Why blood group A individuals are at risk whereas blood group O individuals might be protected from SARS-CoV-2 (COVID-19) infection: A hypothesis regarding how the virus invades the human body via ABO(H) blood group-determining carbohydrates

Peter Arend [a, b, c*]
ORCID ID: https://orcid.org/0000-0002-4000-4167

[a]Philipps University Marburg, Department of Medicine, D-355 Marburg, Lahn, Germany.
[b]Gastroenterology Research Laboratory, University of Iowa, Carver College of Medicine, Iowa City, IA 52242, USA.
[c]Research Laboratories, Chemie Grünenthal GmbH, D-52062 Aachen, Germany.

[*]Correspondence to: Dr. Peter Arend, Am Oberen Stötchen 14, 57462 Olpe, Germany.
E-mail: parend@t-online.de
[a, b, c] where the cited experiments of the author were performed.

Abstract

While the angiotensin converting enzyme 2 (ACE2) protein is defined as the primary severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) receptor, the viral serine molecule might be mobilized by the host's transmembrane protease serine subtype 2 (TMPRSS2) enzyme from the viral spike (S) protein and hijack the host's N-acetyl-D-galactosamine (GalNAc) metabolism. The resulting hybrid, serologically A-like/Tn (T-nouvelle) structure potentially acts as a host–pathogen functional molecular bridge. In humans, this intermediate structure will hypothetically be replaced by ABO(H) blood group-specific, mucin-type structures, in the case of infection hybrid epitopes, implicating the phenotypically glycosidic accommodation of plasma proteins. The virus may, by mimicking the synthetic pathways of the ABO(H) blood groups, bind to the cell surfaces of the blood group O(H) by formation of a hybrid H-type antigen as the potential precursor of hybrid non-O blood groups, which does not affect the highly anti-glycan aggressive anti-A and anti-B isoagglutinin activities, exerted by the germline-encoded nonimmune immunoglobulin M (IgM). In the non-O blood groups, which have developed from the H-type antigen, these IgM activities are downregulated by phenotypic glycosylation, while adaptive immunoglobulins might arise in response to the hybrid A and B blood group structures, suggesting the exertion of autoreactivity. The non-O blood groups thus become a preferred target for the virus, whereas blood group O(H) individuals, lacking the A/B phenotype-determining enzymes and binding the virus alone by hybrid H-type antigen formation, have the least molecular contact with the virus and maintain the critical anti-A and anti-B isoagglutinin activities, exerted by the ancestral IgM, which is considered the humoral spearhead of innate immunity.

Keywords: COVID-19; SARS-CoV-2–human carbohydrate interaction; trans-species glycosylation; A-like/Tn structure; trans-species glycan bridge.

Because the cross-reactive natural anti-A isoagglutinin and the natural anti-Tn (T-nouvelle)* reactivities of nonimmune IgM seem serologically to be inseparable, in this manuscript, I use the term anti-A/Tn-reactive IgM for the nonimmune anti-A reactive IgM and the term A-like/Tn for the corresponding ancestral structure.

*(Moreau R, Dausset J, Bernard J, 1957)

Introduction

Using a translational study, Preece et al. (2002) revealed how measles viruses undergo specific glycosylation when grown in cells that express blood group A, B, and O(H) antigens. Thus, they probably also demonstrated what occurs in human cells after SARS-CoV-2 infection. In view of the unknown native glycoprotein envelope of SARS-CoV-2, in the current paper I present a hypothesis regarding how infection occurs, which apparently is initiated via a trans-species *O*-glycosylation. However, infection does not mean disease, because the invasion of a pathogen does not always lead to symptom expression. Besides innumerable other factors, this phenomenon, which is observed in cases of the human coronavirus disease, results from the quality and degree of different types of physical and chemical bonds that occur between host and pathogen and probably reflects the functioning of the host's phenotype-determining enzymes. The molecular biology of the pathogenesis of an infectious disease determines the susceptibility of a species to the infection. Nevertheless, symptom development or subsequent disease severity might depend on the phenotype. Given that *O*-glycosylation plays a key role in the pathogenesis of the human coronavirus disease, as discussed 14 years ago regarding SARS-CoV-1 infection (Oostra et al., 2006) and is currently predicted for SARS-CoV-2 or COVID-19 (Andersen et al., 2020), would implicate the formation of hybrid, serologically A-like, *O*-GalNAcα1-Ser/Thr-R, Tn antigenic structures (Arend, P., 2020a). This prediction is not consistent with the concept of Guillon et al. (2008), which propose that these viruses bind to hosts using *N*-glycosylation, although their observation that the interaction between the viral S protein of SARS-CoV-1 and the host's cellular receptor was inhibited by natural and monoclonal

anti-A antibodies speaks more for a predominance of the host's *O*-glycoproteome. In fact, this observation indicates the inhibition of a blood group A-like synthesis when the most critical step in SARS-CoV-2 infection appears to be the mobilization of the viral serine molecule (Matsuyama et al., 2010; Hoffmann et al., 2020), representing the key molecule of the classical *O*-glycan and the serologically A-like/Tn structure. Moreover, when Influenza A virus infection induces production of mucinous structures and expression of a GalNAc transferase, encoded by the GALNT3 gene (Nakamura *et al.*, 2016), a mobilization of serine molecules most likely has preceded, and apart from functional *N*-glycosylations, the *O*-glycoproteome of the host apparently plays a key role in the pathogenesis of SARS-CoV-2 infection. The adhesion of the virus to host cells appears to primarily occur independently of the ABO(H) blood group via the genetically undefined intermediate, namely the serologically A-like/Tn evolutionary/developmental structure. In metazoan evolution, up to 20 distinct genetically undefined polypeptide *O*-GalNAc transferases catalyze the first addition of GalNAc to a protein (Bennett et al., 2012), synthesizing the above, serologically A-like *O*-GalNAcα1-Ser/Thr-R Tn, which results from the most complex and differentially regulated step in protein glycosylation (Brockhausen et al., 2009) and representing a normal yet fleeting intermediate structure, characterizing stem cell fidelity (Reisner et al., 1978; Nash et al., 2007). This structure is common to all processes of metazoan development and/or ontogeny, and its hybridized form apparently acts as a host–pathogen functional bridge in different, unrelated infectious diseases: a key function of the serine molecule was also discussed in an earlier article by the author for the non-viral pathogenesis of malaria tropica (Arend, P., 2018b; 2020b). Thus, while the ACE2 protein is

defined as the primary SARS-CoV-2 receptor, evidence suggests that the host and pathogen bind via an intermediate hybrid *O*-glycan, dominated by the pathogen's hydrophilic amino acid serine. In humans, the resulting ontogenetic and blood group-independent structure may, depending on the phenotype, be elongated and/or replaced by mucin-type, blood group-specific, ABO(H) phenotype-determining carbohydrates according to established pathways (Vitiazeva et al., 2015).

Proposed blood group-independent and ABO(H) blood group-specific, mucin-type adhesion of SARS-CoV-2 to human cells

The coevolution of species drives diversity in animals and plants and contributes to natural selection; however, during an infectious disease outbreak, a pathogen may complete an incomplete evolutionary/developmental function by utilizing the host cell's machinery (Arend, P, 2020b). The analysis of older, yet related, data suggests that *Plasmodium falciparum,* the pathogen causing malaria tropica, cannot survive outside its human host because it is unable to initiate protein glycosylation or blood group-independent (serologically A-like) *O*-GalNAcα1-Ser/Thr-R, Tn antigen formation, owing to its inability to synthesize the amino sugar N-acetyl-D-galactosamine (GalNAc) (Dieckmann-Schuppert, Bause and Schwarz, 1993), nor does it possess the genes required for mucin-type *O*-glycan synthesis (Templeton et al., 2004). Thus, the excess of serine production in *Plasmodium falciparum* (Bzik et al., 1988; Aoki et al., 2002; Arisue et al., 2011) might be explainable and in malaria tropica the pathogen's serine molecule seems to be a driver of the infection.

Although the pathogenesis of a non-viral infection cannot be compared to that of a viral infection, the following data strongly

suggests that SARS-CoV-2 invades the human cell by forming an intermediate hybrid *O*-glycan. The virus cannot survive outside its host, and hypothetically utilizes the host cell's machinery by hijacking the A-like/Tn formation after liberation of serine molecules. A recent review published by Watanabe et al. (2019) provided similar suggestions. The virus obviously enters the human body via the ACE2 protein, a polyfunctional protein that represents the binding domain of SARS-CoV viruses. In a complex signaling pathway, ACE2 binds to the S protein on the viral envelope (Inoue et al., 2007) and, after the subsequent cleavage of the ACE–viral S protein complex by cathepsin L (Simmons et al., 2005), the virus enters the cell by receptor-mediated endocytosis (Wang et al., 2008). Regarding which amino acids dominate the host–pathogen fusion process, the most critical molecular step appears to be the mobilization of the viral serine molecule which is performed by the host protease TMPRSS2 (Matsuyama et al., 2010; Hoffmann et al., 2020). The history of this hydrophilic amino acid and its essential involvement in the pathogenesis of SARS-CoV-2 strongly suggest that the binding between pathogen and host occurs via *O*-glycosylation. Although serine-rich repeat proteins (SRRPs) that are involved in the adhesion of different bacteria (Latousakis et al., 2020) to host cell carbohydrates via *O*-glycosylation have not yet been described for the pathogenesis of viral infections, such a mechanism might occur in the pathogenesis of SARS-CoV infections. A new therapeutic observation might reveal the role of the serine molecule in SARS-CoV infections, when an inorganic polymer, polyphosphate, blocks binding of SARS-CoV-2 spike protein to ACE2 receptor, Neufurth, et al. (2020), while the serine molecule is the preferred target also in protein phosphorylation, Ardito et al. (2017). Moreover, the inhibition of the TMPRSS2

enzyme or the blocking of serine mobilization, studied by Meyer et al. (2013) in an E. coli infection, may be utilized by other pharmaceuticals in future therapeutic strategies against SARS-CoV-2 (Wang et al., 2020). Furthermore, inhibiting the interaction between the viral S protein and the host cell receptor by natural and monoclonal anti-A antibodies (Guillon et al., 2008) indicates the inhibition of an *O*-glycan formation. In conclusion, when the most critical molecular step in the pathogenesis of SARS-CoV-2 appears to be the mobilization of the viral serine molecule, the preferential occurrence of severe symptoms in individuals with non-O blood groups suggests that the phenotype-determining saccharides of blood group A and B are the glycosidic targets of serine residues. These residues are preserved on the viral S protein and become available through the action of the host protease TMPRSS2. Apart from functional peptide formations, the essential binding between host and pathogen might not result in a hybrid peptide but an intermediate hybrid *O*-glycan; the ACE2 receptor protein, assumingly codetermined by the ABO(H) phenotype (Cidl et al., 1996; Gassó et al., 2014; Terao et al., 2013), would mediate the activities of the ABO(H) glycan-transferring enzymes to achieve blood group A- and/or B-specific mucin-type (*O*-glycosidic), hybrid binding.

How the ABO(H) blood group phenotype might dominate the humoral innate immunity during SARS-CoV-2 infection

Humoral innate immunity, or the first line of defense, and its complex connection with ABO(H) phenotype formation are considered to play the main role in SARS-CoV-2 infection and in the course of subsequent disease. In contrast to adaptive, environmentally-induced immunoglobulin or B-cell activities, which are controlled by clonal selection, the production of the nonimmune and

polyreactive IgM is not restricted to B cells, but spontaneously occurs in murine (Zhou, O'Hara and Chen, 2011; Shao *et al.*, 2016) and human (F. Hu *et al.*, 2012a) normal and malignant epithelial cells as well. The early ovariectomy of the C57BL/10 inbred mouse (Arend and Nijssen, 1976; 1977a; 1977b) may have revealed that in the mammalian species this IgM molecule is a germline-encoded antibody, which is, through a complex evolutionary process, structurally connected with the ancestral A-like/Tn antigen (detected in water-soluble ovarian glycolipids) and is considered its complementary protein (Arend, P, 2016; P., 2017). Although Larkin and Porter (2005) render the mouse an unsuitable model of the ABO(H)-phenotype discordance in primates, the murine nonimmune IgM has, in the presence of complement, lysed human RBC like an anti-A isoantibody and has even distinguished between the subgroups A_1 and A_2 (Arend and Nijssen, 1977a).

The human nonimmune neonatal IgM molecule is an aggressive anti-glycan-reactive antibody, exerting the innate anti-A and anti-B isoagglutinin activities and suggests the induction of ADCC (antibody-dependent) and/or complement-mediated cytotoxicity, which acts as a bridge to cellular immunity. This antibody molecule undergoes the ABO(H) blood group phenotype formation, occurring on both the cell surfaces and plasma proteins, wherein the evolutionary significance of the serine molecule may become evident in neonatal IgM sequencing; this hydrophile amino acid appears to guaranty alone the polyreactivity of the neonatal IgM molecule (Willis et al, 2013), binding various epitopes via ionic and hydrogen bonding at V-regions (Wang et al. 2016). Thus, serine residues might also serve as acceptors in phenotypical accommodation of autoreactive antibody activities. The metabolic

centre of this hypothetical process could be a blood group-active soluble plasma protein, such as the α2-macroglobulin molecule, whose ABO(H) blood group activity correlates strictly with the phenotype (Matsui *et al.*, 2001). This molecule is considered an evolutionarily conserved arm of the innate immune system (Armstrong and Quigley, 1999) and shows a functional synergism with the structurally related IgM molecule (Stevenson et al., 2015), which will glycosidically be tailored to the phenotypic specificities (Arend, P., 2016; 2017; 2018b), involving the down-regulation of the otherwise autoreactive anti-A activity in blood group A and the anti-B activity in blood group B. Blood group O(H) alone loses the anti-H and maintains the highly anti-glycan aggressive anti-A and anti-B isoagglutinin activities. Both the anti-A levels in blood group B and the anti-B levels in group A are always lower than in blood group O(H) due to cross-reactivities, and it is important to note that environmentally induced anti-A and anti-B antibodies are only produced in blood group O(H) and especially anti-A and anti-B secondary (IgG) immune responses appear to be restricted to this blood group due to clonal selection (Arend, P., 2017; 2018b; Stussi *et al.*, 2005). In the human, the majority of infections occur in view of this innate immunological superiority of the blood group O(H).

SARS-CoV-2 hypothetically evades human immunity by hybridization of the ABO(H) blood groups, or mimicking the above described metabolic pathways. Both blood group O(H) and the non-O blood groups are infected via blood group-independent, trans-species intermediate A-like/Tn (*O*-GalNAc-Ser) binding, which in blood group O(H) is replaced by mucin-type fucosylation, performing the synthesis of a hybrid H-type antigen as the potential precursor of hybrid non-O blood groups. Thus, in

the blood groups A, B and AB, the intermediate A-like/Tn binding is hypothetically replaced by hybrid A- and/or B-allelic mucin type formation via FUT1FUT2 and GalNAc- and/or D-Gal glycosylation, while the anti-B or anti-A-isoagglutinin levels are physiologically reduced (Fig. 1).

The process of viral ABO(H) blood group hybridization, mimicking the syngeneic epitopes as the target of this hybridization, is illustrated in Fig. 2: blood group O(H), which is bound to the virus solely by the formation of the hybrid H-type antigen, has the least contact with the pathogen and is the most protected group when it loses only the anti-H isoagglutinin, but retains the innate anti-A and anti-B isoagglutinin reactivities, involving secondary IgG responses. In the blood group A, the anti-A and anti-H formations are physiologically blocked, while the anti-B shows a reduced level and, as cited above, neither anti-A- nor anti-B-reactive IgG is produced in this group due to clonal selection. In blood group B, the anti-B and anti H-reactivities are blocked respectively, while the anti-A shows a reduced level, as the anti-B in blood group A and neither anti-A nor anti-B-reactive IgG is produced in this blood group. Blood group AB has the strongest contact to pathogen and is the least protected group with respect to innate ABO(H) immunity.

Ultimately, the human ABO(H) blood group phenotypes arise from the evolutionarily oldest genetic system found in primate populations and develop in molecular and functional connection with a special humoral innate immunity, dominated by the nonimmune polyreactive IgM. The ABO(H) phenotypic glycosylation of both the cell surfaces and the plasma proteins occurs identically, which, under normal conditions, physiologically precludes a corresponding natural autoreactivity,

basically exerted by the ancestral nonimmune or neonatal IgM (Arend, P., 2016). However, in the case of infections, such as SARS-CoV-2, this principle enables the formation of foreign hybrid structures. The physiological lack of innate anti-A and anti-B antibodies in the non-O blood groups, namely A, B, and AB, poses an immunological dilemma. On one hand, it protects them from self-reactivity against complementary structures, but on the other hand, it cannot prevent the formation of hybrid, most likely auto-antigenic, structures that arise in a subsequent pathogenic step.

It is assumed that during SARS-CoV-2 infection, especially in the non-O blood groups, the induction of autoimmune processes might contribute to the development of severe symptoms, which may even be dominated by autoimmune inflammation. In fact, this phenomenon has been observed in severe cases of malaria tropica (Hart et al., 2016; Rivera-Correa et al., 2017). It has been explained by the author through hybridization (Arend, 2020b), but awaits elucidation through studies of the complex mechanisms of cellular immunity, which is ignored in this manuscript that exclusively is focused on the ABO(H) blood group phenotype-determined humoral innate immunity and the first steps of SARS-CoV-2 pathogenesis.
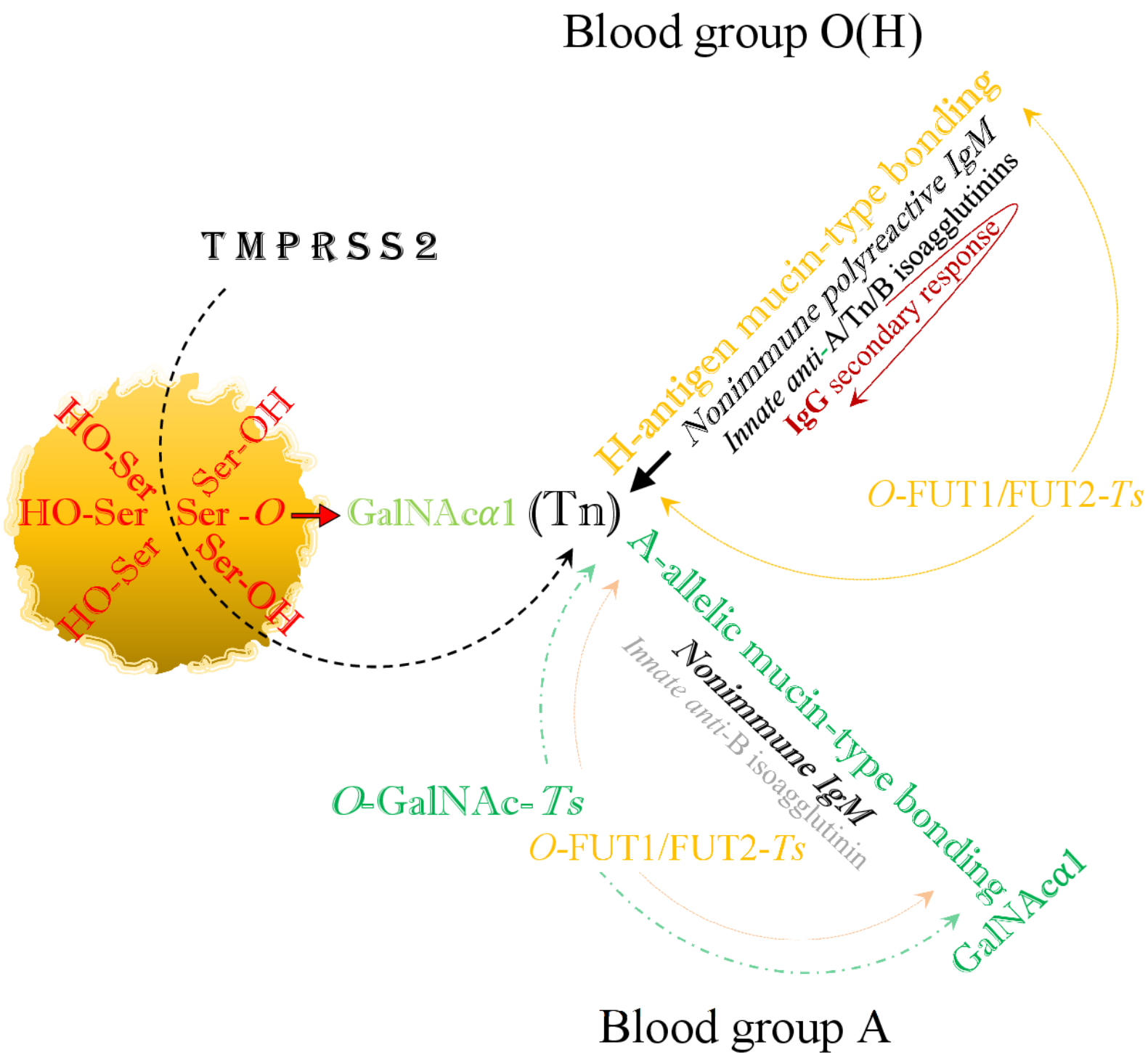
Figure 1.

Figure 1. The viral SARS-CoV-2 serine residues, mobilized by host's TMPRSS2, highjack the host's GalNAc metabolism and both blood group O(H) and blood group A are identically infected via blood group independent, trans-species intermediate *O*-GalNAc-Ser (Tn) glycosylation. In blood group O(H) this intermediate hybrid structure is replaced by mucin-type fucosylation or H-antigen formation, which neutralizes the activity of innate anti-H isoagglutinin but leaves unaffected the activities of innate anti-A/Tn and anti-B isoagglutinins, exerted by the nonimmune polyreactive IgM, implicating a secondary IgG response. In blood group A, the intermediate Tn binding is hypothetically replaced by hybrid A-allelic mucin-type formation via mucin-type fucosylation. This involves the phenotypic accommodation of the polyreactive nonimmune IgM, downregulation of anti-A/Tn IgM (isoagglutinin) activities and decrease in the level of the anti-B IgM (isoagglutinin) activity, while anti-A/B reactive IgG formations are precluded by clonal selection. This figure was constructed according to 'Figure 2' in a previous article (Arend, P., 2018b), in which this mechanism may be similarly utilized by a non-viral pathogen, such as the protozoan parasite *Plasmodium falciparum*.
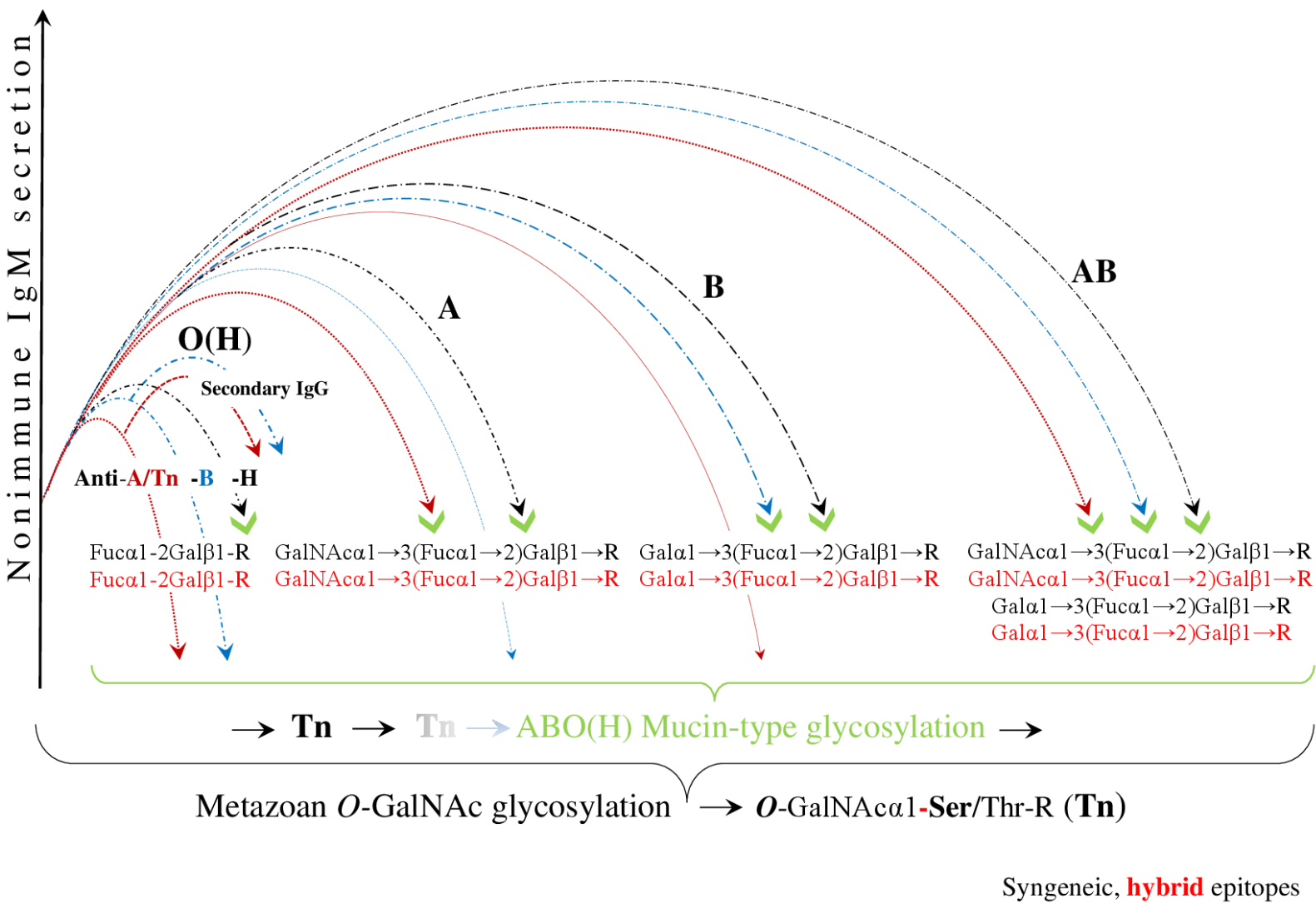
Figure 2.

Figure 2. Evading immunity by ABO(H) blood group hybridization. After SARS-CoV-2 infection, initiated by trans-species, metazoan *O*-GalNAc-glycosylation, the virus hypothetically invades human cells via hybridization of the ABO(H) blood groups. This results in the downregulation and/or reduction of the innate anti-ABO(H) isoagglutinin reactivity, exerted by polyreactive nonimmune IgM. In this figure, the virtual, hybrid mucin-type epitopes mimic syngeneic epitopes, which are the targets of this hybridization. The blood group O(H), which is bound to the virus solely by the formation of the hybrid H-type antigen, has the least amount of contact with the pathogen and is the most protected group when it loses only the anti-H, but retains innate anti-A and anti-B isoagglutinin reactivity, involving secondary IgG responses. In blood group A, the anti-A and anti-H formations are physiologically blocked, a decrease in the level of anti-B isoagglutinin is observed, and neither anti-A- nor anti-B-reactive IgG is produced. In blood group B, the anti-B and anti H-reactivities are blocked, a decrease in the level of anti A is observed, and, similar to blood group A, neither anti-A nor anti-B-reactive IgG is produced. Blood group AB has the strongest contact to the pathogen and is the least protected group with respect to innate ABO(H) immunity.

(Because the first contact between the SARS-CoV-2 virus and the host organism, and the first contact between the protozoan parasite *P-falciparum* and the host organism are hypothetically identical, similar images were used for the illustrations (Arend, P., 2020b))

Central immunological position and evolutionary cross-over point of the human histo (blood group) O(H)
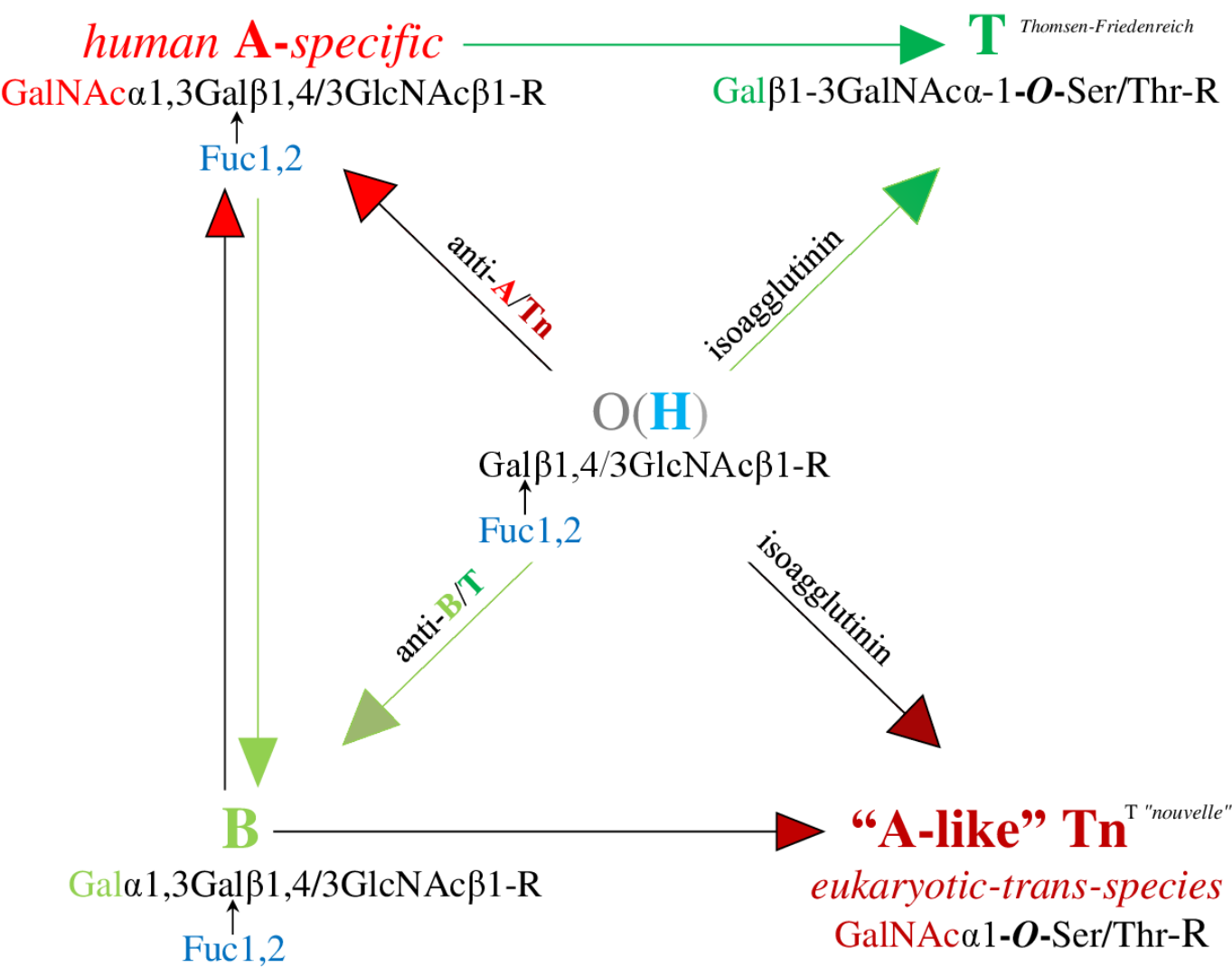
Figure 3.

Figure 3. The central immunological position of 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), 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 and 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 was 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 antigens. This figure was constructed by Arend (2017) and cited in Arend (2018b; 2018a).

Conclusions

The proposed concept of a viral invasion, initiated by the mobilization of the serine molecule from the viral S protein and completed by the formation of a genetically undefined, hybrid A-like/Tn host–pathogen molecular bridge, does not question the established functions of the ACE2 receptor protein based on previous (Wu et al., 2011) and current definitions (Zhou et al., 2020; Armijos-Jaramillo et al., 2020). Rather, it shows an additional and more specific interaction between host and pathogen. Both *N*- and *O*-glycosylation may occur within this complex pathogenic process, and among multiple chemical and physicochemical linkage options, both trans-species ontogenetic, blood group-independent, and blood group-specific binding may predominantly be performed through *O*-glycosylation in two different glycosidic steps, dominated by the pathogen's hydrophilic amino acid serine. The prominent evolutionary position of the serine molecule, most likely even determining the polyreactivity of the neonatal IgM (Willis et al, 2013), is again revealed in SARS-CoV-2 infection and is evident in other unrelated infectious diseases, for example with the serine repeat antigen (SERA) in malaria tropica (Bzik et al., 1988; Aoki et al., 2002; Arisue et al., 2011) and the serine-rich *E. histolytica* protein (STREHP) of *Entamoeba histolytica* (Zhang et al., 1994; Stanley et al., 1995), which dictates the binding and virulence of the parasite (Manochitra, 2017) in amoebic dysentery. Finally, as mentioned above, a new therapeutic observation in SARS-CoV infections might also reveal the role of the serine molecule within this infection, when an inorganic polymer, polyphosphate, blocks binding of SARS-CoV-2 spike protein to ACE2 receptor, Neufurth, et al. (2020), while the serine molecule is the preferred target also in protein phosphorylation, Ardito et al. (2017).

Intriguingly, the susceptibility of blood group A individuals to infections with *Plasmodium falciparum*, the pathogen of malaria tropica, is similar to infections with SARS-CoV-2, and since the ABO(H) phenotype development is molecularly

connected to the development of humoral innate immunity, it might be tempting to speculate that both the viral and the non-viral pathogenesis will be initiated via a hybrid, developmental classical A-like/Tn *O*-glycan.

Again, it is proposed that SARS-CoV-2 infection is initiated via a functional host–pathogen molecular bridge by forming an intermediate and genetically undefined, serologically A-like/Tn 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, which, together with the B-allele, determines the risk of developing life-threatening diseases in the non-O blood groups (Arend, 2020b). Interactions between different pathogenic viruses and human ABO(H) glycans have been recognized for decades and can be explained by molecular biological models of a similar nature. A human rotavirus interacts with the A-type histo-blood group antigen, and its infectivity is specifically abrogated by anti-A antibodies (L. Hu *et al.*, 2012b). Appropriately, the comprehensive study by Guillon et al. (2008) and their analysis of a SARS-CoV-1 outbreak in Hongkong in 2003, demonstrated that blood group O(H) was associated with low risk of infection, while the interaction between the viral S protein and the host cell receptor was inhibited by natural and monoclonal anti-A antibodies in vitro. Finally, the actual and first statistical study of SARS-CoV-2 indicated that individuals with blood group A have a significantly higher risk of acquiring SARS-CoV-2 or COVID-19 infection, whereas people with blood group O have a significantly lower risk of infection compared to non-O blood groups (Zhao et al., 2020).

In the "true blood group O", O(h) or *Bombay* type (Bhende et al., 2008), which lacks any ABO(H) blood group carbohydrate synthesis and, consistent with the presented concept, demonstrates the strongest isoagglutinin activities (involving complement-dependent anti-H with an optimum effect at 37^OC), the susceptibility to SARS-CoV-2 cannot be assessed by statistical standards due to the extremely small

Bombay population, whereas the central immunological position of the classical human blood group O(H), (Fig. 3) might have already become evident in a small study conducted 50 years ago (Arend and Fehlhaber, 1969); in this study an adaptive, via the gut microbiome occurring isoagglutinin induction, was statistically documented exclusively for the histo (blood) group O(H), dominated by the IgG class. However, the individual risk of becoming infected with SARS-CoV-2 or becoming seriously ill cannot be predicted based on a person's ABO(H) blood group affiliation alone because many other risks exist and blood group O(H) is no longer considered a genetic entity (Arend, P., 2016). Nevertheless, SARS CoV-2 (COVID-19) infection can be regarded as an evolutionary selective disease, contributing to the current global distribution in terms of human blood groups O(H), A, B, and AB, which according to Springer and Wiener (1962) developed over millions of years mainly in connection with ABO(H) blood group-related life-threatening diseases, such as malaria (Cserti and Dzik, 2007; Cserti-Gazdewich, 2010; Arend, P., 2018b; 2020b). The synthesis of blood group AB enables the strongest contact with a pathogen and molecularly precludes any isoagglutinin activity, making this group the least protected and the smallest among the ABO(H) blood groups. In contrast, individuals with blood group O(H), which are prone to other infections, especially cholera, have survived all infectious diseases in an immunological balance with many pathogens and remain the largest blood group worldwide despite extensive historical cholera pandemics (Echenberg, 2011; Mutreja et al., 2011; Chowdhury et al., 2017). These individuals rarely develop severe diseases classified as blood group A/B-related infections. They maintain anti-A/Tn cross-reactive and anti-B complement-dependent isoagglutinin activities, exerted by the polyreactive, nonimmune IgM, which is regarded as the humoral spearhead of innate immunity and the first line of defense.

This is the first description of a hypothetical pathogenesis, wherein the contact between host and pathogen is initiated by formation of a hybrid trans-species, devel-

opmental A-like/Tn O-glycan, which (in its normal or syngeneic version) through its intermediate function plays a key role in the evolution of species, in the human is replaced by (in the case of infections) hybrid ABO(H) phenotypic epitopes and is controlled by its molecularly and functionally connected innate immunity. The in this manuscript described hypothetical formation of hybrid ABO(H) blood group-specific epitopes requires experimental evidence, while the hybrid A-like/Tn structure might become the basis for a new pharmacological and immunological therapeutic target, independent of the ABO(H) blood group phenotype.

References

Andersen, K.G., Rambaut, A., Lipkin, W.I., et al., 2020. The proximal origin of SARS-CoV-2. Nat. Med. 26, 450–452. https://doi.org/10.1038/s41591-020-0820-9.

Aoki, S., Li, J., Itagaki, S., et al., 2002. 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 (49), 47533–47540. https://doi.org/10.1074/jbc.m207145200.

Ardito F, Giuliani M, Perrone D, Troiano G, Muzio L Lo. The crucial role of protein phosphorylation in cell signalingand its use as targeted therapy (Review). *Int J Mol Med*. 2017. doi:10.3892/ijmm.2017.3036

Arend, P., 2016. ABO (histo) blood group phenotype development and human reproduction as they relate to ancestral IgM formation: A hypothesis. Immunobiology 221 (1), 116–127. https://doi.org/10.1016/j.imbio.2015.07.003.

Arend, P., 2017. 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 (7), 1601–1613. https://doi.org/10.1002/cam4.1079.

Arend, P., 2018a. ABO phenotype-protected reproduction based on human specific α1,2 L-fucosylation as explained by the Bombay type formation. Immunobiology 223 (11), 684–693. https://doi.org/10.1016/j.imbio.2018.07.015.

Arend, P., 2018b. Position of human blood group O(H) and phenotype-determining enzymes in growth and infectious disease. Ann. N. Y. Acad. Sci. 1425 (1), 5–18. https://doi.org/10.1111/nyas.13694

Arend, P., 2020a. 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-determining carbohydrates). figshare, v175.

Preprint. https://doi.org/10.6084/m9.figshare.12019035.v34

[dataset]

Arend, P., 2020b. Malaria tropica evades host immunity through ABO blood group hybridization. figshare, v120. https://doi.org/10.6084/m9.figshare.8208689.v120

Arend, P., Fehlhaber, G. 1969. 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]. Klin. Wochenschr. 47 (10), 535–541. https://doi.org/10.1007/bf01715818.

Arend, P., Nijssen, J., 1976. Significance of specific ovarian receptors for syngeneic naturally-occurring haemagglutinating anti-A antibodies. J. Immunogenet. 3 (6), 373–382. https://doi.org/10.1111/j.1744-313x.1976.tb00598.x.

Arend, P., Nijssen, J., 1977a. Age dependent appearance of A-specific ovarian glycolipids and syngeneic "natural" anti-A hemolysin in mice. Zeitschrift fur Immunitatsforschung: Immunobiology 153 (1), 74–84. https://doi.org/10.1016/S0340-904X(77)80028-6.

Arend, P., Nijssen, J., 1977b. A-specific autoantigenic ovarian glycolipids inducing production of "natural" anti-A antibody. Nature 269 (5625), 255–257. https://doi.org/10.1038/269255a0.

Arisue, N., Kawai, S., Hirai, M., et al., 2011 Clues to evolution of the SERA multigene family in 18 *Plasmodium* species. PLoS One 6 (3), e17775. https://dx.doi.org/10.1371%2Fjournal.pone.0017775.

Armijos-Jaramillo, V., Yeager, J., Muslin, C., et al., 2020. SARS-CoV-2, an evolutionary perspective of interaction with human ACE2 reveals undiscovered amino acids necessary for complex stability. Evol. Appl. 00, 1–11. https://doi.org/10.1111/eva.12980.

Armstrong, P.B., Quigley, J.P., 1999. Alpha2-macroglobulin: an evolutionarily

conserved arm of the innate immune system. Dev. Comp. Immunol. 23 (4–5), 375–390. https://doi.org/10.1016/s0145-305x(99)00018-x.

Bennett, E., Mandel, U., Clausen, H. et al. (2012) 'Control of mucin-type O-glycosylation: A classification of the polypeptide GalNAc-transferase gene family.', *Glycobiology*, 22(6), pp. 736–56. doi: doi: 10.1093/glycob/cwr182.

Bhende, Y.M., Deshpande, C.K., Bhatia, H.M. et al., 2008. A "new" blood-group character related to the ABO system. 1952. Natl. Med. J. India 21 (5), 3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8005636.

Brockhausen, I., Schachter, H. & Stanley P in Varki, A. (2009) 'Essentials of Glycobiology. 2nd. Edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009, Varki, A. Ed. et al. Chapter 9, O-GalNAc Glycans', in Ajit Varki, Richard D Cummings, Jeffrey D Esko, Hudson H Freeze, Pamela Stanley, Carolyn R Bertozzi, Gerald W Hart, and M. E. E. (ed.). Cold Spring Harbor Laboratory Press.

Bzik, D.J., Li, W.B., Horii, T., et al., 1988. Amino acid sequence of the serine-repeat antigen (SERA) of Plasmodium falciparum determined from cloned cDNA. Mol. Biochem. Parasitol. 30 (3), 279–288. https://doi.org/10.1016/0166-6851(88)90097-7.

Chowdhury, F.R., Nur, Z., Hassan, N., et al., 2017. Pandemics, pathogenicity and changing molecular epidemiology of cholera in the era of global warming. Ann. Clin. Microbiol. Antimicrob. 16 (1), 10. https://doi.org/10.1186/s12941-017-0185-1.

Cídl, K., Strelcová, L., Znojil, V., et al., 1996. Angiotensin I-converting enzyme (ACE) polymorphism and ABO blood groups as factors codetermining plasma ACE activity. Exp. Hematol. 24 (7), 790–794.

Cserti, C. M. and Dzik, W. H. (2007) 'The ABO blood group system and Plasmodium falciparum malaria', *Blood*, pp. 2250–2258. doi: 10.1182/blood-2007-

03-077602.

Cserti-Gazdewich, C. M. (2010) 'Plasmodium falciparum malaria and carbohydrate blood group evolution', *ISBT Science Series*. doi: 10.1111/j.1751-2824.2010.01380.x.

Dieckmann-Schuppert, A., Bause, E., Schwarz, R.T., 1993. Studies on O-glycans of Plasmodium-falciparum-infected human erythrocytes. Evidence for O-GlcNAc and O-GlcNAc-transferase in malaria parasites. Eur. J. Biochem. 216 (3), 779–788. https://doi.org/10.1111/j.1432-1033.1993.tb18198.x.

Echenberg, M.J., 2011. Africa in the time of cholera: a history of pandemics from 1815 to the present (African studies). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511976599.

Gassó, P., Ritter, M.A., Mas, S., et al., 2014. Influence of ABO genotype and phenotype on angiotensin-converting enzyme plasma activity. –J. Renin-Angiotensin-Aldosterone Syst. 15 (4), 580–584. https://doi.org/10.1177/1470320313510583.

Guillon, P., Clément, M., Sébille, V., et al., 2008. Inhibition of the interaction between the SARS-CoV Spike protein and its cellular receptor by anti-histo-blood group antibodies. Glycobiology 18 (12), 1085–1093. https://doi.org/10.1093/glycob/cwn093.

Hart, G.T., Akkaya, M., Chida, A.S., et al., 2016. The Regulation of Inherently Autoreactive VH4-34–Expressing B Cells in Individuals Living in a Malaria-Endemic Area of West Africa. J. Immunol. 197 (10), 3841–3849. https://doi.org/10.4049/jimmunol.1600491.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., et al., 2020. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell 181 (2), 271–280. https://dx.doi.org/10.1016%2Fj.cell.2020.02.052.

Hofmann, B. T. *et al.* (2014) 'ABO Blood Group IgM isoagglutinins interact with

tumor-associated O-glycan structures in pancreatic cancer', *Clinical Cancer Research*, 20(23), pp. 6117–6126. doi: 10.1158/1078-0432.CCR-14-0716.

Hu, F., Zhang, L., Zheng, J., et al., 2012a. Spontaneous Production of Immunoglobulin M in Human Epithelial Cancer Cells. PLoS One 7 (12), e51423. https://doi.org/10.1371/journal.pone.0051423.

Hu, L., Crawford, S.E., Czako, R., et al., 2012b. Cell attachment protein VP8* of a human rotavirus specifically interacts with A-type histo-blood group antigen. Nature 485 (7397), 256–259. https://doi.org/10.1038/nature10996.

Inoue, Y., Tanaka, N., Tanaka, Y., et al., 2007. Clathrin-Dependent Entry of Severe Acute Respiratory Syndrome Coronavirus into Target Cells Expressing ACE2 with the Cytoplasmic Tail Deleted. J. Virol. 81 (16) 8722–8729. https://doi.org/10.1128/jvi.00253-07.

Larkin, J., Porter, C., 2005. Mice are unsuitable for modelling ABO discordance despite strain-specific A cross-reactive natural IgM. Br. J. Haematol. 130 (2), 310–317.

Latousakis, D., MacKenzie, D.A., Telatin, A., et al., 2020. Serine-rich repeat proteins from gut microbes. Gut Microbes 11 (1), 102–117. https://doi.org/10.1080/19490976.2019.1602428.

Manochitra, K., Parija, S.C., 2017. *In-silico* prediction and modeling of the *Entamoeba histolytica* proteins: Serine-rich *Entamoeba histolytica* protein and 29 kDa Cysteine-rich protease. PeerJ 5, e3160. https://doi.org/10.7717/peerj.3160.

Matsui, T., Hamako, J., Ozeki, Y., et al., 2001. Comparative study of blood group-recognizing lectins toward ABO blood group antigens on neoglycoproteins, glycoproteins and complex-type oligosaccharides. Biochim. Biophys. Acta Gen. Subj. 1525 (1–2), 50–57. https://doi.org/10.1016/S0304-4165(00)00170-7.

Matsuyama, S., Nagata, N., Shirato, K., et al., 2010. Efficient Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein by the

Transmembrane Protease TMPRSS2. J. Virol. 84 (24), 12658–12664. https://doi.org/10.1128/jvi.01542-10.

Meyer, D., Sielaff, F., Hammami, M., et al., 2013. Identification of the first synthetic inhibitors of the type II transmembrane serine protease TMPRSS2 suitable for inhibition of influenza virus activation. Biochem. J. 452 (2), 331–343. https://doi.org/10.1042/bj20130101.

Moreau R, Dausset J, Bernard J, M. J. (1957) 'Acquired hemolytic anemia with polyagglutinability of erythrocytes by a new factor present in normal blood (Article in French)', *Bull Mem Soc Med Hop Paris.*, 73(20–21), pp. 569–87.

Mutreja, A., Kim, D.W., Thomson, N.R., et al., 2011. Evidence for several waves of global transmission in the seventh cholera pandemic. Nature 477 (7365), 462–465. https://doi.org/10.1038/nature10392.

Nash, R., Neves, L., Faast, R. et al. (2007) 'The Lectin Dolichos Biflorus Agglutinin Recognizes Glycan Epitopes on the Surface of Murine Embryonic Stem Cells: A New Tool for Characterizing Pluripotent Cells and Early Differentiation', *Stem Cells*, 25(4), pp. 974–982. doi: 10.1634/stemcells.2006-0224.

Neufurth M, Wang X, Tolba E, et al. The inorganic polymer, polyphosphate, blocks binding of SARS-CoV-2 spike protein to ACE2 receptor at physiological concentrations. *Biochem Pharmacol*. 2020. doi:10.1016/j.bcp.2020.114215

Oostra, M., de Haan C.A., de Groot, R.J., et al., 2006. Glycosylation of the Severe Acute Respiratory Syndrome Coronavirus Triple-Spanning Membrane Proteins 3a and M. J. Virol. 80 (5), 2326–2336. https://doi.org/10.1128/jvi.80.5.2326-2336.2006.

Preece, A.F., Strahan, K.M., Devitt, J., et al., 2002. Expression of ABO or related antigenic carbohydrates on viral envelopes leads to neutralization in the presence of serum containing specific natural antibodies and complement. Blood 99 (7), 2477–2482. https://doi.org/10.1182/blood.v99.7.2477.

Reisner, Y., Itzicovitch, L., Meshorer, A. & Sharon, N. (1978) 'Hemopoietic stem cell transplantation using mouse bone marrow and spleen cells fractionated by lectins.', *Proc Natl Acad Sci U S A*, 75(6), pp. 2933–6. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=392680&tool=pmcentr ez&rendertype=abstract.

Rivera-Correa, J., Guthmiller, J.J., Vijay, R., et al., 2017. Plasmodium DNA-mediated TLR9 activation of T-bet+ B cells contributes to autoimmune anaemia during malaria. Nat. Commun. 8, 1282. https://doi.org/10.1038/s41467-017-01476-6.

Shao, W., Hu, F., Ma, J., et al., 2016. Epithelial cells are a source of natural IgM that contribute to innate immune responses. Int. J. Biochem. Cell Biol. 73, 19–29. https://doi.org/10.1016/j.biocel.2016.01.017. Bennett, E., Mandel, U., Clausen, H. et al. (2012) 'Control of mucin-type O-glycosylation: A classification of the polypeptide GalNAc-transferase gene family.', *Glycobiology*, 22(6), pp. 736–56. doi: doi: 10.1093/glycob/cwr182.

Simmons, G., Gosalia, D.N., Rennekamp, A.J., et al., 2005. Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. Proc. Natl. Acad. Sci. U. S. A. 102(33):11876–11881. https://doi.org/10.1073/pnas.0505577102.

Springer, G. F. and Wiener, A. S. (1962) 'Alleged causes of the present-day world distribution of the human abo blood groups', *Nature*, 193(4814), pp. 444–445. doi: 10.1038/193444a0.

Stanley, S.L., Tian, K., Koester, J.P., et al., 1995. The Serine-rich *Entamoeba histolytica* Protein Is a Phosphorylated Membrane Protein Containing *O*-Linked Terminal *N*-Acetylglucosamine Residues. J. Biol. Chem. 270 (8), 4121–4126. http://dx.doi.org/10.1074/jbc.270.8.4121.

Stevenson, L., Laursen, E., Cowan, G.J., et al., 2015. α2-Macroglobulin Can Cross-link Multiple *Plasmodium falciparum* Erythrocyte Membrane Protein 1 (PfEMP1)

Molecules and May Facilitate Adhesion of Parasitized Erythrocytes. PLoS Pathog. 11 (7), e1005022. https://doi.org/10.1371/journal.ppat.1005022.

Stussi, G., K. Huggel, H. U. Lutz, U. Schanz, R. Rieben, and J. D. Seebach, "Isotype-specific detection of ABO blood group antibodies using a novel flow cytometric method," *Br. J. Haematol.*, vol. 130, no. 6, pp. 954–963, 2005, doi: 10.1111/j.1365-2141.2005.05705.x.

Templeton, T.J., Iyer, L.M., Anantharaman, V., et al., 2004. Comparative analysis of apicomplexa and genomic diversity in eukaryotes. Genome Res. 14 (9), 1686–1695. https://dx.doi.org/10.1101%2Fgr.2615304.

Terao, C., Bayoumi, N., McKenzie, C.A., et al., 2013. Quantitative variation in plasma angiotensin-I converting enzyme activity shows allelic heterogeneity in the ABO blood group locus. Ann. Hum. Genet. 77(6):465–471. https://doi.org/10.1111/ahg.12034.

Vitiazeva, V., Kattla, J.J., Flowers, S.A., et al., 2015. The *O*-linked glycome and blood group antigens ABO on mucin-type glycoproteins in mucinous and serous epithelial ovarian tumors. PLoS One 10 (6), e0130197. https://doi.org/10.1371/journal.pone.0130197.

Wang, H., Yang, P., Liu, K., et al., 2008. SARS coronavirus entry into host cells through a novel clathrin- and caveolae-independent endocytic pathway. Cell Res. 18 (2), 290–301. https://doi.org/10.1038/cr.2008.15.

Wang, H., Coligan, J. E. and Morse, H. C. (2016) 'Emerging functions of natural IgM and its Fc receptor FCMR in immune homeostasis', *Frontiers in Immunology*. doi: 10.3389/fimmu.2016.00099.

Wang, X., Dhindsa, R., Povysil, G., et al., 2020. Transcriptional inhibition of host viral entry proteins as a therapeutic strategy for SARS-CoV-2. Preprints. https://doi.org/10.20944/preprints202003.0360.v1.

Watanabe, Y., Bowden, T.A., Wilson, I.A., et al., 2019. Exploitation of

glycosylation in enveloped virus pathobiology. Biochim. Biophys. Acta Gen. Subj. 1863 (10), 1480–1497. https://doi.org/10.1016/j.bbagen.2019.05.012.

Willis JR1, Briney BS, DeLuca SL, Crowe JE Jr, M. J. (2013) 'Human Germline Antibody Gene Segments Encode Polyspecific Antibodies', *PLoS Computational Biology*, (9(4):e1003045.). doi: 10.1371/journal.pcbi.1003045.

Wu, K., Chen, L., Peng, G., et al., 2011. A Virus-Binding Hot Spot on Human Angiotensin-Converting Enzyme 2 Is Critical for Binding of Two Different Coronaviruses. J. Virol. 85 (11) 5331–5337. https://doi.org/10.1128/JVI.02274-10.

Zhang, T., Cieslak, P.R., Foster, L., et al., 1994. Antibodies to the serine rich Entamoeba histolytica protein (SREHP) prevent amoebic liver abscess in severe combined immunodeficient (SCID) mice. Parasite Immunol. 16 (5), 225–230. https://doi.org/10.1111/j.1365-3024.1994.tb00344.x.

Zhao, J., Yang, Y., Huang, H., et al., 2020. Relationship between the ABO Blood Group and the COVID-19 Susceptibility. medRxiv. https://doi.org/10.1101/2020.03.11.20031096.

Zhou, P., Yang, X., Wang, X., et al., 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579, 270–273. https://doi.org/10.1038/s41586-020-2012-7.

Zhou, R., O'Hara, S.P., Chen, X.M., 2011. MicroRNA regulation of innate immune responses in epithelial cells. Cell. Mol. Immunol. 8 (5), 371–379. https://doi.org/10.1038/cmi.2011.19.