How blood group A might be a risk and blood group O be protected from Coronavirus (COVID-19) infections (how the virus invades the human body via ABO(H) blood group carbohydrates)

by Peter Arend

Abstract.

Although the angiotensin-converting-enzyme 2 (ACE2) is defined as the primary SARS-CoV receptor, it is the history of the amino acid serine, suggesting the actual or additional binding via an intermediate hybrid *O*-glycan: the TMPRSS2 (trans-membrane serine) host protease-mobilized, virus-encoded serine molecule gets access to the host's N-acetyl-D-galactosamine (GalNAc) metabolism and the resulting intermediate, hybrid A-like/Tn structure performs the adhesion of the virus to host cells primarily independent from the ABO(H) blood group. In humans, this genetically undefined structure may be replaced by blood group ABO(H)-specific, mucin-type hybrid epitopes. While the phenotype formation occurs in evolutionary connection with humoral innate immunity, in the blood group O(H), hybrid epitopes become exposed to the highly anti-glycan-aggressive ABO(H) isoagglutinin activities, exerted by the ancestral, nonimmune immunoglobulin M (IgM). In the non-O blood groups these IgM activities are downregulated by phenotypic glycosylation, while the formation of adaptive IgG is excluded by clonal selection. The non-O blood groups thus become a preferred target for the virus, whereas blood group O(H) individuals, lacking the AB phenotype-determining enzymes, have the least molecular contact with the pathogen; they maintain the isoagglutinins and the power of ancestral IgM, considered the humoral spearhead of innate immunity.

Introduction.

When according to the numbers of Wikipedia in countries like Chile, Ecuador, Colombia, Simbabwe and Mexico 59 to 85 percent of the people have blood group O and these countries officially publish extremely low COVID-19 cases and death rates per 1 million inhabitants, this indicates a lower susceptibility of blood group O to the disease, although insufficient investigations and reports cannot be excluded. Moreover, this blood group can no longer be considered a genetic entity.[1] Analysis of genetic statistics from any given contemporary population may be burdened with hidden unusual O alleles including O2, described at the ABO locus, which implicates unexpected phenotypes, involving weak blood group A alleles and further undefined blood group diversities.[2]

The molecular biology of an infection pathogenesis determines the susceptibility of a species to the infection, while the phenotype-determining enzymes decide about the difference between infection and severity of the disease. In the case that *O*-glycosylation plays a key role in the pathogenesis of coronavirus disease, as was discussed already 14 years ago in a SARS-CoV infection[3] and is currently again predicted for SARS-CoV-2 or COVID-19,[4] this would involve the formation of hybrid, serologically A-like, *O*-GalNAcα1-Ser/Thr-R, Tn ("T nouvelle") antigenic structures. Sustaining this prediction, the interaction between the viral spike protein and the host's cellular receptor was inhibited by natural and monoclonal anti-A antibodies *in vitro*.[5]

The adhesion of the virus to host cells primarily occurs independent of the ABO(H) blood group through the genetically undefined intermediate, serologically A-like/Tn evolutionary/developmental structure, which is common to all metazoan growth and/or

ontogenetic processes and apparently acts as a host-pathogen functional bridge in different, unrelated infectious diseases.[6, 7] Thus, although the angiotensin-converting-enzyme 2 (ACE2) protein is defined as the primary SARS-CoV receptor, the actual and/or additional binding between host and pathogen appears to occur via an intermediate hybrid *O*-glycan, dominated by the pathogen's hydrophilic amino acid serine. In the human species this intermediate structure will be elongated and/or replaced by mucin-type ABO(H) phenotype-determining carbohydrates.

How blood group A might be a risk and blood group O be protected from COVID-19 infection.

The virus cannot survive outside of its hosts and hypothetically utilizes the host cell's machinery via hijacking its A-like/Tn formation by serine-rich motifs. Similar suggestions are subject of the recent review, published by Watanabe et al. (2019).[8] Although susceptibility to any infection and its severity depend on uncountable factors and an individual risk for a person cannot be predicted solely on the basis of his blood group, individuals with blood group A could not respond with either acquired or innate antibodies to the synthesis of hybrid A-like structures due to clonal selection and phenotypic, glyosidic accommodation of plasma proteins.[9] The ABO(H) blood group phenotype formation occurs on both the cell surfaces and plasma proteins in evolutionary, developmental connection with innate immunity. Thus, in the human blood group O(H), the ancestral, polyreactive nonimmune or innate immunoglobulin M (IgM), exerting the highly anti-glycan-aggressive anti-A, B and H-isoagglutinin activities, is considered the complementary protein of the syngeneic intermediate A-like/Tn structure and controls the expression and qualities of this structure, whereas in the non-O blood groups it is neutralized through the ABH-

phenotype-determining enzymes. Therefore, blood group A individuals become a preferred target for the virus, which hypothetically mimicks the ABO phenotype pathways.

The virus enters the human body via the human ACE2 receptor protein, which is a polyfunctional protein and, among various other functions, represents the binding domain of the SARS-CoV viruses. In a complex signalling pathway the human ACE2 binds to the spike (S) protein of the virus envelope,[10] and after subsequent cleavage of the ACE-virus S-protein complex by cathepsin L,[11] the virus enters the cell by receptor-mediated endocytosis.[12]

Within the ongoing discussion as to which amino acids dominate the host-pathogen fusion, the most critical molecular step appears to be the mobilization of the viral serine molecule by TMPRSS2 serine protease,[13, 14] and it is the history of this hydrophilic amino acid and its obviously essential involvement in SARS-CoV-2 pathogenesis, which strongly suggest that the binding between pathogen and host occurs by *O*-glycosylation. Although serine-rich repeat proteins (SRRPs) performing the adhesion of different bacteriae to host cell carbohydrates via *O*-glycosylation[15] have not been described for a virus infection pathogenesis, a similar mechanism may take place in SARS-CoV pathogeneses and the observation that the adhesion of the SARS-CoV spike protein to its cellular receptor was inhibited by monoclonal and natural anti-histo-blood A group antibodies[5] should prove the inhibition of the classical *O*-GalNAc binding to a serine molecule. Moreover, the preferential infection or severe disease of an individual with blood group A or other non-O blood groups demonstrates that ABO(H) blood group phenotype-determining sugars are the glycosidic target. Finally, serine residues, preserved on the viral spike protein, become available through the action of the host's transmembrane serine protease TMPRSS2,[14] while the ACE2 receptor protein,

hypothetically codetermined by the ABO phenotype,[16, 17, 18] might mediate the ABO glycan transferring enzyme activity and performing a further (blood group-A-specific mucin-type) hybrid binding. Analogously, the binding to blood group O cells occurs by mucin-type fucosylation via fucosyltranferases 1 and 2 (FUT1/FUT2) activities and performance of a hybrid H-type antigen,[6, 7] which in contrast to A antigen formation does not affect the innate and adaptive anti-A isoagglutinin levels (Fig. 1). Again, within the complex molecular pathogenic process, the most critical molecular step appears to be the mobilization of the viral serine molecule,[13, 14, 19] which gets access to both the host's blood-group independent (ontogenetic) and blood group-specific (A-allelic-encoded) *O*-GalNAcα1-Ser/Thr-R (Tn) formations, whose evolutionary relationship remains unknown. This means, that although the ACE2 protein is defined as the SARS-CoV virus receptor and mediates the transferring enzyme activities, the actual and/or additional binding between host and pathogen appears to be an intermediate hybrid *O*-glycan.

Conclusions.

The proposed concept of virus invasion via glycosidic binding on ABO(H) blood group-determining sugars does not question the established functions of the ACE2 receptor protein in its previous[20] and current definitions[21, 22] but rather shows an additional and more specific interaction between host and pathogen. Both *N*- and *O*-glycosylations may occur within this complex pathogenic process, and among multiple chemical and physicochemical linkage options, both a trans-species ontogenetic, blood group-independent and blood-group specific binding may be performed through intermediate *O*-glycosylations in two different glycosidic steps, dominated by the pathogen's hydrophilic amino acid serine; moreover, the blood

group-specific, mucin-type formation appears to elongate and/or replace the trans-species, ontogenetic binding. When again the prominent evolutionary position of the serine molecule has been revealed for SARS-CoV-2 infection, this has become evident also in other, unrelated infectious diseases, such as malaria tropica, with the discovery of the serine repeat antigen SERA,[23, 24, 25] or *Entamoeba histolytica* infection, in which the serine-rich *E. histolytica* protein STREHP [26, 27] dictates the binding and virulence of the parasite.[28] It is proposed that Covid-19 infection is initiated by intermediate hybrid A-like/Tn antigen formation, hijacking the physiological, genetically undefined, serologically A-like/Tn ontogenetic structure, which must be differentiated from the human blood group A-specific epitope. This is encoded by the A allele of the ABO gene located on chromosome 9q34, and which together with the B-allele determines the risk of developing life-threatening disease in the non-O blood groups.[7] Interactions between different pathogenic viruses and human ABO(H) glycans are known for decades and are explained by similar molecular biological models. A human rotavirus interacts with A-type histo-blood group antigen and its infectivity was specifically abrogated by anti-A antibodies.[29] Appropriately, the analysis of a SARS-CoV outbreak in Hongkong 2008 revealed that blood group O was associated with low risk of infection, and the interaction between the viral spike protein and the host's cellular receptor was inhibited by natural and monoclonal anti-A antibodies *in vitro*, cited above.[5] Finally, the actual and first statistical study indicates that people with blood group A have a significantly higher risk for acquiring SARS-CoV-2 or COVID-19 infection, whereas people with blood group O have a significantly lower risk for the infection compared with non-O blood groups.[30] While this observation awaits confirmation, the central immunological position of blood group O

may have become evident in a small study 50 years ago,[31] which has been discussed in recent work (Fig. 2).[32, 6] Blood group O individuals, lacking the blood group-A-determining enzyme, may develop the least molecular contact with the pathogen and maintain the anti-A/Tn cross-reactive, complement-dependent isoagglutinin activities, which are exerted by the polyreactive, nonimmune immunoglobulin M (IgM),[6, 7, 32] representing the humoral spearhead of innate immunity and a first line of defense. SARS-Cov-2 or COVID-19 infection may be considered an evolutionary selective disease, contributing to the present-day world distribution of the human ABO(H) blood groups, which most likely has arisen through evolutionary selective diseases[34] over millions of years. While an ABO(H) blood group-related infection primarily occurs regardless of the blood group, the synthesis of the blood group AB enables the strongest contact with a pathogen and molecularly precluding any isoagglutinin activity, makes this group the least protected and the smallest among the ABO blood groups. In contrast, blood group O(H) individuals rarer develop severe disease and survive such infections as the largest blood group worldwide.

References.

1. Yazer, M.H., Hosseini-Maaf, B. & Olsson, M. L. Blood grouping discrepancies between ABO genotype and phenotype caused by O alleles. *Curr. Opin. Hematol.* **15**, 618–624 (2008).
2. Arend, P. ABO phenotype-protected reproduction based on human specific α1,2 L-fucosylation as explained by the Bombay type formation. *Immunobiology* vol. 223 684–693 (2018).
3. Oostra, M., de Haan, C. A. M., de Groot, R. J. & Rottier, P. J. M. Glycosylation of the Severe Acute Respiratory Syndrome Coronavirus Triple-Spanning Membrane Proteins 3a and M. *J. Virol.* (2006) doi:10.1128/jvi.80.5.2326-2336.2006.
4. Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C. & Garry, R. F. The proximal origin of SARS-CoV-2. *Nat. Med.* (2020) doi:10.1038/s41591-020-0820-9.
5. Guillon, P. *et al.* Inhibition of the interaction between the SARS-CoV Spike protein and its cellular receptor by anti-histo-blood group antibodies. *Glycobiology* (2008) doi:10.1093/glycob/cwn093.
6. Arend, P. Position of human blood group O(H) and phenotype-determining enzymes in growth and infectious disease. *Annals of the New York Academy of Sciences* (2018) doi:10.1111/nyas.13694.
7. Arend, P. *Malaria tropica evades host immunity through ABO blood group hybridization*. https://doi.org/10.6084/m9.figshare.8208689.v120 (2020) doi:10.6084/M9.FIGSHARE.8208689.V120.
8. Watanabe, Y., Bowden, T. A., Wilson, I. A. & Crispin, M. Exploitation of glycosylation in enveloped virus

pathobiology. *Biochimica et Biophysica Acta - General Subjects* (2019) doi:10.1016/j.bbagen.2019.05.012.

9. Arend, P. ABO (histo) blood group phenotype development and human reproduction as they relate to ancestral IgM formation: A hypothesis. *Immunobiology* **221**, (2016).
10. Inoue, Y. *et al.* Clathrin-Dependent Entry of Severe Acute Respiratory Syndrome Coronavirus into Target Cells Expressing ACE2 with the Cytoplasmic Tail Deleted. *J. Virol.* (2007) doi:10.1128/jvi.00253-07.
11. Simmons, G. et al. J., Reeves, J. D., Diamond, S. L. & Bates, P. Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. *Proc. Natl. Acad. Sci. U. S. A.* (2005) doi:10.1073/pnas.0505577102.
12. Wang, H. *et al.* SARS coronavirus entry into host cells through a novel clathrin- and caveolae-independent endocytic pathway. *Cell Res.* (2008) doi:10.1038/cr.2008.15.
13. Matsuyama, S. *et al.* Efficient Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein by the Transmembrane Protease TMPRSS2. *J. Virol.* (2010) doi:10.1128/jvi.01542-10.
14. Hoffmann, M. *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* (2020) doi:10.1016/j.cell.2020.02.052.
15. Latousakis, D., MacKenzie, D. A., Telatin, A. & Juge, N. Serine-rich repeat proteins from gut microbes. *Gut Microbes* (2020) doi:10.1080/19490976.2019.1602428.
16. Cidl, K., Strelcova, L., Znojil, V. & Vacha, J. Angiotensin I-converting enzyme (ACE) polymorphism and ABO blood groups as factors codetermining plasma ACE activity. *Exp. Hematol.* (1996).

17. Gassó, P., Ritter, M. A., Mas, S. & Lafuente, A. Influence of ABO genotype and phenotype on angiotensin-converting enzyme plasma activity. *JRAAS - J. Renin-Angiotensin-Aldosterone Syst.* (2014) doi:10.1177/1470320313510583.
18. Terao, C. *et al.* Quantitative variation in plasma angiotensin-i converting enzyme activity shows allelic heterogeneity in the ABO blood group locus. *Ann. Hum. Genet.* (2013) doi:10.1111/ahg.12034.
19. Hoffmann, M., Marx, K., Reichl, U., Wuhrer, M. & Rapp, E. Site-specific O-Glycosylation Analysis of Human Blood Plasma Proteins. *Mol. Cell. Proteomics* **15**, 624–641 (2016).
20. Wu, K. *et al.* A Virus-Binding Hot Spot on Human Angiotensin-Converting Enzyme 2 Is Critical for Binding of Two Different Coronaviruses. *J. Virol.* (2011) doi:10.1128/jvi.02274-10.
21. Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* (2020) doi:10.1038/s41586-020-2012-7.
22. Armijos-Jaramillo, V., Yeager, J., Muslin, C. & Perez-Castillo, Y. SARS-CoV-2, an evolutionary perspective of interaction with human ACE2 reveals undiscovered amino acids necessary for complex stability. *bioRxiv* (2020) doi:10.1101/2020.03.21.001933.
23. Bzik, D. J., Li, W. bo, Horii, T. & Inselburg, J. Amino acid sequence of the serine-repeat antigen (SERA) of Plasmodium falciparum determined from cloned cDNA. *Mol. Biochem. Parasitol.* (1988) doi:10.1016/0166-6851(88)90097-7.
24. Palacpac, N. M. Q., Arisue, N., Tougan, T., Ishii, K. J. & Horii, T. Plasmodium falciparum serine repeat antigen 5 (SE36) as a malaria vaccine candidate. *Vaccine* (2011)

doi:10.1016/j.vaccine.2011.06.052.

25. Aoki, S. *et al.* Serine repeat antigen (SERA5) is predominantly expressed among the SERA multigene family of Plasmodium falciparum, and the acquired antibody titers correlate with serum inhibition of the parasite growth. *J. Biol. Chem.* **277**, 47533–47540 (2002).
26. ZHANG, T., CIESLAK, P. R., FOSTER, L., KUNZ-JENKINS, C. & STANLEY, S. L. Antibodies to the serine rich Entamoeba histolytica protein (SREHP) prevent amoebic liver abscess in severe combined immunodeficient (SCID) mice. *Parasite Immunol.* (1994) doi:10.1111/j.1365-3024.1994.tb00344.x.
27. Stanley, S. L., Tian, K., Koester, J. P. & Li, E. The Serine-rich *Entamoeba histolytica* Protein Is a Phosphorylated Membrane Protein Containing *O* -Linked Terminal *N* -Acetylglucosamine Residues. *J. Biol. Chem.* **270**, 4121–4126 (1995).
28. Manochitra K, P. S. In-silico prediction and modeling of the Entamoeba histolytica proteins: Serine-rich Entamoeba histolytica protein and 29 kDa Cysteine-rich protease. *PeerJ.* **5**, 3160 (2017).
29. Hu, L. *et al.* Cell attachment protein VP8* of a human rotavirus specifically interacts with A-type histo-blood group antigen. *Nature* (2012) doi:10.1038/nature10996.
30. Zhao, J. *et al.* Relationship between the ABO Blood Group and the COVID-19 Susceptibility. *medRxiv* (2020) doi:10.1101/2020.03.11.20031096.
31. Arend, P. & Fehlhaber, G. Varying influence of increased enteral antigen absorption on the behavior of ‘natural’ antibodies in O and A blood group subjects. Comparative

blood group serological studies on patients with ulcerative colitis and healthy persons [Article in German]. *J. Mol. Med. (Klinische Wochenschrift)* **47**, 535–541 (1969).

32. Arend, P. Early ovariectomy reveals the germline encoding of natural anti-A- and Tn-cross-reactive immunoglobulin M (IgM) arising from developmental O-GalNAc glycosylations. (Germline-encoded natural anti-A/Tn cross-reactive IgM). *Cancer Med.* **6**, (2017).
33. Hofmann, B. T. *et al.* ABO Blood Group IgM isoagglutinins interact with tumor-associated O-glycan structures in pancreatic cancer. *Clin. Cancer Res.* **20**, 6117–6126 (2014).
34. Springer, G. F. & Wiener, A. S. Alleged causes of the present-day world distribution of the human abo blood groups. *Nature* (1962) doi:10.1038/193444a0.

Proposed (virtual) adhesion and/or response

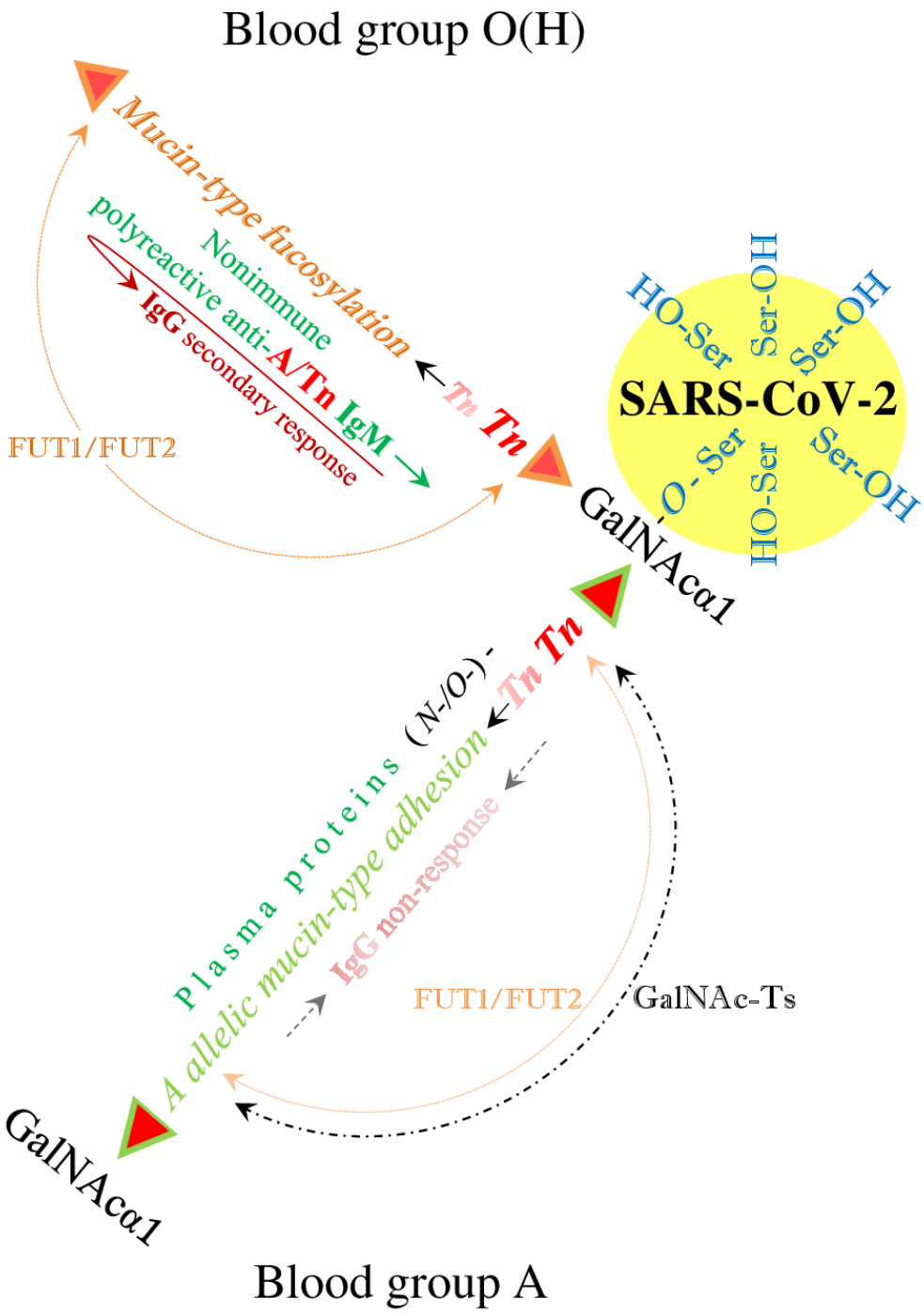

Figure 1

Figure 1.

The virus hypothetically mimicks the ABO phenotype synthetic pathways and utilizes the host cell's machinery via hijacking the host's A-like/Tn formation and phenotype-determining enzymes by serine-rich motifs. The ABO(H) phenotype formation occurs on the both cell surfaces and plasma-proteins.[9] While in the blood group A the naturally-occurring anti-A and anti-H antibody activities, exerted by the polyreactive, nonimmune IgM molecule, are neutralized by the phenotype-determining glycotransferases (FUT1/FUT2 and GalNAc-Ts), in blood group O the anti-A activites remain unaffected. Thus, the intermediate A-like/Tn structure has been elongated and/or replaced by phenotype-determining epitopes. This figure was constructed according to figure 2 in a previous article, in which this mechanism may be similarly utilized by a non-viral pathogen, such as the protozoan parasite *Plasmodium falciparum* (See references 6 and 7).

Central immunological position and evolutionary
cross-over point
of the human histo (blood group) O(H)

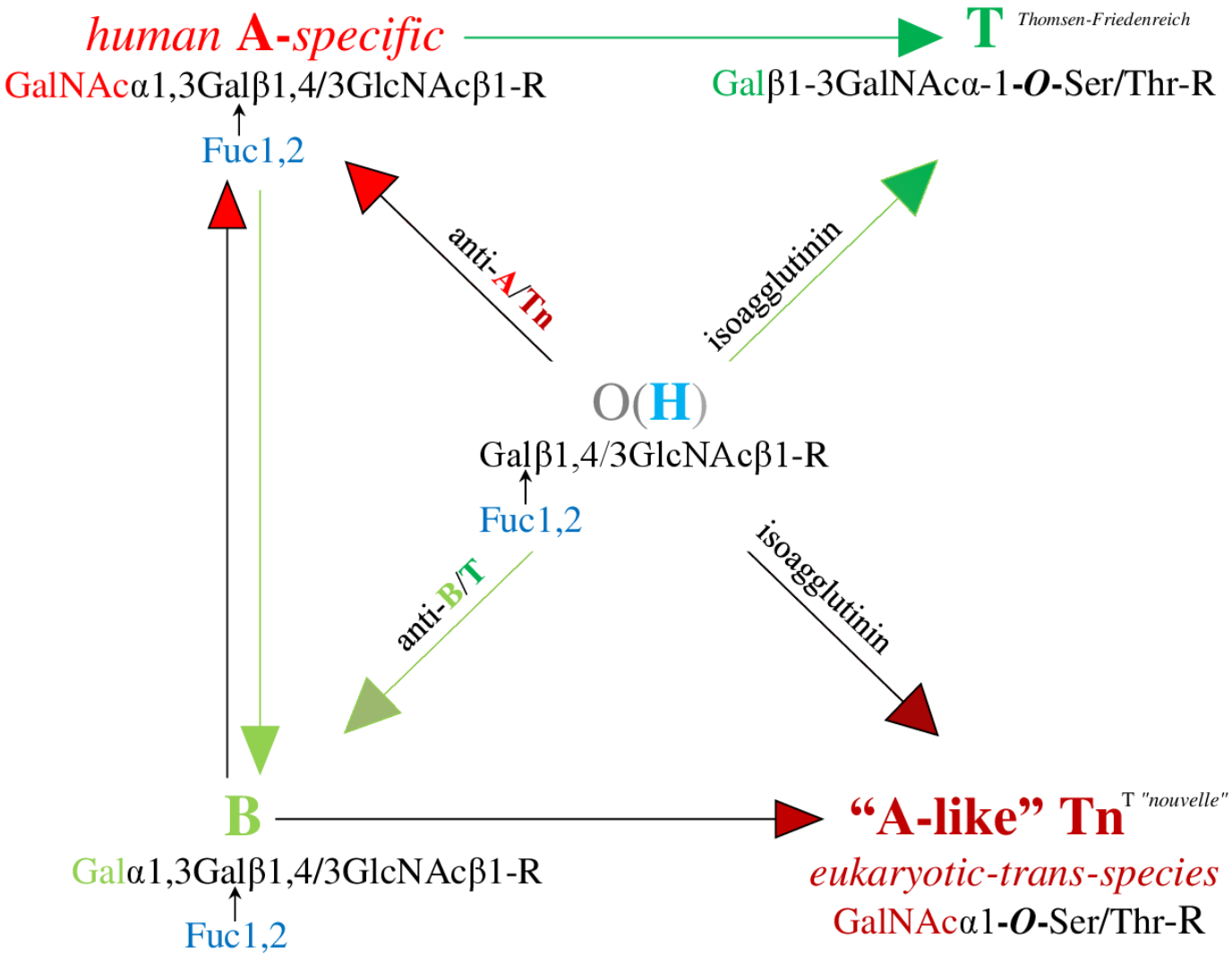

Figure 2.

Figure 2.

The central immunological position of the blood group O(H) is evident in its comprehensive production of "natural" antibodies against all of the mature A and B glycans and their cross-reactive developmental structures Tn and T. The human A-specific (A-allelic) glycosylation and the trans-species "A-like"/Tn formation are developmentally connected via the formation of cross-reactive anti-A/Tn isoagglutinin. According to Hofmann et al. (2014),[33] blood O(H) sera bind to both Tn and T antigens, and the anti-A isoagglutinin levels in blood group O(H) and blood group B sera are associated with the anti-Tn antibody, which does not react with blood group B red cells or T glycoconjugates. By contrast, the anti-B antibodies of blood group A sera and of blood group O(H) sera bind to B and T glycoconjugates but not to A or Tn glycoconjugates. This selective cross-reactivity of isoagglutinins with Tn and T antigens has been explained by the authors through the phenotype-specific terminal moieties; the terminal N-acetylgalactosamine is shared by A and Tn antigens, and the terminal galactose is, although with a different configuration, shared by B and T antigen. This figure was constructed by Arend (2017)[32], cited in Arend (2018).[6]