OPEN ACCESS Freely available online

Journal of Physical Chemistry & Biophysics

Research Article

Hypothetical Pathways on GlycoRNA Biochemistry

Santiago Alonso Gil*

Department of Structural and Computational Biology, Max Perutz Labs, University of Vienna, Campus Vienna Biocenter 5, 1030, Vienna, Austria

ABSTRACT

GlycoRNAs are recently discovered biomolecular structures whose medical relevance is key to understanding cell-cell communication. Research fields such as immunology and tumor research are developing theories around these novel conjugates located on the cell surface. The work of Bertozzi is a reference where the experiments show relevant physicochemical properties of the sialylated RNA conjugates. Going through the current enzymatic and chemical available data, in this work, we hypothesize on the possible biochemical pathways leading to N-glycosylation of RNA nucleobase. Based on the reactivity of glycoRNA with oligosylsaccharide transferases and the PNGase F hydrolase, we base our hypothesis on the conjugation of modified versions of the canonical nucleobases and the GlcNAc moiety through an amide-functionalized small linker. We assume that every nucleobase (A, C, G, and U) is modified following a different enzymatic pathway, but purine bases and pyrimidine bases share the same link to the anomeric carbon of the sugar (Ca), respectively: C_7-CONH-Ca and C_5-CH_2-CONH-Ca.

Keywords: GlycoRNAs; RNA modifications; N-glycosylation; Biochemical structures; Glycosidases; Glycosyl transferases

INTRODUCTION

Since the discovery of the N-linked glycosylated RNAs (glycoRNAs) [1], the scientific community has focused its attention on the potential biomedical relevance of these molecular structures. From that moment, N-glycosylation became part of the select group of post-transcriptional modifications (phosphorylation, methylation, and acetylation) of RNA. As a novel epitranscriptomic modification, RNA glycosylation shows a relevant role in membrane localization and cell-cell communication [2]. One of the most interesting hypothetical functionality of glycoRNAs is their paper on neuron-microglia interaction after an ischemic stroke: The sialylated neuronal Y-RNAs are detected by the microglial Siglec-11 receptor producing an anti-inflammatory phenotype [2]. As it is well-known, sialic acids show multiple functions in tumor biology, such as immune evasion, proliferation and metastasis, promoting angiogenesis or resisting apoptosis and therapy [3]. GlycoRNAs may also be involved in those processes. These novel conjugates opened a new paradigm in biology, and at the same time, many incognita are still in the air. What do we know, and what do we ignore about the chemical structure of glycoRNAs?

The exhaustive work directed by Bertozzi showed relevant discoveries which helped to decipher the chemical nature of the link between RNA and carbohydrates [1]:

a) GlycoRNA production is contributed by the canonical N-glycan biosynthetic machinery: A dose-dependent loss of glycoRNA is caused by the inhibition of Oligosaccharyl Transferase (OST) with NGI-1. Furthermore, this result is also observed when the system is treated with kifunensine and swainsonine (α-mannosidases I and II inhibitors).

b) The resulting glycoRNAs show sensitivity to PNGase F, Endo F2 and Endo F3 hydrolases: Given the specificity of OST and PNGase F, the RNA- carbohydrate bond must be an amide bond-containing linker.

c) Treating the glycoRNAs with Endo Hf, O-glycosidase, and StcE mucinase shows no effect: Confirming the exclusive N-glycosylation of the obtained glycoRNAs and the absence of high-mannose glycans.

d) Siglec receptors and anti-RNA antibodies recognize all surface glycoRNAs: The binding of Siglec 11 and 14 showed sensitivity to RNAse A, meaning that cell surface glycoRNAs could be direct Siglec receptor ligands. In conclusion, these glycoconjugates are sialic acid-capped and confirm the covalent conjugation between RNA and carbohydrates.

e) Observed sedimentation of glycoRNAs in the sucrose gradient suggests a relatively small molecular weight for the RNA-saccharide linker.

Correspondence to: Santiago Alonso Gil, Department of Structural and Computational Biology, Max Perutz Labs, University of Vienna, Campus Vienna Biocenter 5, 1030, Vienna, Austria, E-mail: santiago.alonso.gil@univie.ac.at

Received: 04-Aug-2023, Manuscript No. JPCB-23-25939; **Editor assigned:** 07-Aug-2022, PreQC No. JPCB-23-25939 (PQ); **Reviewed:** 21-Aug-2023, QC No. JPCB-23-25939; **Revised:** 28-Aug-2023, Manuscript No. JPCB-23-25939 (R); **Published:** 04-Sep-2023, DOI: 10.35248/2161-0398.23.13.359.

Citation: Gil SA (2023) Hypothetical Pathways on GlycoRNA Biochemistry. J Phys Chem Biophys. 13:359.

f) Limitations: The selective metabolic labelling of sialic acids with Ac4ManNAz concludes with the possibility of other RNA-glycans conjugations. Furthermore, the precise chemical structure of the linkage is still unknown.

Extra biostatistical analysis of the experimental data by Cui et al. [4], also showed:

g) The positional sequence pattern and the RNA-binding protein-related motifs are located following a non-random distribution: The sequence context of RNA glycosylation sites was used as a predictive tool for RNA glycosylation, building the base of the GlyinsRNA webserver [5]. Interestingly, the appearance of guanosine in the glycosylation site is enriched by +30%, while adenosine (-5%), uracil (-15%) and cytosine (-30%) show depletion.

Taking into account the listed data, in this work, we want to focus our attention on the chemical nature of the RNA-glycan linker and hypothesize its chemical structure based on the relevant evidence.

METHODOLGY

Molecular modelling

Computational details: The cluster models were constructed from the available PNGase F complex with di-N-acetylchitobiose (PDB ID 1PNF) [6]. The substrate, water molecules and the side-chain of the residues at 10 Å from the substrate formed the initial cluster model. The two water molecules occupying the glycosidase subsite -1 were deleted, and the Asn-$[GlcNAc]_2$ model was built after manually attaching a β-oriented neutral asparagine (CH_3-CH_2-C=O-NH-) to the anomeric carbon of the sugar in the subsite +1. The cluster model was optimized using Gaussian 16 electronic structure program [7], fixing all the atoms except hydrogens, water molecules and the substrate and described using the available Universal Force Field [8]. The same method was used to optimize the ncm^5U-$[GlcNAc]_2$ and sangivamycin-$[GlcNAc]_2$ cluster models.

Disentangling the glycoRNA structural puzzle

One of the main pieces of evidence is that the formation of the RNA-glycan bond is catalyzed by OST, and its cleavage is catalyzed by the PNGase F enzyme. OST is an enzymatic complex whose substrate recognition depends on a substrate consensus sequence (-Asn-X-Thr/Ser-) [9]. The Peptide-N4-(N-acetyl-β-D-glucosaminyl) asparagine amidase F (PNGase F) converts the asparagine residue to an aspartic acid making this enzyme specific for cleaving amide groups [6]. Thus, before the formation of the amide group, the reactants of OST are a primary amide (X-Asn-$CONH_2$) and a carbohydrate. As for glycosylated proteins, the common reactant in this equation is the glycan, so we should base our hypothesis on finding a nucleobase functionalized with an amide group.

Taking into account that the formation of glycoRNAs uses the same machinery as the N-glycosylation process in proteins, one can speculate on the idea that the nucleobase modification process also follows an already well-known biochemical pathway. But, there is a main difference between protein and RNA glycosylation: The number of molecules that can become glycosylated, one case for proteins (Asn) and four for RNA (A, C, G and U). One may assume that each nucleobase is functionalized differently or, at least, pyrimidine nucleobases will be differentiated from purine bases.

RESULTS

To find a possible modification pathway starting from each of the four canonical RNA nucleobases and resulting in an amide-functionalized derivative, we checked the last version of the MODOMICS database [10]. After intensive literature research, we found three main amide derivatives (Figure 1): 7-amido-7-deazaguanine (ADG, guanine derivative), sangivamycin (adenosine derivative) and 5-carbamoylmethyluridine (ncm^5U, uracil derivative). All the derivatives provide a linker with a relatively small molecular mass. Despite the absence of amide-functionalized cytosine, this nucleobase can be converted in uracil by a tRNA-specific cytidine deaminase (CDAT8) [11]. GTP is converted in ADG employing GCHI, QueD, QueE and QueC (enzymatic pathway) [12]. While ncm^5U is the amide-derivative of the cm^5U, which acts as the reactant of the ALKBH8 tRNA methyltransferase [13]. The enzymatic pathway connecting U → cm^5U → ncm^5U remains unknown. Besides sangivamycin is an adenosine derivative, the known enzymatic pathway resulting in the derivative starts in guanosine. GTP is converted to sangivamycin through three synthetic processes: ToyD, ToyB, ToyC, and ToyM perform the deazapurine synthesis, following the purine salvage is catalyzed by ToyH, ToyE, ToyG, ToyF, and ToyI, and finally, the side-chain modification from toyocamycin to sangivamycin is done by ToyJ, ToyK and ToyL [14].

Assuming that these three derivatives are synthetically available for an organism, the final glycoRNA product of the OSL should be a proper reactant for the PNGase F. Fortunately, a crystal structure of this enzyme in complex with di-N-acetylchitobiose (product complex) where the substrate is located in the sugar subsites +1 and +2 [6]. Using molecular modelling, we were able to reconstruct a cluster model of the N-asparagine-di-N-acetylchitobiose complex (reactant complex, Figure 2A), and we built two more complexes by substituting the asparagine with the ncm^5U (Figure 2B) and sangivamycin (Figure 2C). Both Asn and ncm^5U contain a flexible CH_2 group connecting the amide group with the rest of the chain, and they fit similarly in the glycosidase subsite -1. For sangivamycin (same for ADG), the amide group is directly connected to the nucleobase ring and slightly affects the orientation of the amide in the subsite -1. Taking into account that our protein model is rigid, the author assumes that a slight readjustment of the catalytic residues would be enough to facilitate the hydrolytic reaction.

Hypotheses

After the points discussed previously, the author listed the following hypotheses regarding the chemical nature of the RNA-glycan linker:

a) Each RNA nucleobase (A, C, G and U) goes through a different enzymatic N-glycosylation pathway.

b) The linker for A and G positions (purine bases) is a direct secondary amide connection between the anomeric carbon of a GlcNAc moiety and the C_7 atom of the corresponding modified nucleobase (N7 is changed by carbon, Figure 3-up).

c) The linker for U and C positions (pyrimidine bases) connects the C_5 atom of a uracil base with a CH_2 group connected to the secondary amide connection, and the nitrogen of the amide is connected to the anomeric carbon of a GlcNAc moiety (Figure 3-down).

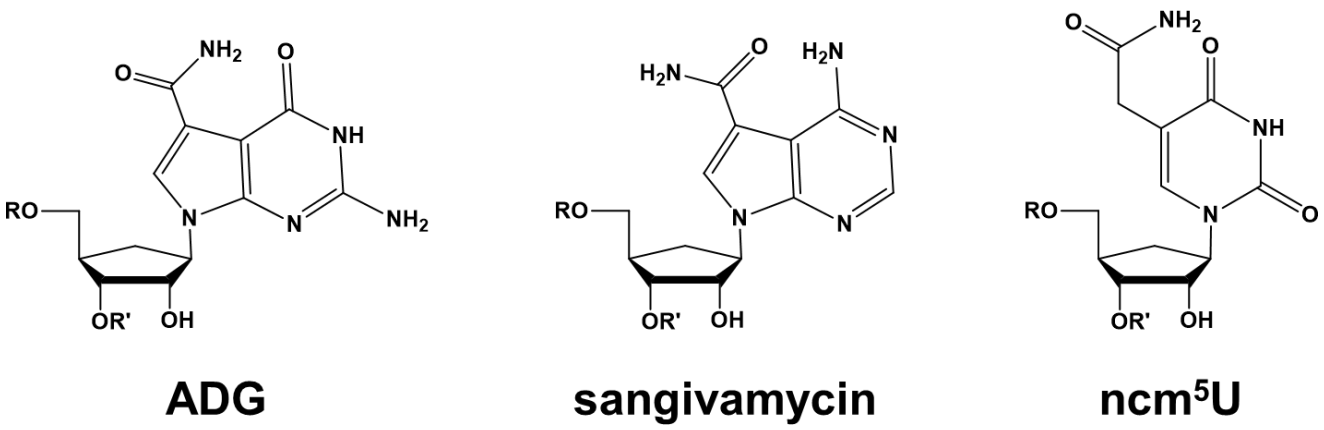

Figure 1: Amide-containing RNA nucleobase derivatives.

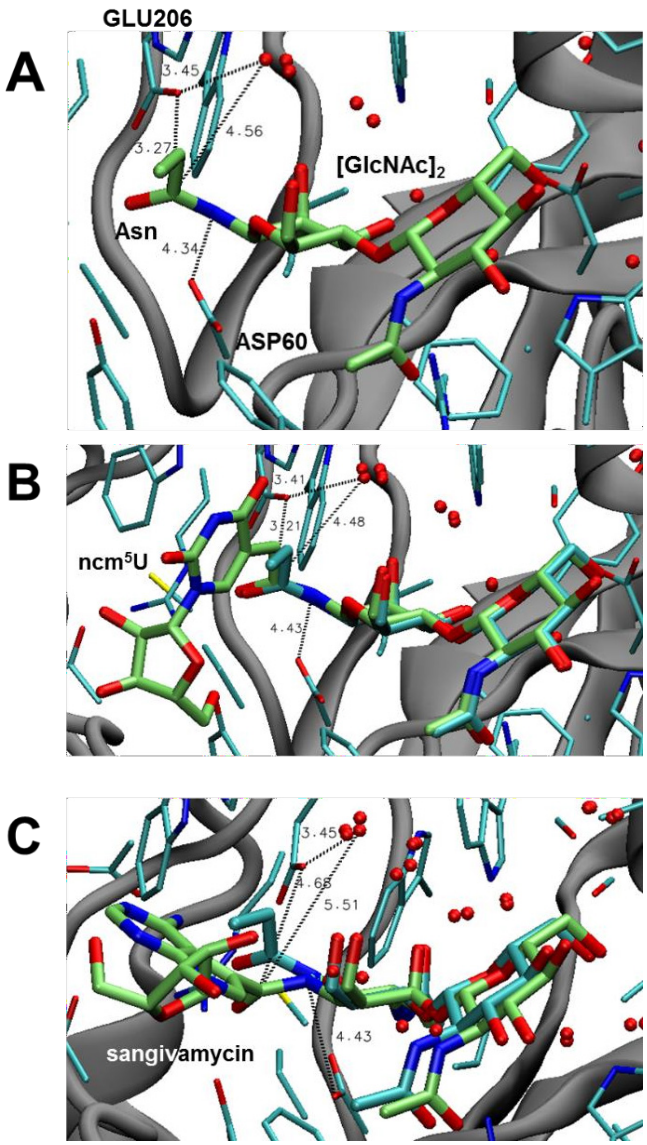

Figure 2: Atomistic representation of (A) the cluster model of the PNGase F-Asn-$[GlcNAc]_2$ complex, Superposition of the PNGase F-Asn-$[GlcNAc]_2$ with (B) the ncm^{5}U-$[GlcNAc]_2$ and (C) the sangivamycin-$[GlcNAc]_2$ models, Glu206 and Asp60 are catalytic residues, Hydrogens are not shown for the sake of clarity.

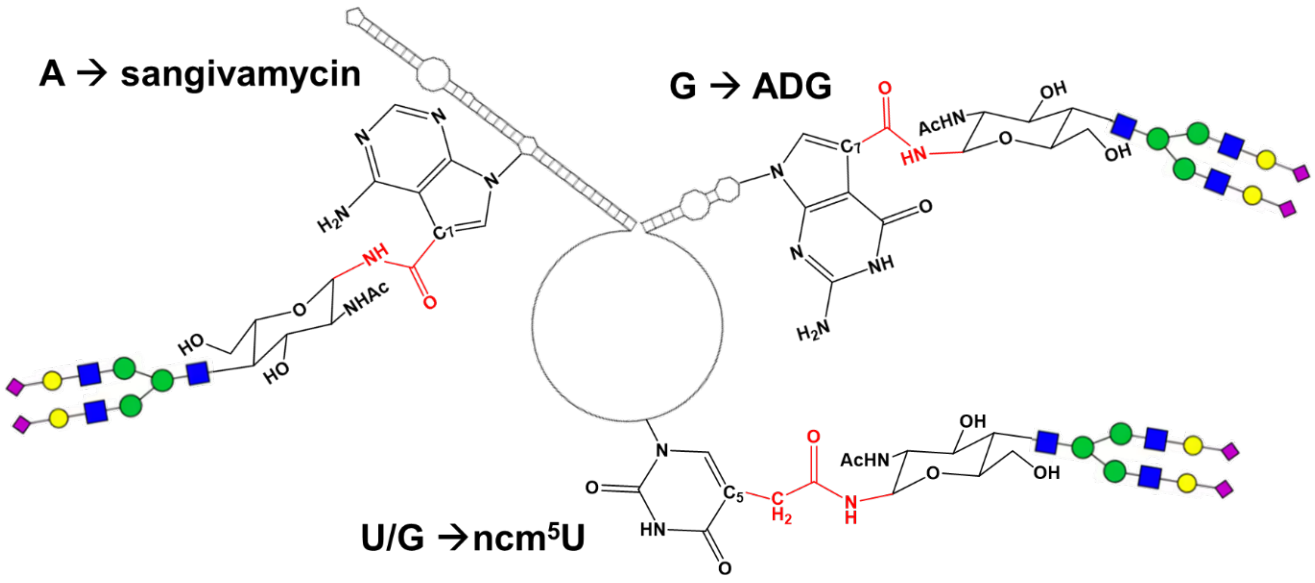

Figure 3: Schematic representation of the hypothetical chemical link between modified nucleobases and sialylated N-glycans. **Note:** (■) GlcNAc, (●) Mannose, (○) Galactose, (◆) Neu5Ac.

DISCUSSION

As it has been described in this work, glycoRNAs have a preferential sequence pattern to overcome N-glycosylation. As in the case of proteins, where OST prefers –Asn-X-Thr/Ser-regions, the same enzyme may have a binding affinity for a certain RNA sequence. Furthermore, as it happens between amino acids and nucleobases [15-17], the intrinsic affinity between a given modified RNA nucleobase and carbohydrates may affect the composition of the preferred N-glycosylation sites. Once the chemical nature of the RNA-glycan linker is deciphered, we are interested in tackling the following works:

a) Answering the question of the effect of the novel modifications on the physicochemical properties of RNA. Following the strategy performed by Zagrovic et al. [15-17], we plan to quantify the intrinsic affinity between modified nucleobases and GlcNAc.

b) Developing a proper force-field to simulate glycoRNAs, properly describing the experimental behaviour of these novel glycoconjugates using classical molecular dynamics simulations.

c) Simulation of the glycoRNA-Siglec 11 structure with accuracy using both classical and QM/MM molecular dynamics methods.

CONCLUSION

As a hypotheses publication, this work does not shed light on any concise conclusion. However, the glycoRNA world is a young field with many interesting open questions. This work may bring inspiration for future experiments and may be used as a review of the current state-of-the-art of the chemical nature of sialic acid-based glycoRNAs.

ACKNOWLEDGEMENTS

S.A-G acknowledges the European Union's Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under the Marie Curie Skłodowska Grant Agreement No. 847548.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

1. Flynn RA, Pedram K, Malaker SA, Batista PJ, Smith BA, 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.
2. Chokkalla AK, Jeong S, Subramanian S, Vemuganti R. Immunomodulatory role of glycoRNAs in the brain. J Cereb Blood Flow Metab. 2023;43(4):499-504.
3. Zhou X, Yang G, Guan F. Biological functions and analytical strategies of sialic acids in tumor. Cells. 2020;9(2):273.
4. Cui C, Wu X, Zhou Y. GlyinsRNA: A webserver for predicting glycosylation sites on small RNAs. RNA Biol. 2021;18(sup2):600-603.
5. Cui C, Wu X, Zhou Y. GlyinsRNA: Prediction of glycosylation sites in small RNA. 2023.
6. Kuhn P, Guan C, Cui T, Tarentino AL, Plummer TH, van Roey P. Active site and oligosaccharide recognition residues of peptide-N4-(N-acetyl-β-d-glucosaminyl) asparagine amidase F. J Biol Chem.1995;270(49):29493-29497.
7. Frisch ME, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, et al. Gaussian 16, revision C. 01. Gaussian. 2016.
8. Rappé AK, Casewit CJ, Colwell KS, Goddard III WA, Skiff WM. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. J Am Chem Soc. 1992;114(25):10024-10035.
9. Yan A, Lennarz WJ. Unraveling the mechanism of protein N-glycosylation. J Biol Chem. 2005;280(5):3121-3124.

10. Boccaletto P, Stefaniak F, Ray A, Cappannini A, Mukherjee S, Purta E, et al. MODOMICS: A database of RNA modification pathways. 2021 update. Nucleic Acids Res. 2022;50(D1):231-235.

11. Randau L, Stanley BJ, Kohlway A, Mechta S, Xiong Y, Söll D. A cytidine deaminase edits C to U in transfer RNAs in Archaea. Science. 2009;324(5927):657-659.

12. Yuan Y, Hutinet G, Valera JG, Hu J, Hillebrand R, Gustafson A, et al. Identification of the minimal bacterial 2′-deoxy-7-amido-7-deazaguanine synthesis machinery. Mol Microbiol. 2018;110(3):469-483.

13. Songe-Møller L, van den Born E, Leihne V, Vågbø CB, Kristoffersen T, Krokan HE, et al. Mammalian ALKBH8 possesses tRNA methyltransferase activity required for the biogenesis of multiple wobble uridine modifications implicated in translational decoding. Mol Cell Biol. 2010;30(7):1814-1827.

14. McCarty RM, Bandarian V. Deciphering deazapurine biosynthesis: Pathway for pyrrolopyrimidine nucleosides toyocamycin and sangivamycin. Chem Biol. 2008;15(8):790-798.

15. Zagrovic B, Bartonek L, Polyansky AA. RNA-protein interactions in an unstructured context. FEBS Lett. 2018;592(17):2901-2916.

16. Hajnic M, Ruiter AD, Polyansky AA, Zagrovic B. Inosine nucleobase acts as guanine in interactions with protein side chains. J Am Chem Soc. 2016;138(17):5519-5522.

17. Hajnic M, Alonso-Gil S, Polyansky AA, de Ruiter A, Zagrovic B. Interaction preferences between protein side chains and key epigenetic modifications 5-methylcytosine, 5-hydroxymethycytosine and N6-methyladenine. Sci Rep. 2022;12(1):19583.