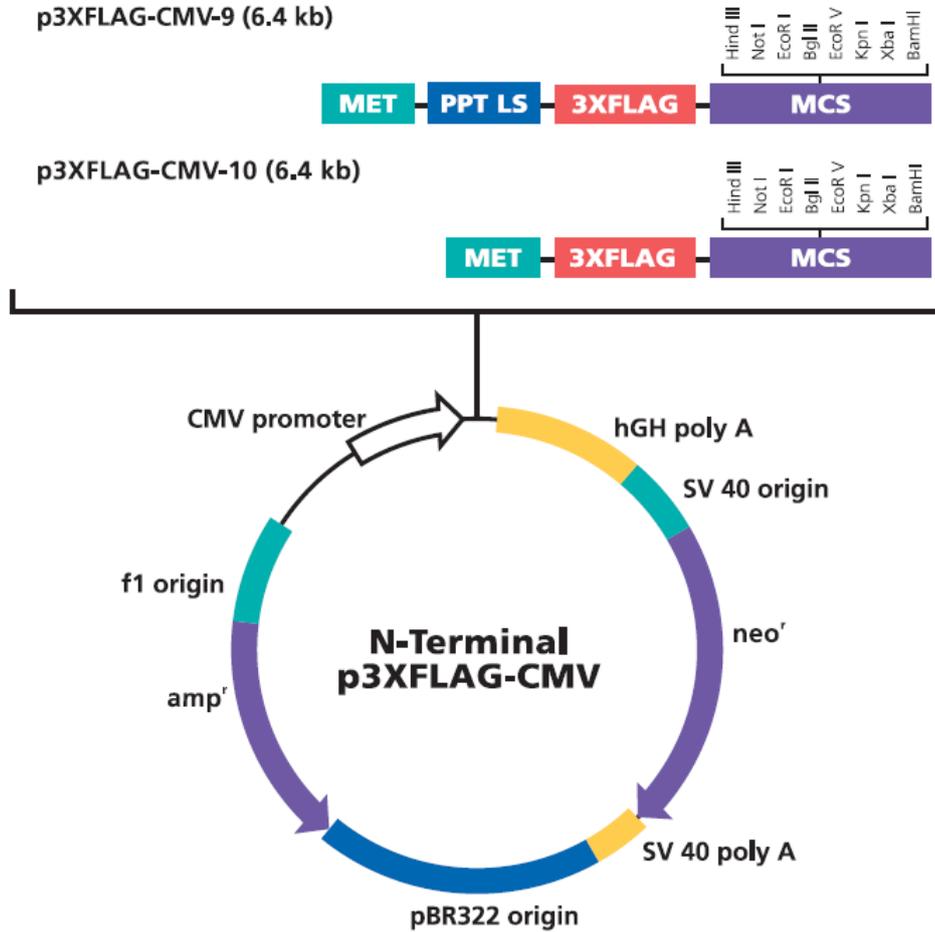
 Glyco3D	<p>Databases and tools with 3D structures of mono, di, oligo and polysaccharides and carbohydrate recognizing proteins (lectins, antibodies, glycosyltransferases and glycosaminoglycan binding proteins), drawing structures and functions informations.</p>	<p>http://glyco3d.cermav.cnrs.fr/home.php</p>
 Glyco.me	<p>Database to explore the glycogenome, enzymes and glycoproteins and their involvements in cellular functions in both health and diseases.</p>	<p>https://glyco.me/</p>
 GlycoDomain Viewer	<p>Tool to study site protein glycosylation with respect to protein context and conservation.</p>	<p>http://glycodomain.glycomics.ku.dk</p>
 GlycoEnzymes	<p>Database of glycoenzymes and expression constructs for the production of human glycosylation enzymes.</p>	<p>http://glycoenzymes.ccruc.uga.edu</p>
 GlycoEpitope	<p>Database listing information on carbohydrate antigens, glyco-epitopes and antibodies.</p>	<p>https://www.glycoepitope.jp</p>
 GlyGen	<p>Database for computational and informatics resources with glycans and glycoconjugates data integration and harmonization from multiple international data sources.</p>	<p>https://www.glygen.org/</p>
 GlycoGlyph	<p>Tool kit for glycan drawing with SNFG compatible symbols.</p>	<p>https://glycotoolkit.com/Tools/GlycoGlyph/</p>
 GlycoPedia	<p>Library of 150 monosaccharides 1D, 2D and 3D structures found as components of bio-active glycans, oligo and polysaccharides.</p>	<p>https://www.glycopedia.eu/resources/library-of-bio-active-monosaccharides-1d-2d-3d-structures/article/presentation</p>
 Glycosciences.de	<p>Databases and tools for glycosciences such as MonosaccharideDB, GlycoMaps, CARP and PDB-CARE for conformational maps of glycans, GlyProt to connect N-glycans on a 3D protein structure or NMR SE for Shift Estimation of glycans.</p>	<p>http://glycosciences.de/index.php http://www.monosaccharidedb.org/start.action</p>

 GlyCosmos	<p>Database with lots of resources of glycogenes, glycoproteins, lectins, glycolipids and glycans and providing tools like SugarDrawer, GlycanBuilder and GlycoMaple for pathways, GlycomeAtlas for human, mouse and zebrafish models or with glycomics MS data with GlycoPOST, UniCarb-DR or GALAXY.</p>	<p>https://glycosmos.org/ https://glycopost.glycosmos.org/ https://unicarb-dr.glycosmos.org/ https://glycoanalysis.info/system</p>
 GlycoStore	<p>Database of analytical data acquired from chromatographic, electrophoretic and MS composition of N-, O-, GSL glycans and free oligosaccharides associated with a range of glycoproteins, glycolipids and biotherapeutics.</p>	<p>https://glycostore.org/collections</p>
 GlyTouCan	<p>Database for registry of glycans structures and assignment of unique accession number.</p>	<p>https://glytoucan.org/</p>
 GOLD	<p>Software tool for Protein-Ligand Docking Software, particularly use for chemical drug design and glycans interactions.</p>	<p>https://www.ccdc.cam.ac.uk/solutions/csd-discovery/components/gold/</p>
 GTEx Portal	<p>Resource to study human gene expression and regulation and its relationship to genetic variation as for glycosyltransferases.</p>	<p>https://gtexportal.org/home/</p>
 Prism GraphPad Prism	<p>Software for comprehensive analysis and powerful statistics useful to represent scientific graphics after sialylation reactions.</p>	<p>https://www.graphpad.com/scientific-software/prism/</p>
 ImageJ	<p>Software for processing and analyzing scientific images, useful for WB sialyltransferases and polysialylation signals quantifications.</p>	<p>https://imagej.net/software/imagej/</p>
 KEGG Glycan	<p>Database with a collection of glycan structures integrated with other KEGG resources, including KEGG Pathway, Module, Network and Disease.</p>	<p>https://www.genome.jp/kegg/glycan/</p>
 LiteMol	<p>Tool for visualization of 3D macromolecular data such as glycosyltransferases.</p>	<p>https://www.litemol.org/</p>
 MestReNova	<p>Software to process analytical chemistry data, especially from NMR spectra and techniques after CMP-Sia synthesis.</p>	<p>https://mestrelab.com/software/mnova/</p>

 NetNGlyc - 1.0 NetNGlyc	Tool for prediction of N-glycosylation sites in glycoproteins using artificial neural networks examining the sequence context of NxS/T sequons.	https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0
NetOGlyc - 4.0 NetOGlyc	Tool for prediction of O-GalNAc (mucin type) glycosylation sites in mammalian glycoproteins using neural network predictions.	https://services.healthtech.dtu.dk/service.php?NetOGlyc-4.0
 PDBe-KB	Database for integration and enrichment of 3D-structure data and functional annotations to enable basic and translational research.	https://www.ebi.ac.uk/pdbe/pdbe-kb/
 pDRAW32	Software for DNA analysis, molecular cloning, determination of enzyme restriction site and primers for glycosyltransferases engineering.	https://www.acaclone.com/
 PhyloFish	Database for phylogenomic analysis of gene duplications in teleost fish by RNAseq approach.	http://phylofish.sigena.org/index.html
 Phyre2	Tool for prediction and analyse of protein structure, function and mutations.	http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index
 PolyS	Software package for building 3D structures of polysaccharides.	http://www.models.life.ku.dk/polys
 PyMOL	Software for visualization and animation of 3D chemical and protein structures such as glycosyltransferases.	https://pymol.org/2/
 RCSB	Database for protein structure annotations of large biological molecules such as glycosyltransferases.	https://www.rcsb.org/
 RINGS	Database providing algorithmic and tools for glycobiology research such as Glycan Miner tool for mining subtrees from a set of glycan structures.	https://rings.glycoinfo.org/
 SNFG	Database for standardization with Symbol Nomenclature for drawing glycan structures.	https://www.ncbi.nlm.nih.gov/glycans/snfg.html

 SugarBindDB	Database on interaction between carbohydrate sequences and pathogenic organisms specifically adherence via lectins or adhesins.	https://sugarbind.expasy.org/
 SweetUnityMol	Software to display 3D structures of carbohydrates, polysaccharides and glycoconjugates	http://sourceforge.net/projects/un-itymol/files/
THE HUMAN PROTEIN ATLAS The Human Protein Atlas	Database which map all the human proteins at different levels : cells, tissues, organs and integration all various omics technologies.	https://www.proteinatlas.org/
 UniCarbKB	Database with annotation of glycans from glycoproteins, including tools as SPRINT-Gly for prediction of N- and O-glycosylation sites of human and mouse proteins.	http://unicarbkb.org/
 UniLectin	Database classifying lectins in different families based on their origin, X-ray structures and the structure of their carbohydrate-binding domains or monosaccharide specificity.	https://www.unilectin.eu/
 UniProt	Database for protein sequence with informations about identification, glycosylation and biological functions derived from the research litterature.	http://www.uniprot.org/
 YASARA	Software for modeling and simulation based on artificial reality to generate molecular graphics.	http://www.yasara.org/
 Zen Zeiss	Software for microscopy imaging and processing, notably used for acquisition after sialylation on cell surfaces.	https://www.zeiss.com/microscopy/en/products/software/zeiss-zen-lite.html#getzenlite

Appendix 2: Plasmid constructs used for molecular cloning of STs. Each vector contains 3XFLAG tag sequence close to the Multiple Cloning Site (MCS) where restriction sites were chosen to insert polySTs sequences. In contrast to CMV10, CMV9 vector has a preprotrypsin leader sequence (PPT-LS) directing secretion of the fusion protein into the culture medium.



Multiple Cloning Site

(p3XFLAG-CMV-9* and p3XFLAG-CMV-10)

3XFLAG Peptide Sequence															
Met*	Asp	Tyr	Lys	Asp	His	Asp	Gly	Asp	Tyr	Lys	Asp	His	Asp	Ile	
ATG	GAC	TAC	AAA	GAC	CAT	GAC	GGT	GAT	TAT	AAA	GAT	CAT	GAC	ATC	
TAC	CTG	ATG	TTT	CTG	GTA	CTG	CCA	CTA	ATA	TTT	CTA	GTA	CTG	TAG	
3XFLAG Peptide Sequence															
Asp	Tys	Lys	Asp	Asp	Asp	Asp	Lys								
GAT	TAC	AAG	GAT	GAC	GAT	GAC	AGG	CTT	GGG	GCC	GCG	AAT	TCA	TCG	ATA
CTA	ATG	TTC	CTA	CTG	CTA	CTG	TTC	GAA	CGC	CGG	CGC	TTA	AGT	AGC	TAT
								Hind III							
Bgl II		EcoR V			Kpn I		Xba I		Bam HI						
GAT	CTG	ATA	TCG	GTA	CCA	GTC	GAC	TCT	AGA	GGA	TCC	CGG	GTG		
CTA	GAC	TAT	AGC	CAT	GGT	CAG	CTG	AGA	TCT	CCT	AGG	CCC	CAC		

*For pFLAG-CMV-9, the Met-preprotrypsin leader sequence (PPT LS) precedes the FLAG coding sequence.

Appendix 3: Table of sense and antisense oligonucleotides used to generate cDNA encoding truncated form of the *C.ma* st8sia2 r-1, st8sia2 r-2 and st8sia4 lacking their first 28, 35 or 50 aa residues. cDNAs were amplified by PCR using previous construct p3×FLAG-CMV10 as DNA templates for *st8sia4* and *st8sia2 r-1* and pcDNA3.1 for *st8sia2 r-2*, leading to Δ28 ST8Sia II-r1, Δ35 ST8Sia II-r1 or Δ50 ST8Sia II-r1, and Δ28 ST8Sia II-r2, Δ35 ST8Sia II-r2 or Δ50 ST8Sia II-r2 or Δ28 ST8Sia IV constructs.

<i>st8sia4:</i>	Δ28	5'-GCAAAGCTTGAAGAACATCAGGAAGCTCAAGTC-3' (sense)
		5'-CGGGGTACCTTAAGATTCGCACTTCGAAGTCG-3' (antisense)
<i>st8sia2 r-1:</i>	Δ28	5'-GTTAAGCTTGAGGAAGAAATTGCGAATATTGGAG-3' (sense)
	Δ35	5'-AAAAAGCTTGGAGGTTCCAGAACATTGTACTTG-3' (sense)
	Δ50	5'-AAAAAGCTTAACAGAAATGTAGCAGTGAAAGCCAATC-3' (sense)
		5'-CTTGCCCCCTTGCTCCATACCAC-3' (antisense)
<i>st8sia2 r-2:</i>	Δ28	5'-AAAAAGCTTGAAGAAGAAATTGCGAATATTGGAG-3' (sense)
	Δ35	5'-AAAAAGCTTGGAGGTTCCAGAAAATTGTACATGCAC-3' (sense)
	Δ50	5'-AAAAAGCTTAACAGAAATGTAGCAGTGAAGGCCAATC-3' (sense)
		5'-AAAGGTACCTTAAGTTCCTGCATCACAAGAGCCAGT-3' (antisense)

Appendix 4: Full nucleotide and aa sequences of sialyltransferases used and acceptors produced. Deletions in aa sequences for enzymes used in experiments are indicated in red.

>Cormar_GT29_ST8Sia2-r1_NUC : 1137 bp

ATGCAGTTAGAATTC CGAACGGTGATGTTTGGAAATTGTAACACTGCTGGTCATATTTCTGATCATTTGCTGATATTGCTGAAATCGAGGAAGAAATTCGGAATATTGGAGGTTCCAG
AACATTTGACTTGCACAGCCTCATCCCCAAGCCCAACAGAAATGTAGCAGTGAAGGCcAATCCTACACCTTTAATCAGTGAAGGGGAGGACAAGAGCCCTGCTTCTCCTTCAGGTT
TGAATAATACCACCAGGCTCTCTTCAGATAACTGGACCTTCAACAGAACCCTCTCCAGCCTCATCAgAAGAACATCCTGAGGTTCTTTGACCCAGAGAGGGACATCTCTATTCTG
AAAGGCACATTGAAACCTGGAGATGTCATCCACTACATCTTTGACCGTCAGAGCACTACCAACATCTCAGAGAACCCTGTACCGGTTGCTACCCACTGCGTCCCCCATGAAGAACCA
GCACCACAGACGCTGTGCCATCGTAGGGAACCTCTGGGATCCTACTCAACAGCAGCTGTGGGCCAGAGATAGACTCCCATGACTTTGTTATCAGATGTAACCTGGCGCCAGTGGAGG
AGTATGCTGgGGATGTGGGGCGGCGCACCAATCTGGTGACCATGAACCCGTGGTGGTTTCAGCGGGCCTTCCATGACCTGGCCAGTGGAGGAGCGCTTCCCTGCAGCGG
CTCCGGGGCCTCAGCGGCAGCGTACTGTGGATCCCAGCCTTTCATGGCCAAGGGAGGAGAGGGTGGAGTGGGCCATCCGTCTCATCCTGCTGCACACGGTGGATGTGCACAC
CGCCTTCCCCTCACTGCGCCTTCTCCATGCTGTGAGAGGATTTGGCTGACTAACCAACGTCCAGATCAAGAGACCCACCCTGGTCTGCTGATGTACACCATGGCCACCCGCTTCT
GTGAGGAGATCCACCTGTACGGCTTCTGGCCTTCCCCTCGGGATTCACAGGGGATACCAGTCAAGTACCATTACTATGACACTCTGACCTACGAGTACACATCCCATGCCAGTCCA
CATACTATGCCACTGGAGTTCAGGACACTGAGTTCACCTCCACAGACAGGGGGCGCTGCGGCTCAACACTGGCTCTTGTGATGCAGGAATGTAA

>Cormar_GT29_ST8Sia2-r1 : 378 aa (Δ 35)

MQLEFRTMVFGIVTLLVIFLI**IADIAE**IEEEI**IANI**IGGSRTLYLHSLIPKPNRNVAVKANPTPLISEGEDKSPASPSGLNNTTRLSSDNWTFNRTLSSLIRKNIILRFFDPERDISIL
KGTLPKPGDVIHYIFDRQSTTNI SENLYRLLPTASPMKNQHRRCAIVGN SGILLNSSCGPEIDSHDFVIRCNLAPVEEYAGDVGRRTNLVMTNPSV VQRAFHDLASEQWREFLQR
LRGLSGSVLWIPAFMAKGGEERVEWAIRLILLHTVDVHTAFPSLRLLHAVRGYWL TNNVQIKRPTTGLLMYTMATRFCEEIHLYGFWPFPRDSQGI PVKYHYDYDTLT YEYTS HASP
HTMPLEFR TLSSLHRQ GALRLNTGSCDAGM

>Cormar_GT29_ST8Sia2-r2_NUC : 1137 bp

ATGCAGTTAGAATTC CGAACATTGATGTTTGGAAATTGTAACAGTCTGGTAATATTTCTGATCATTTGCTGATATTGCTGAAGTCGAAGAAGAAATTCGGAATATTGGAGGTTCCAG
AAAATTGTACATGCACAGCCTCATTTCCCAAGCCGAACAGAAATGTAGCAGTGAAGGCCAATCCTAAACCATTAGTCAAGTGAAGGGGAGGACAAGAGCCCTGCTTCTCCTTCATACT
CAAATAATACCACCAAGCTCTCTTCAGATAACTGGACGTTCAACAGGAGCCTCTCCAATTTCTATAGGGAAGAACATCCTGAGGTTCTTCGACCCAGAGAGGGACATCTCTATTCTG
AAGGGCACATTGAAACCTGGAGATGTCATCCACTACATCTTTGACCGTCAGAGCACTACCAACATCTCAGAGAACCCTGTACCGGTTGCTGCCCCACTGTGTCCCCCATGAAGAACCA
GCACCACAGGCAGTGTGCCATCGTAGGGAACCTCTGGGATCCTACTCAACAGCAGCTGTGGGCCAGAGATAGACTCCTATGACTTTGTTATCAGATGTAACCTGGCGCCAGTGGAGG
AGTATGCTGGGGACGTGGGCCGGCGCACCAACCTGGTGACCATGAACCCATCTGTGGTGCAGCGGGCCTTCCAGGACCTGGCCAGCGAGGAGTGGAGGGAGCGCTTCCCTGCAGCGG
CTCCGGGGCCTCAGCGGCAGCGTACTGTGGATCCCAGCCTTTCATGGCCAAGGGAGGGGAGGAGGGTGGAGTGGGCCATCCGTCTCATCCTGCTGCACACAGTGGATGTGCACAC
CGCCTTCCCCTCACTGCGCCTTCTCCATGCTGTGAGAGGATTTGGCTGACCAACAACGTCCAGATCAAGAGACCCACCCTGGTCTGCTGATGTACACCATGGCCACTCGCTTCT
GTGAGGAGATCCACCTGTACGGTTCCTGGCCTTCCCCTCAGGATTCACAGGGGAAACCAGTGAAGTACCCTACTATGATACTCTGACCTACACGTACACGTCCCATGCCAGTCCA
CACACTATGCCACTGGAGTTCAGGACCTGAGTTCACCTCCACAGACAGGGGGCGCTGCGGCTCCACACTGGCTCTTGTGATGCAGGAAC TTAA

>Cormar_GT29_ST8Sia2-r2 : 378 aa (Δ 35)

MQLEFRTLMFIVTVLVIFLI**IADIAE**VEEEE**IANI**IGGSRKLYMHSLIPKPNRNVAVKANPKPLVSEGEDKSPASPSYSNNTTKLSSDNWTFNRSLSNSIGKNIILRFFDPERDISIL
KGTLPKPGDVIHYIFDRQSTTNI SENLYRLLPTVSPMKNQHRRCAIVGN SGILLNSSCGPEIDSYDFVIRCNLAPVEEYAGDVGRRTNLVMTNPSV VQRAFDLASEEWREFLQR
LRGLSGSVLWIPAFMAKGGEERVEWAIRLILLHTVDVHTAFPSLRLLHAVRGYWL TNNVQIKRPTTGLLMYTMATRFCEEIHLYGFWPFQDSQGKPVKYHYDYDTLT YTYTSHAS P
HTMPLEFR TLSSLHRQ GALRLHTGSCDAGT

>Homsap_GT29_ST8Sia2_NUC : 1128 bp

ATGCAGCTGCAGTTCGGGAGCTGGATGCTGGCCGCGCTCACGCTGCTCGTGGTCTTCCCTCATCTTCGCAGACATCTCAGAGATCGAAGAAGAAATCGGGAAATTCGGGAGGCAGAGG
TACAATCAGATCAGCTGTGAACAGCTTACATAGCAAATCTAATAGAGCTGAAGTTGTAATAAACGGCTCCCTCATCACCAGCTGTTGTTGACAGAAGTAATGAAAGCATCAAGCACA
ACATCCAGCCAGCCTCGTCCAAATGGAGACATAACCAGACGCTCTCTCTGAGGATCAGGAAGCAGATTTTAAAGTTCTTGGATGCTGAAAAGGACATTTCTGTCTAAAGGGAACC
CTGAAGCCTGGAGATATTTATTCATTACATCTTCGATCGAGACAGCACCATGAATGTGTCCCAGAACCCTACGAGCTCCCTCCAGGACTTCGCCACTGAAGAATAAGCACTTTGG
GACTTGTGCCATCGTGGGCAACTCGGGGGTCTTGTGAACAGCGGCTGTGGGCAGGAGATTGACGCCACAGCTTCGTCATCAGGTGCAACCTGGCCCCAGTACAGGAGTATGCC
GGGATGTGGGGCTCAAGACAGACCTGGTAACCATGAACCCCTCGGTCATCCAGCGGGCCTTTGAGGACTTGGTCAATGCCACGTGGCGGGAGAAGCTGCTGCAACGGCTGCACAGC
CTCAATGGCAGCATCCTGTGGATCCCTGCCCTTCATGGCCCCGGGGCGCAAGGAGCGTGTGAGTGGGTCAACGAGCTTATCCTGAAGCACCACGTCAACGTGCGCACTGCATACCC
CTCGCTGCGCCTGCTGCACGCCGTTTCGCGGATACTGGCTGACCAACAAAGTCCACATCAAAAGACCCACCACCGCCTCTTGATGTATACCCCTGGCCACACGTTTCTGCAACAAA
TCTACCTCTACGGCTTCTGGCCCTTCCGCTGGATCAGAACCAGAACCAGTCAAGTACCCTATTATGACAGCCTCAAGTATGGCTACACCTCCCAGGCCAGCCCCGATACCCATG
CCCTTGGAGTTTAAAGGCCCTCAAGAGCCTACATGAGCAGGGGGCTTTGAACTGACTGTTCGGCCAGTTCGATGGGGCCACGTAG

>Homsap_GT29_ST8Sia2 : 375 aa ($\Delta 23$)

MQLQFRSWMLAALTLVVFLIFADISEIEEEIGNSSGRGTRISAVNSLHKSNNRAEVIINGSSSPAVVDRSNESIKHNIQPASSKWRHNQTLNLRIRKQILKFLDAEKDISVLKGT
LKPGDI IHYIFDRDSTMNVSQNLVYELLPRTSPLKNKHFGTCAIVGNSGVLNLSGCGQEIDAHSFVIRCNLAPVQYARDVGLKTDLVTMNPVSIQRAFEDLVNATWREKLLQRLHS
LNGSILWIPAFMARGGKERVEWVNELILKHHVNVRTAYPSLRLLHAVRGYWLTKVHIKRPPTGLLMYTLATRFCKQIYLYGFWPFPLDQNPVKYHYDLSLKYGYTSQASPHM
PLEFKALKSLHEQGALKLTVGQCDGAT

>Danrer_GT29_ST8Sia3_NUC : 1122 bp

ATGGTTCGGGTGGTTCAGTGTTCGGGGCTGGTGTATGTTTCAGCGTTGCTTTATTAATTTTGTTCATTTGATCAGCTACGCTCCATCAAGAAGGACTTCATCTTACCAGCCCCGAAATA
CGCCAACGCCGGCGGCCCGCGGATGTACATGTTCCACGCGGGATTTTCGGTCCCAACTGGCCATGAAGTTTCTGGATCCTGCGTTCACATCGCTGAATACTGCGCTCAACGAGAACC
TGCAGGAGTCCAGCAACTGGAGGTTCAACAGGAGCGCTTATGCAGAGCTCAGTAAAGAGATCGCGCAGCATATTGACGTGCCGCACAACCTCACCTGACCAAGAACAGCGTCCGC
GTGGGTTCAGTTAATGCACCTACGACTACTCCAGTCAAAAATACGCTTCTCCATTTGGGGAGAACCCTCAGGTCTCTGCTTCCGGATGCTTCTCCCGTCCCAACAAACGCTACAACAC
TTGCGCGGTAGTGGGAAACAGCGGGATCCTCACCGGCAGCCGATGTGGTCTTCAAATCGACAAAATATGACTTTTGTCTTCCGCTGTAATTTTCGCGCCCACCGAAGTCTTCCGTCGAG
ACGTGGGGCGCCGGACCAACCTCACACCTTTAAACCCAGTATTTCTGGAGAAATACTACAACAACCTGCTGACCATCCAGGACAGGAACAACCTTCTTCCCTCAGTTTGAAGAAGTTG
GACGGCGCCATCTTATGGATCCCGGCTTTTCTTTCACACGCTTGCCACAGTGACGCGAACACTTGTTCGATTTCTTTGTTCGAGCACAGGGGGCAGCTGAAAGTACAGCTAGCCTG
GCCGGGAAACATCATGCAGTATGTCAACAGGTATTTGGAAAACGAAGCAGCTTTCTCCGAAGCGTCTGAGCACGGGGATCCTGATGTTACGCTGGCCCTCGTCTCTGTGTGAACAGG
TGCATTTGTACGGATTTCTGGCCCTTCGGCTGGGACCCCAACACCGGCAAGGAGCTGCCGTATCATTACTATGACAAGAAGGGCACCAAAATTCACCACCAAGTGGCAGGAGTTCGCAC
CAGCTGCCACCCGAGTTCAAGCTGCTCTTCAAGATGCACGCAGACGGAGTCTCAAACCTGAGCCTGTCCCCTGCGCC

>Danrer_GT29_ST8Sia3 : 374 aa ($\Delta 83$)

MVRVSVLGLVMFVSALLILSLISYVSIKKDFIFTAPKYANAGGPRMYMFHAGFRSGLAMKFLDPAFTSLNLTALNENLQESSNWRFNRSAYAELSKEIAQHIDVPHNFTLTKNSVR
VGQLMHYDYSSHKYVFSIGENLRSLLPDASPVLNKRYNTCAVVGNSGILTGSRCPGPEIDKYDFVFRCNFAPTEVFRRDVGRRTNLTFNPSILEKYNNLLTIQDRNNFFLSLKKL
DGAILWI PAFFFHTSATVTRTLVDFFVEHKGQLKVQLAWPGNIMQYVNRWYKTKQLSPKRLSTGILMFTLASSLCEQVHLYGFWPFWDPNTEGKELPYHYDYDKKGTFTTKWQESH
QLPTEFKLLFKMHADGVLKLSLSHCA

>Cormar_GT29_ST8Sia4_NUC : 1074 bp

CGGAGAAATGCAAAAATCTGCAATATGGCATTGAGTCTTTCTCTAACAAAACGAAAGGGTATTCAGAGAACGACTACCTTCAGATTATCACAGATATACAGAGTTGTCCATGGAAA
CGGCAAGCAGAAGAATATGCAAATTTTAGAGCCAAACTTGCTTCCTGCTGTGATGCTGTTCAAACCTTTGTTGTTTTCTCAGAATAACACTCCAGTTGGGACTAATATGAGTTACGA
GGTGGAAAGCAAAAAAGAAATCCCAATTAAGAAGAACATTTTTTCATATGTTTTCCAGTGTCCCAGCCTTTTGTGGACTACCTTATAATCAGTGTGCAGTGGTCGGAAATGGGGGAA
TTCTGAATAAGTCTCTCTGTGGAACCTGAAATAGATAAAATCCGACTTCGTTTTTTAGGTGTAACCTACCCCAACCACAGGAGATGTTAGTAAAGATGTTGGCAGTAAAACAAATCTT
GTGACTATAAATCCAAGCATCATAACTCTGAAATATGGGAACCTAAAGGAAAAAAAAGCCCTATTTCTGGAGGACATTGCAACCTATGGAGATGCATTTTTTCTTCTGCCAGCATT
TTCTTCAGGGCCAACACGGGTACCTCTTTCAAAGTATACTACACGCTCGAAGAGTCTAAAGCAAGACAAAAGGTTCTATTTTTTCCATCCCAAGTACCTGAAAGATCTGGCCCTTT
TCTGGAGAACTAAAGGTGTGACTGCATACCGCTTGTCCACCGGCTTGATGATCACAAGTGTGTCAGTGGAACTGTGTAATAATGTGAAGCTGTATGGATTTCTGGCCCTTCTCTAAA
ACTGTAGAAGACATACCTGTCAGCCATCACTATTATGACAACAAGCTACCTAAACATGGTTTCCATCAGATGCCCAAAGAATACAGCCAGATCCTCCAACCTCACATGAAAGGAA
CCTCAAACCTGCAATTTAGCAAATGTGAAGTCGCC

>Homsap_GT29_ST8Sia6 : 398 aa (Δ 27)

MRPGGALLALLASLLLLLLLLRLLWC PADAPGRARILVEESREATHGTPAALRTLRS PATAVPRATNSTYLNEKSLQLTEKCKNLQYGIESFSNKTKGYSENDYLQIITDIQSCPWK
RQAEYANFRAKLASCCDAVQNFVVSQNNTPVGTNMSYEVESKKEIPIKKNIFHMFVPSQPFVDYPYNQCAVVGNGGILNKSLCGTEIDKSDVFRCNLPPTTGDVSKDVGSKTNL
VTINPSIITLKYGNLKEKKALFLEDIATYGD AFFLLPAFSFRANTGTSFKVYYTLEESKARQKVLFFHHPKYLKD LALFWRTKGV TAYRLSTGLMITSVAVELCKNVKLYGFWPF SK
TVEDI PVSHHYDNKLPKHGFHQMPKEYSQILQLHMKGILKQLFSKCEVA

>Danrer_GT29_ST8Sia8_NUC : 1080 bp

ATGCGCGTTATGAGGACTCTTATGCGCTGGCTGCTTCCTGTATCCTTCTGTGCTCTTTCTGCAGTGTGGCCTTCTGGATTTTCATCTCCAACAATAATGTAATCCCTCATCCTGC
ATCTAGGATTCCTCAAAAAGCCTCCAACACACAATCTTGTAAAGCATGCAAAGACAGTGTAAATCATTGGCAAAGCAGTGGGAAACTATTCAAACAGCTGGAAAAACACGAGGCAA
ACTATAAAAGATTACAGGTTATTACTAAATGAGAAATGCCATGCTGTGTCTAAAGCTGTGGTCCACACAGAATAACACTCCTCTGGGGTCAAATGTTGTCTATGATGGCGAAAGA
AAGCCACTACAGGTGACCCAAGCCCTGTACAACATCTTGGCTAAGGAACAACCATTTGGAAATGCGACATGGGAGTCATGTGCTGTAGTGGGGAATGGAGGCGTTTTGGCCAAATAG
CAGTTGTGGAGAAGAAATCAATTCAGCCCAGTTTGTCTATTAGATGCAACCTTCCACCCTTGTATGACAGATATGAGAAAGATGTGGGAAACAAAACCAATCTTGTACAGCAAAATC
CCAGCATCCTACATGAGAAGTACAGTGGTCTTATGGAGCGGCGGCTCCATTTGTGGAAAGCCTTATTCATACGGTCAAGCCTTATTACTGCTGCCAGCCTTCTCCTACGGTTCAC
AACACGCCTGTATCCCTGCGTGCCTTTTACACACTGGAGGACTTTGGCAGAGAGGGCCCGCTGCCATTTTTCTGAATCCAGAATACTTGAGAAAAGTTGACAAAAGTTTTGGCGCA
GCAAGGCCTGAACTCTGTTCGCCCCAGCACTGGGCTCATCATGGCCAGCTTGGCTCTAGAAATCTGTACTAACGTTTATTTGTATGGTTTTTGGCCCTTTGGTAAACATCCGAACG
ACAGTCGACCCATCACCAACCATATTACGATAATCGGGAGAGCAAGAAAAATGTGCATTCAATGCCGAGCGAGTTCGAACAATTACTGAAACTCCATAAGCAAGGTGTGGTTCAC
ATACACCTGGGAGAATGTCAACCCGCTCATAGATGA

>Danrer_GT29_ST8Sia8 : 359 aa (Δ 33)

MRVMRTLMRWLLPVILLCSFCSVAFWIFISNNNVIHPASRI PQKASNTQSKACKDSVIIGKALGNYSNSWKKHEANYKRFRLLLNEKCHAVSKAVVTQNNTPLGSNVVDGERR
KPLQVTQALYNILAKEQPFNGATWESCAVVGNGGVLANS SCGEEINSAQFVIRCNLPPLDDRYEKDVGNKTNLVTANPSILHEKYSGLMERRRPFVESLHSYGQALLLLPAFSYGH
NTPVSLRAFYTLED FREGPLPIFLNPEYLRKLT KFWREQGLNSVRPSTGLIMASLALEICTNVHLYGFWPF GKHPNDSRPI TNHYDNRESKKNVHSMPSSEFEQLLKLHKQGVVH
IHLGECQPAHR

>Danrer_GT29_ST6Gal1_NUC : 1455 bp

ATGGATCGTGTTAGCCTACTTTGGCGATTGAGGCGTCGGGCTCGTTCGAGGGGCCCTATGCATGGCACTCTTCTGTTTACTATGGCTCTCCTGTACTCTTTGTGCTGAAAACAG
CATTCCTGTACAGATGCCATTTTCGGAGTCAAGGCCAGAACTCGAGCTCAACCAAGAGCACACACAATGTTAAAGTATTGCGGGGAACAGGAGGGTCAAAGCCATGTATACTG
ACCCTCAAAGCTCCCTGGTGTCTTCTGGAGATCCTCAAAGCCTATTTCCATCCTCTCATCATCTAACACTCAATGGAGAGCACTTCTAAAGACCACAGTCTCGGGGGCAA

AAGAGAGAGCGTGGGCTCTTTTACTGGCTTCTGGCACAACCACTGACTATTTTTGGAGGCAGACGGAGAGGAGATATGGGGACCGCAATTCGGGATGCTGATGTTTTTAAGCCTAA
TGGAGCACTCGGAGAGGTCTGGAATGAAGAGATGTCGAGCAGCATGCTGGGAAAAAGGCTGAGGAAAGTTGTGCAGAACTATCAGGCAATGAATAAATATGGAGTCAAGTACCCTG
CAACAGTGAGAGCAGCACACCGCTACAAACTCAGTGGACCAGAAATCTTTTGTGAAATTAAGGAGAAAGTTCAGGTCACAACCTTAAACACCAGATATGGAGCCATTTCTGGGTTT
CCATGGGGCTCTCAGTTACCTCCACGCCAAATAACCTCTGATGTCGGCCCGTTCAAACCTGTGCAGTGGTGTGCTCTGCAGGATCTCTCAAAAACCTCTGGATTAGGAAAGGAAAT
AGATTCTCATGACGCAGTGATACGATTCAATGCCGCACCGACAGCTGGTTTTGAGACAGACGTGGGCTCCAAAACAACAGTGCAGCTTATTAATTCACAGTTAATGGCATCTGAAG
ACCACCATTTCTCTCCAGCTCTTTGTACAGTGCAGGGATTTTAGTGAGCTGGGATCCTAGCCCTTATTCTTCAGATCTATGGGAGTGGTTCAACAAGACAGATTACCCCATATTT
AAACAGTATCAGCGATATCGCCGCTTGCACCTCAGCAGCCATTTTACATTTGTACATCCGCGAATGGAATGGCAGCTGTGGCAGAGAATACAAGACAACATGGGTGAAGCCATCCA
AAAAATCCACCTCATCAGGCTGTGGGCACAGTTTTAATGATGTCCCTGTGTGAAGTGGTACACGTGTACGAGTTCCTGCCATCGCGCGGAAAAACCGAGCTGTGCCACTACT
ACCAGCGCTCAGTGATGCTGCCTGCACACTCGGAGCCTACCACCATTTACTCTATGAGAAAAACCTGGTCAAGAGGATGAACCAAGGCTCCGACCGAGATATCTACACACTTGGC
AGGGTCACTCTTCTGGGTTTCGCCACATTTAACTGCCTTCCAGCACATTCAAAAACGTAA

>Danrer_GT29_ST6Gal1 : 484 aa (Δ 135)

MDRVSLRWRLRRRARRGALCMALFCLTMALLYTLCAENSI PVTDAIFGVKARTRAQ PRAHTIVKVL RGTGGSKPMY TDPQKLP GVI PGDPQKPIPI LSSSNYSME STSKDHSLGGK
KRERGLFYWLLAQPLTIFGRRRRGDMGTAIRDADVFKPNGALGEVWNEEMSS SMLGKRLRQV VQNYQAMNKYGVKY PATVRAAHRYKLSGPEILCEIKEKVQV TLTLPDMEPFSGF
PWGSQ LPPRQITSDVGPFKTCAVVS SAGSLKNSGLGKEIDSHDAVIRFNAAPTAGFETDVGSKTTVRLINSQLMASEDHHFLSSSLYSAGILVSWDPSPYSSDLWEWFNKTDYPIF
KQYQRYRRLHPQQPFYIVHPRMEWQLWQRIQDNMGEAIQKNPSSGLLGTVLMMSLCEVHVHYEFLPSRRKTELCHYYQRFSDAACTLGAYHPLLYEKNLVKRMNQGS DRDIYTLG
RVTLPGFATFNCTSSSTHKT

>Homsap_GT29_ST6Gal1_NUC : 1221 bp

ATGATTACACCAACCTGAAGAAAAAGTTCAGCTGCTGCGTCCCTGGTCTTTCTTCTGTTTGCAGTCATCTGTGTGTGGAAGGAAAAGAAGAAAGGGAGTTACTATGATTCCCTTTAA
ATTGCAAACCAAGGAATTCAGGTGTAAAGAGTCTGGGAAAATTGGCCATGGGGTCTGATTTCCAGTCTGTATCCTCAAGCAGCACCCAGGACCCCCACAGGGGCCCGCAGACCC
TCGGCAGTCTCAGAGGCC TAGCCAAGGCCAAACCAGAGGCCCTCTTCCAGGTGTGGAACAAGGACAGCTCTTCCAAAACCTTATCCCTAGGCTGCAAAAGATCTGGAAGAATTAC
CTAAGCATGAACAAGTACAAAGTGTCTTACAAGGGGCCAGGACCAGGCATCAAGTTCAGTGCAGAGGCCCTGCGCTGCCACCTCCGGGACCATGTGAATGTATCCATGGTAGAGGT
CACAGATTTTCCCTTCAATACCTCTGAATGGGAGGGTTATCTGCCAAGGAGAGCATTAGGACCAAGGCTGGGCCCTGGGGCAGGTGTGCTGTTGTGTCGTGACGCGGATCTCTGA
AGTCTTCCCAACTAGGCAGAGAAATCGATGATCATGACGCAGTCTTGGGTTTAAATGGGGCACCCACAGCCAACCTTCCAACAAGATGTGGGCACAAAAACTACCATTCGCCTGATG
AACTCTCAGTTGGTTACCACAGAGAAGCGCTTCCCAAAGACAGTTTGTACAATGAAGGAATCCTAATTGTATGGGACCCATCTGTATACCACTCAGATATCCCAAAGTGGTACCA
GAATCCGGATTATAATTTCTTTAACAAC TACAAGACTTATCGTAAGCTGCACCCCAATCAGCCCTTTTACATCCTCAAGCCCCAGATGCCCTTGGGAGCTATGGGACATTTCTCAAG
AAATCTCCCCAGAAGAGATTCAGCCAAACCCCCATCCTCTGGGATGCTTGGTATCATCATCATGATGACGCTGTGTGACCAGGTGGATATTTATGAGTTCCCTCCCATCCAAGCGC
AAGACTGACGTGTGCTACTACTACCAGAAGTCTTTCGATAGTGCCCTGCACGATGGGTGCC TACCACCCGCTGCTCTATGAGAAGAATTTGGTGAAGCATCTCAACCAGGGCACAGA
TGAGGACATCTACCTGCTTGGAAAAGCCACACTGCCTGGCTTCCGGACCATTCACTGCTAA

>Homsap_GT29_ST6Gal1 : 406 aa (Δ 56)

MIHTNLKKKFSCCVLVFLFLFAVICVWKEKKGSSYYDSFKLQTKEFQVLKSLGK LAMGSDSQSVSSSSTQDPHRGRQTLGSLRGLAKAKPEASFQVWNK DSSSKNLI PRIQKIWKNY
LSMNKYKVS YKGP GPGIKFSAEALRCHLRDHVNVSMVEVTDFFPNTSEWEGYLPKESIRTKAGPWGRCAVVS SAGSLKSSQLGREIDDHDAVLRFN GAPTANFQQDVGKT TIRLM
NSQLVTTTEKRFLKDSLYNEGILIVWDPSVYHSDI PKWYQNPDYNFNNYKTYRKLHPNQPFYIILKPQMPWELWDILQEISPEEIQPNPSSGMLGIIIMMTLCDQVDIYEF LPSKR
KTDVCCYYQKFFDSACTMGAYHPLLYEKNLVKHLNQGTDEDIYLLGKATLPGFRTIHC

>Homsap_GT29_ST3Gal1_NUC : 1020 bp

ATGGTGACCCTGCGGAAGAGGACCCTGAAAGTGCTCACCTTCCTCGTGCTCTTCATCTTCCTCACCTCCTTCTTCCTGAACTACTCCCACACCATGGTGGCCACCACCTGGTTCCC
CAAGCAGATGGTCCCTGGAGCTCTCCGAGAACCTGAAGAGACTGATCAAGCACAGGCCTTGACCTGCACCCACTGCATCGGGCAGCGCAAGCTCTCGGCCTGGTTCGATGAGAGGT
TCAACCAGACCATGCAGCCGCTGCTGACCGCCAGAACCGCTCTTGGAGGACGACACCTACCGATGGTGGCTGAGGCTCCAGCGGGAGAAGAAGCCCAATAACTTGAATGACACC
ATCAAGGAGCTGTTTCAGAGTGGTGCCTGGGAATGTGGACCCATATGCTGGAGAAGAGGTCCGGTGGGCTGCCGGCGCTGCGCCGTTGTGGGCAACTCGGGCAACTGAGGGAGTCTTC
TTATGGGCCTGAGATAGACAGTACGACTTTGTCCCTCAGGATGAACAAGGCGCCACGGCAGGGTTTGAAGCTGATGTTGGGACCAAGACCACCCACCATCTGGTGTACCCCTGAGA
GCTTCCGGGAGCTGGGAGATAATGTCAGCATGATCCTGGTGCCTTCAAGACCATCGACTTGGAGTGGGTGGTGGAGCGCCATCACCACGGGCACCATTTCCACACCTACATCCCCG
GTTCCTGCAAAGATCAGAGTGAAACAGGATAAGATCCTGATCTACCACCCAGCCTTCATCAAGTATGCTTTTGACAACTGGCTGCAAGGGCACGGGCGATACCCATCTACCGGCAT
CCTCTCGGTATCTTCTCAATGCATGTCTGCGATGAGGTGGACTTGTACGGCTTCGGGGCAGACAGCAAAGGGAACTGGCACCCTACTGGGAGAACAACCCATCCGCGGGGGCTT
TTCGCAAGACGGGGGTGCACGATGCAGACTTTGAGTCTAACGTGACGGCCACCTTGGCCTCCATCAATAAAAATCCGGATCTTCAAGGGGAGA

>Homsap_GT29_ST3Gal1 : 340 aa (Δ56)

MVTLRKRRLKVVTFVLVFI FLTSFFLNYSHTMVATTWFPKQMVLELSENLRLIKHRPCTCTHCIGQRKLSAWFDERFNQTMQPLLLTAQNALLEDDTYRWWLRLQREKKPNNLNDT
IKELFRVVPGNVDPMLEKRSVGCRRCAVVGNSGNLRESSYGPEIDSHDFVLRMKNKAPTAGFEADVGTKTTHHLVYPESFRELGDNVSMILVFPKTI DLEWVVSATTTGTI SHTYIP
VPAKIRVKQDKILIIYHPAFIKYVFDNWLQGHGRYPSTGILSVIFSMHVCDEVLDYGFADSKGNWHHYWENNP SAGAFRKTGVHDADFE SNTATLASINKIRIFKGR

>Danrer_ALCAM_NUC : 1668 bp

ATGCATTCGGTTATCTGCCTTTTCGGTGCCTTCATAGCTGCCGCTGTGTTTGCTCCAGGGAGCTGCCCTGCCGACGGTTATAGGTCTGTACGGTGGAGCCATCGAAGTGCCATGCAA
CAATGGAAATAACAAGCCAGATGGCCTTATTTTACCCTAAATGGAAATATGTGAAAGATGACGGCTCCCTGGCGATCTACTAATAAAGCAGGCACAGAAAGATGATCCGACTGTGT
CTGCCATGGACGGCTACAAAAC TAGAGTTAGCATCGCTGCTAATTCAGCTTACTGATTGCCAGGGCTCTTTGACTGACCAGAGAGTCTTACCTGCATGGTGGTGTCTTCAACT
AATCTGGAGGAATTCCTCTGTGGAAGTTAAAGTTACAAAAAACCATCAGCCCCGTGTAATCAAAAAACAAAGTAAAAGAACTGGAAAATGGCAAACCTGACGCAGTTGGGGGAATGTGT
GGTGGAGAGCGCCAACCCAGCAGCAGATCTCATCTGGATGAAGAACAACCAGGCTCTGGTGGATGATGGCAAGACGATCATTATCACATCAGATGTCACCAAGGACCCAGTCAACCG
GTCTGTCCAGCACCTCTTCCAGACTGCAGTACACAGCAAGGAAAGAGGATGTGGCATCACAGTTCACCTGCGTTGCAAAGCACGTGACGGGACCCAAACAGGTTTCAACACCCGAT
ACCTTCCAAATTCGCTATCCCACTGAGAAGGTGAGTCTACAGGTTGTCTCTCAGAGCCCCATTAGGGAAGGTGATGATGTGACTCTGAAATGCCAGGCGGACGGAAACCCCTCCTCC
TACCAGCTTCAACTTTAACAATTAAGGGAAAGAAGGTACGGTGACGGACAAGGATGTCTACACACTGACCGGCGTACCCGAGCCGACAGCGGTGTGTACAAGTGCTCTCTGCTTG
ACAACGATGTGATGGAGTCCACTCAGATCGTCACAGTGAGCTTTCTGGATGCAAGCCTCACTCCTACAGGCAAGGTGTTAAAAAAGCTCGGGGAAAACCTGGTAGTGTCTTTGGAG
AGAATGCCCTCTTCTGAAGTAAAAGTGACGTGGACTAAGGATAAACCCTGAACTGGACAAACTGCCGTGATTTCTCCAGTTGAGATACAGCGACGCGGGCTTATACGTGTGTGATGT
GTCCATTGAAGGAATCAAACACAGCTTTTCCCTCAGACTTACTGTGGAAGGTGGTCCAAGAATTACCGCCTGACAAAGCATCGCAGCAATGACGGAAAACACAAAGTGTGACGT
GCGAGGCAGAAGGTTACCTAAACCTGAAGTGCAGTGGAGTGTCAACGGAACCGACGATGAAACATCGTATGTCAACGAAAAGCCACATACAAACTGACGGTGGTCCCGAGTAAG
AACCTCACCGTACAGCTGCCTCGTGACCAATAAACTCGGTTTTCGACACAAAGGACATCAGCGTGTTCCTCCGAAAAAATGAAGATGGCGCAGACCAGGCCAAAGTGATTTGGGGTGT
CGTGGTTGGACTGTTTCTAGCCGCTGCCCTGGTGGGACTCATCTACTGGTTGTATATCAAGAAAACAAGACAAGGCAGCTGGAAGACCGGAGAGAAGGAGACTGGCACTTTCAGAGG
AGAGTAAGAACTGGAGGAGAACAATCATAAAGCAGATGTCTAA

>Danrer_ALCAM : 555 aa

MHSVICLFGAFIAAAVFPAGSCLPTVIGLYGETIEVPCNNGNNKPDGLIFTKWKYVKDDGSPGDLLIKQAQKDDPTVSAMDGYKTRVSI AANSSLLIAQGSLLTDQRVFTCMVVSST
NLEEF SVEVKVHKKPSAPVIKKNVKELENGKLTQLGECVVESANPAADLIWMKNNQALVDDGKTI I I TSDVTKDPVTGLSSTSSRLQYTARKEDVASQFTCVAKHVTPGNQVSTPD
TFQIRYPTEKVS LQVVSQSPIREGDDVTLKQADGNPPPTS FNFNIKGGKVTVDKDVYTLTGVTTRADSGVYKCSLLDNDVMESTQIVTVSFLDASLPTGKVLKKGLENLVVSL E
KNASSEVKVTWTKDNRKLDKLPDFSQLRYS DAGLYVCDVSI EGIKHSFSFELTVEGGPRIITGLTKHRSNDGKHKVLTCEAE GSPKPEVQWSVNGTDDETS YVNGKATYKLT VVPSK
NLTVSCLVTNKLGFDTKDISVFSRKNEDGADQAKVIVGVVGLFLAAALVGLIYWLYIKKTRQGSWKTGEKETGTSEESK KLEENNHKAD

PERMISSION TO REPRODUCE FIGURES

All the figures coming from scientific publications and reproduced in this thesis manuscript obtained, when required, permissions to be reused. Fig. 3, 14, 18, 19, 20, 23, 24 (MDPI), fig. 6 (Springer Nature), fig. 30 (ChemBioChem) were taken from Open Access publications and required any special permission to be reused. Fig. 5 was partly generated using Servier Medical Art, provided by Servier, licensed under a Creative Commons Attribution 3.0 unported license. Fig. 7, 27 (Elsevier), fig. 10 (Frontiers), fig. 31 (ACSPublications) were taken from publications under Creative Commons Attribution License (CC BY) and required any special permission to be reused for non-commercial use. Fig. 15, 16, 25 and 26 have required specific permissions to be reused from Springer Nature as reported below.



License Number	5547590749537	5547590095873	5547581090053	5547590979849
License Number	May 14, 2023	May 14, 2023	May 14, 2023	May 14, 2023
Licensed Content Publisher	Springer Nature	Springer Nature	Springer Nature	Springer Nature
Licensed Content Publication	Nature Structural & Molecular Biology	Nature Structural & Molecular Biology	Nature Structural & Molecular Biology	Springer eBook
Licensed Content Title	Crystal structure of the polysialic acid-degrading endosialidase of bacteriophage K1F	Structure of human ST8SiaIII sialyltransferase provides insight into cell-surface polysialylation	Structural insight into mammalian sialyltransferases	Endosialidases: Versatile Tools for the Study of Polysialic Acid
Licensed Content Author	Katharina Stummeyer et al	Gesa Volkers et al	Francesco V Rao et al	Elina Jakobsson, David Schwarzer, Anne Jokilammi et al
Licensed Content Date	Dec 19, 2004	Jul 20, 2015	Oct 11, 2009	Jan 1, 2012
Type of Use	Thesis/Dissertation	Thesis/Dissertation	Thesis/Dissertation	Thesis/Dissertation