 doi:10.20944/preprints202109.0501.v2

Review: Sagacious epitope selection for vaccines, and both antibody-based therapeutics and diagnostics: Tips from virology and oncology.

Authors: Samuel Ken-En Gan[1,2,*], Ser-Xian Phua[1], Joshua Yi Yeo[1]

Affiliation:

[1] Antibody & Product Development Lab, EDDC-BII, Agency for Science, Technology and Research (A*STAR), Singapore 138672, Singapore; phuasx@gmail.com (S.-X.P.); joshua_yeo@eddc.a-star.edu.sg (J.Y.Y.)

[2] APD SKEG Pte Ltd, Singapore 439444, Singapore

*Corresponding author: samuel_gan@eddc.a-star.edu.sg; samgan@apdskeg.com

Antibody & Product Development Lab

60 Biopolis Street, #B2 Genome, Singapore 138672

Tel: +65 6407 0584

Keywords: Antigen Selection; Epitope Selection; Antibody Targeting; Epitope Accessibility; Antibody Engineering; Protein Engineering; Drug Targeting; Model-Informed Drug Discovery and Development

 doi:10.20944/preprints202109.0501.v2

Abbreviations:

MID3 (Model Informed Drug Discovery and Development), HIV (Human immunodeficiency virus), HBV (Hepatitis B virus), RT (Reverse transcriptase), SARS-CoV-2 (Severe acute respiratory syndrome coronavirus 2), MERS (Middle East Respiratory Syndrome), SARS (Severe acute respiratory syndrome, RdRp (RNA-dependent RNA polymerase), MHC (Major histocompatibility complex), LSTM (Long short-term memory), CDR (Complementarity-determining region), MD (Molecular dynamics), IgM (Immunoglobulin M), Fab (Fragment antigen-binding), BiTEs (Bispecific T-cell engagers), CH (Constant heavy), IgA (Immunoglobulin A), IgE (Immunoglobulin E), HPV (Human papillomavirus), HCV (Hepatitis C virus), HFV (Haemorrhagic fever virus), GISAID (Global initiative on sharing all influenza data), Fc (Fragment crystallizable), C (Constant), V (Variable), HFMD (Hand, foot and mouth disease), EV (Enterovirus), NNRTIs (Non-nucleoside reverse transcriptase inhibitors), MIS-C (Multisystem inflammatory syndrome in children), ART (Antiretroviral therapy), PTM-SD (Post Translational Modification Structural Database), NMR (Nuclear magnetic resonance).

Statement of Significance: Advances in protein engineering and antibody development has allowed focus on the target antigen for antibody-based design thinking to maximize the success of antibody developments. Based on Model-Informed Drug Discovery and Development, considerations of epitope factors such as accessibility and locality allow for better epitope selection and interventions.

 doi:10.20944/preprints202109.0501.v2

ABSTRACT

The target of an antibody plays a significant role in the success of antibody-based therapeutics and diagnostics, and to an extent, that of vaccine development. This importance is focussed on the target binding site – epitope, where epitope selection as a part of design thinking beyond traditional antigen selection using whole cell or whole protein immunisation can positively impact success. With purified recombinant protein production and peptide synthesis to display limited/selected epitopes, intrinsic factors that can affect the functioning of resulting antibodies can be more easily selected for. Many of these factors stem from the location of the epitope that can affect accessibility of the antibody to the epitope at a cellular or molecular level, direct inhibition of target antigen activity, conservation of function despite escape mutations, and even non-competitive inhibition sites. Through the incorporation of novel computational methods for predicting antigen changes to model-informed drug discovery and development, superior vaccines and antibody-based therapeutics or diagnostics can now be more easily designed to mitigate failures. With detailed examples, this review highlights the new opportunities, factors and methods of predicting antigenic changes for consideration in sagacious epitope selection.

 doi:10.20944/preprints202109.0501.v2

Introduction

Antibodies and their fragments are increasingly important in diagnostics and therapeutics development as evidenced in the ongoing COVID-19 pandemic (1,2). An already expensive process, diagnostics can fail owing to escape mutations on the epitope that compromise primer-based kits (2-4) or diagnostic antibody binding (5) even with sagacious rational antibody design and engineering (6). On therapeutics, antigenic epitope changes leading to escape mutations contribute to the expensive drug failures. Thereby, to improve success, Model Informed Drug Discovery and Development (MID3) (7,8) has been in pilot by the U.S. Food and Drug Administration since 2018 (9) to support drug development (8,10).

The ability to select the right single antigen for diagnostics, therapeutics, and to an extent that for vaccines targeting (e.g., choosing only the Spike over a whole virus), was augmented through recombinant technology, where purified target antigens could be produced and either injected into animals or used with *in vitro* antibody display methods e.g., phage display for antibody selection. The same technology also supported the targeting of specific epitopes on the antigen, where having antibodies specific to an epitope in a diagnostic kit can improve selectivity and specificity. This specificity is useful when differentiating between similar antigens e.g., between viruses such as human immunodeficiency virus (HIV) and Hepatitis B (HBV), even when targeting Reverse Transcriptase (RT) or distinguishing severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) from Middle East Respiratory Syndrome (MERS), Influenza, and severe acute respiratory syndrome (SARS) when targeting the RNA-dependent RNA polymerase (RdRp). Yet, being too specific in diagnostics, as opposed for vaccinations, can also result in false negatives when the target epitope on the antigen mutates beyond antibody recognition (11).

With advances in peptide technology, short/stapled peptides can also be used without necessarily going to the deoxyribonucleic acid (DNA) level for recombinant expression and

 doi:10.20944/preprints202109.0501.v2

the commonly used whole cells or antigens (12,13). Recent developments in cyclic peptides and peptide vaccines further allow for immunization against specific conformational structures in epitopes instead of whole antigens through mimotopes (14) as B cell peptides (15). Such higher selective methods can support the development of therapeutics to reduce off-target effects, although some level of lower specificity could be of value for vaccines, and to an extent, diagnostics to target variants.

The selection of epitopes when incorporated into antibody design thinking, is thus a paradigm shift from chance-dependent antibody development to a more rational purposeful approach. Give the dependence on the intended application in guiding towards high specificity or to have cater to minor changes, there are two categories (16) of 1) linear/continuous, defined as a stretch of amino acids sequences; 2) conformational/discontinuous, defined as sequence distal residues in close proximity through protein folding, with the latter conformational type more prevalent as B-cell epitopes.

About 96% of monoclonal antibody therapeutic candidates fail to make it to the market (17), costing close to tens of millions of dollars (18) for each failure, augmenting the MID3 approach that includes epitope prediction. Epitope prediction, traditionally based on amino acid physicochemical properties such as hydrophobicity, flexibility, solvent accessibility and antigenicity (19-22) have seen augmentation by machine learning methods to show promise for cancer (23) and even hybrid experimental-computational approaches (24) involving deep neural network for major histocompatibility complex (MHC) binding (25) and attention-based long short-term memory (LSTM) networks (26).

 doi:10.20944/preprints202109.0501.v2

Epitopes factors

Accessible Epitopes (Cellular)

For epitope selection particularly for therapeutics, accessibility of the epitope by the antibodies is perhaps the first and foremost consideration. In maximising success, extracellular targets are typically picked for antibody therapeutic candidates with the exception for 'intrabodies' (where target cells produce the antibodies against intracellular antigens within itself (27)). Apart from intrabodies, whole antibodies against intracellular targets often have limited penetration of membranes (knowledge known to frequent flow cytometry users) which will require intracellular delivery methods (28,29) that are not always feasible for therapeutics. Although unknown in prevalence, there are instances where intracellular oncogenic markers can be targeted because of externalization e.g. PRL3 (30). Despite the cellular penetration handicap, antibodies do have intrinsic advantages for undruggable targets that are not suitable for small molecules (29) in specificity, requiring both the selection and effectiveness of targeting intracellular proteins to have specific assessments on a case-by-case basis.

With antibody intracellular delivery hurdles yet to be overcome, most current therapeutics and diagnostic whole antibody targets are extracellular antigens. There are many factors within extracellular targets such as post-translational modification which can underlie the suitability and recognition of the epitopes. In one example, glycosylation has been shown to impede antibody recognition in both the ongoing COVID-19 pandemic (31) and HIV (32) when the modification occurred at epitope sites. Thus, even with good epitope bioinformatics-calculated prediction scores or good experimental results from bacterial produced non-glycosylated proteins (with rare exceptions), care must be taken for the effects of post-translational modification on occluding or interfering with antibody recognition.

Epitope occlusion, in the form of cryptic epitope or cryptotopes (33), has been reported in numerous viruses such as the Norovirus (34), Influenza (35), Ebola (36) and HIV **(Fig. 1A**),

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 November 2021 doi:10.20944/preprints202109.0501.v2

where apart from post-translational modifications methods, HIV utilised the hypervariable regions in the Env to occlude gp41 epitopes (37) from immune detection. On self-antigens, such occlusion of inducible cryptic epitopes play a role in reducing autoimmunity (38), but in some cases, can be exploited for differentiating disease states when the cryptotopes are exposed during pathogenesis e.g. prion disease (39). In the unlikely situation where targeting the viral receptor is not easily achieved, blocking the viral target (host cell receptor) from viral spike binding can be performed, as in the case for SARS-CoV-2 (40,41) and poliovirus (42), although care needs to be taken not to end up affecting host cell activity by the antibody by over-activation (e.g. mimicking receptor stimulation and inducing of altered signalling, triggering uncontrolled microthrombosis, cell lysis and neutrophil activation (43)) or by preventing activation (e.g. inhibiting hyaluronan clearance by liver cells (44)).

Antigens have the potential to bind antibodies in non-conventional complementarity determining regions (CDR) dependent manner by inducing binding pockets or the formation of stretches/patches on antibodies. This phenomenon was observed for non-conventional immunogenic molecules such as nickel (45) binding to Trastuzumab and Pertuzumab that can possibly underlie the disease pathogenesis of nickel type-I allergy. The other example of induced binding is in the molecular dynamics (MD) simulation of Trastuzumab binding to Her2, inducing a cryptotopes that facilitated Pertuzumab binding (46). While the synergism between the two clinical therapeutics were shown experimentally to be due to the different epitopes on the same antigen without evidence of an induced Pertuzumab epitope (47), the possibility of occluded epitopes ought to be sagaciously considered depending on the desired application and utility.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 November 2021 doi:10.20944/preprints202109.0501.v2

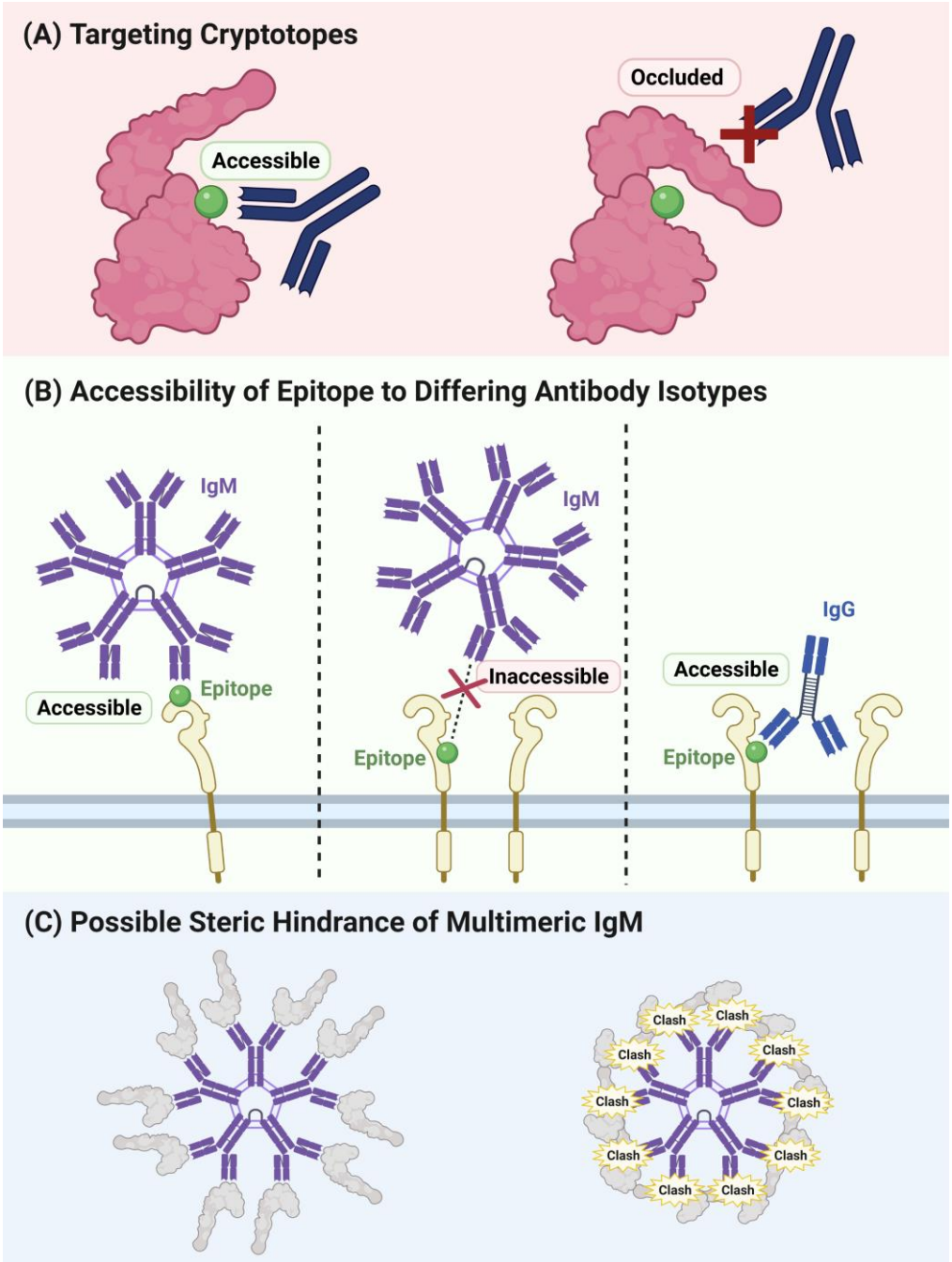

Figure 1. Accessibility of Epitopes. (**A**) Targeting cryptotopes. (**B**) Accessibility of epitope to differing antibody isotypes. The location of the epitope on the target antigen can affect its accessibility by antibodies that are conjugated or multimeric in nature due to steric hindrances. This is evident in internal epitopes that would not be accessible to multimeric IgMs but accessible to monomeric Igs or antibody fragments. (**C**) Steric hindrances for multivalent binding by antibodies due to the location of the epitope on the target antigen. Created with BioRender.com.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 November 2021 doi:10.20944/preprints202109.0501.v2

Accessible Epitopes (Molecular): Lessons from Immunoglobulin M (IgM) for multi-specific antibodies

Apart from access at the intra/extracellular level and occluded/induced epitopes, obvious steric hindrances at the molecular level to access epitopes can impact the efficacy of the antibody **(Fig. 1B)**. While the earlier mentioned example (46,47) showed synergistic binding of Trastuzumab and Pertuzumab simultaneously to their different epitopes on Her2, steric hindrances resulting from multiple whole or conjugated or multimeric antibodies are known to arise in flow cytometry (48).

When made into multimeric IgMs for multiple antigen binding, Trastuzumab IgM could not have full occupancy of its Fab regions due to steric hindrances binding to multiple Her2. This was however, not in the case for Pertuzumab IgM (49), which showed higher avidity effects (50). The sheer size of multimeric antibodies, can be advantageous in agglutination for immune clearance and result in a better therapeutic or diagnostic than the traditional IgG, as evidenced in a nasal delivery for SARS-CoV-2 (51). Yet, this advantage needs to be studied for possible steric hindrance **(Fig. 1C)**. With IgM being typically used in hemagglutination assays (52-55), one needs to consider these steric effects on the accuracy of such IgM based assays.

For this reason, the checklist to selecting any epitope for diagnostic or therapeutic application needs to ensure molecular accessibility, especially when utilizing larger antibodies (whether conjugated or multimeric or whole). It should be noted that the textbook primary antibody response is typically IgM and steric hindrances may underlie (in part at least) why IgMs tend to have lower affinities for the antigen.

Such steric hindrances also apply to the development of bi-specific antibodies requiring additional optimization (56). Potentially limiting the effectiveness of multi-specific antibodies, especially those intended to engage whole cells (57) (the promises of bispecific T-cell engagers

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 November 2021 doi:10.20944/preprints202109.0501.v2

(BiTEs) for oncology (58)), accessibility must be considered to maximize success. It is in this area that perhaps a more flexible antibody hinge at the constant heavy (CH) 1 domain of antibodies may alleviate some structural constraints as evidenced in immunoglobulin A (IgA) (59) and other isotypes (50).

Conservation of Epitopes: Lessons from Viruses and Omalizumab-Immunoglobulin E (IgE)

Since escape mutations in the epitope results in antibody recognition failure, one key main criterion of epitope selection is its conservation and this occurs in at least two levels within the antigen of 1) in the presence of mutations; and 2) conservation within the family to allow broad-spectrum targeting **(Fig. 2)**. The earlier level of being within the species or viral type is exemplified in the recent COVID-19 pandemic where SARS-CoV2 spike mutations led to decreased effectiveness of the early vaccines to novel variants (5). On the level for broad-spectrum protection, other viral vaccines such as that against the human papillomavirus (HPV) are found to induce cross-neutralizing antibodies (60) to its close relations.

Likely due to faster escape mutations in microbial pathogens, there are more monoclonal antibody therapeutics used clinically against cancer than infectious diseases (61), where those targeting the latter tend to be polyclonal (62), perhaps in an attempt to cover more epitopes to mitigate escape. Nonetheless, in the search for conserved regions in pathogens, large sequence databases can provide some insights (such as HBVdb (63), Los Alamos Hepatitis C Virus (HCV) (64), Haemorrhagic Fever Virus (HFV) (65), HIV (66,67), Stanford HIV Drug Resistance (68), Global Initiative on Sharing All Influenza Data (GISAID) (69) and Nextstrain (70)), but bearing in mind the constraints of extracellular targets for both neutralization of viruses for therapeutics and detection in diagnostics that do not require pre-processing to release internal intracellular contents. It should be noted that broad-spectrum targets of such nature are expectedly limited due to viral trophism.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 November 2021 doi:10.20944/preprints202109.0501.v2

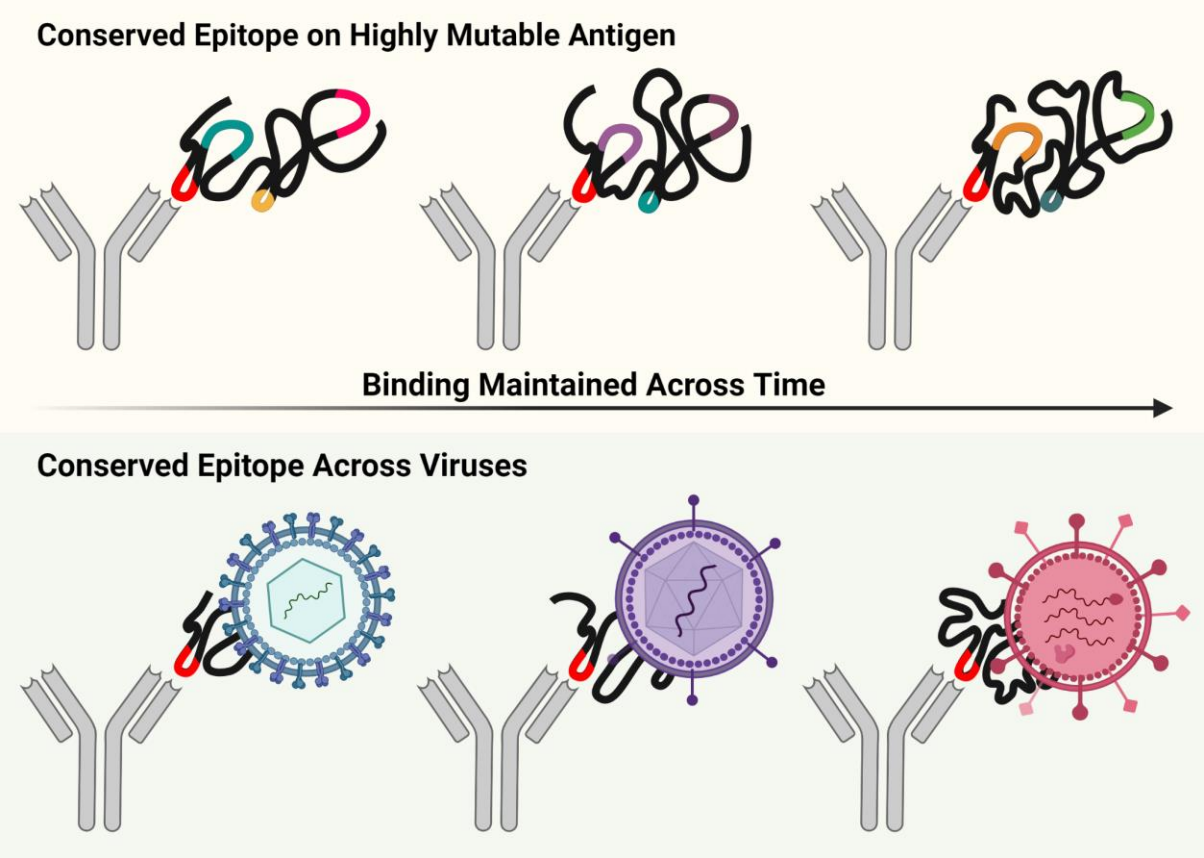

Figure 2. Conservation of epitopes across highly mutable target antigens (top) would reduce the chance of escape mutations even as the pathogen accumulates mutations in the target antigen to prolong the effectiveness of the antibody in either detecting or neutralizing the target. Similarly, targeting conserved epitopes that are conserved in similar proteins (e.g., Reverse Transcriptase) across viruses or pathogens can allow broad-spectrum detection or therapeutic intervention (bottom), improving its market use while also useful against emerging viruses utilizing replication enzymes e.g., RT or RdRp of the same protein family. Created with BioRender.com.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 November 2021 doi:10.20944/preprints202109.0501.v2

With some foresight, overcoming cellular level accessibility issues e.g. intrabodies or novel delivery methods to target intracellular targets, conserved viral epitopes can expand to viral enzymes (71) for better search of potential broad-spectrum antivirals. This can be performed through a search of substrate analogues sites that indicate the presence of a functional domain (72,73) that would be more conserved due to the preservation of enzyme functions.

In antibodies, the broad conserved target of Omalizumab (Xolair®) to IgE fragment crystallizable (Fc) enabled it to be effective against various IgE type I allergies (74). Since the constant (C) region of antibodies is generally more conserved, Omalizumab was sagaciously raised against the C-region, overlapping with the FcεRIα binding site. While the existence of allotypes in the Cε-regions (75) can be an issue, other possible effects recently reported includes the allosteric communication from the variable (V) region that could influence FcεRIα interaction (76). Despite overlapping with the binding site of FcεRI, the V-region distal allosteric effects had negligible effects on Omalizumab binding. This make the Omalizumab binding side a sagacious epitope that displayed the following benefits: 1) at a conserved region across a variety of pathogenic factors; 2) at the allosterically insulated regions that is unaffected by the changes at hypervariable regions; and 3) relatively immutable.

Predicting New Mutated Epitopes: Lessons from Viruses

The need for conserved epitopes in the face of mutations is evidenced by the challenge to existing COVID-19 vaccines by new variants (77). In the lack of conserved epitopes, the alternative option is to pre-emptively predict possible escape mutations in the epitopes and design interventions that would remain effective against them. This is not only easily performed through recombinant or peptide methods discussed above, but even easier to implement as vaccines with the clinical use of mRNA vaccines (78).

 doi:10.20944/preprints202109.0501.v2

Despite the lack of their reported incorporation for therapeutic or diagnostic development, the use of prediction methods in vaccines have been in place for a while with the example of the successful polio virus RdRp in vitro platform (79) and similar attempts for HFMD causative Enterovirus (EV) A71 (80). These methods of predicting new mutations exhibit great promise despite being *in vitro* and can benefit from *in silico* augmentation involving network analysis (81), robust statistical model building (82), time series (83), stacking models (84) and random signature analysis (85,86). Nonetheless, it should be noted that EVs and lentiviruses like HIV, mutate through recombination whereas Influenza utilize the Orthomyxovirus family method of assortment, requiring further research to get to a unified computational method capable of dealing with RNA viruses of different mutation methodologies. For reassortment viruses, it is important to consider that *in silico* simulations of the reassortment would be more efficient and safer to generate given the need of co-infection which can be constrained experimentally by biosafety and bioethics concerns in possible gain-of-function experiments. Regardless of the methodology approach, key features of trophism changes or polymerase activity in respiratory viruses, such as that mentioned for H5N8 (87) would be areas of importance to focus in species jumping and pandemic potential.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 November 2021 doi:10.20944/preprints202109.0501.v2

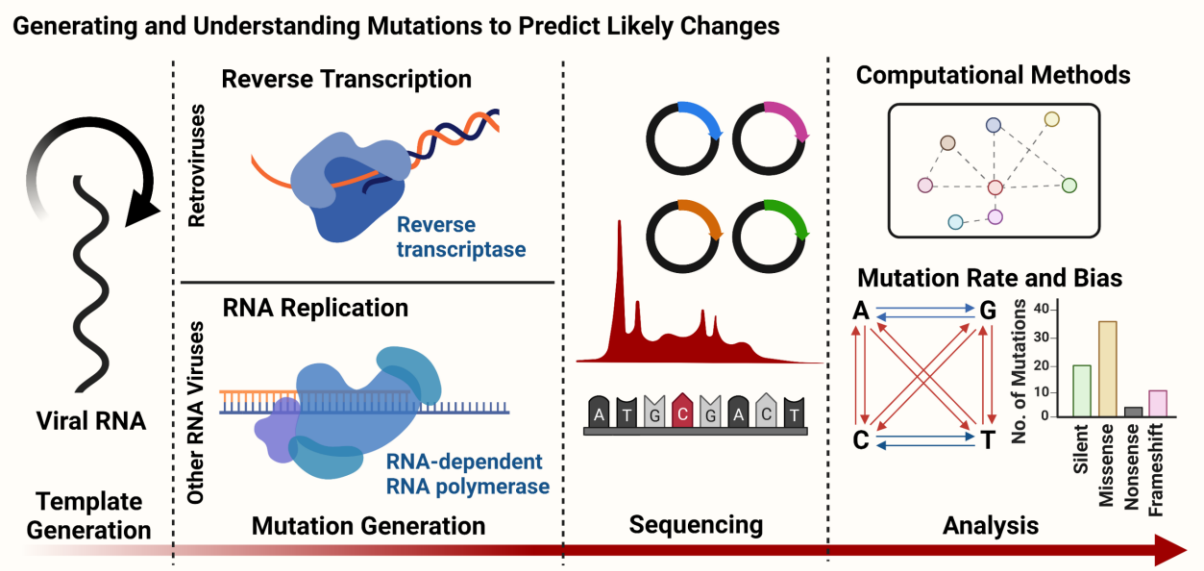

Figure 3. Schematic of mutation prediction of RNA viruses that can be generated *in vitro* (adapted from (88)) to determine innate viral polymerase mutation rates, hotspots and biases for the incorporation into computational methods for better prediction. Such a platform would not only provide insights to the viral polymerases for augmentation towards lethal mutagenesis, but also improve preparedness to identify conserved regions and identify potential mutations of concern that can lead to escape to guide drug and diagnostic kit development. Created with BioRender.com.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 November 2021 doi:10.20944/preprints202109.0501.v2

Experimentally, there is already notable progress in predicting mutations through the understanding of molecular biology and the innate biases of polymerases involved in replication (**Fig. 3**). While such predictions are often deemed stochastic and unpredictable (89), recent evidence has shown certain propensities in the genetic code (90) can confer a certain degree of predictability. It is certain that with more knowledge about the target biology, the less stochastic and more predictable mutations in the given target will be. This is demonstrated recently where even without immune or drug selection pressures, HIV reverse transcriptase displayed clear mutational biases in specific locations in HIV genes within a mimic of a single replication cycle (88). Given that a high percentage of the generated mutations were found in patients, the rise of drug-resistant mutations in the HIV genes were not entirely stochastic nor required intense immune or drug selection pressures to emerge. As the 'Godzilla' of fast mutating viruses, studies showed that the cross resistance for protease inhibitors (91) and reverse transcriptase inhibitors (92) could be investigated by network analysis. Such analysis has showed how mutations can confer not just resistance to one inhibitor, but to also cross-resistance. Thereby, in the vein of MID3, structural modelling studies can reveal the possible effects of emerging mutations with respect to drug resistance given the constraints of the mutations to the functional fitness of the target. It is also in this area that therapeutics which augment lethal mutagenesis for error catastrophe (93,94) could be an alternative strategy for undruggable targets.

 doi:10.20944/preprints202109.0501.v2

Enabling better direct inhibition: Lessons from Her2 and IgM

While alternative strategies exist in non-competitive non-nucleoside RT inhibitors or non-nucleoside reverse transcriptase inhibitors (NNRTIs) (95) and the above-mentioned lethal-mutagenesis-based (96), the direct inhibition of targets is the most common therapeutic approach. This approach puts the focus of epitopes at the active/functional sites as key targets, to which the consideration of the antibody can augment the success given that larger molecules may inhibit binding better.

From the example of steric hindrances of Trastuzumab and Pertuzumab IgM to Her2, Pertuzumab IgM inhibited cell proliferation better than Pertuzumab IgG1 (49). This was interesting because Pertuzumab IgM being larger/multimeric, may have prevented Her2 homodimerization/activation (necessary for its oncogenic effects (97)) better than its IgG1 counterpart based on the stoichiometry of binding sites. This is especially so given that the experiments were performed *in vitro* without effector immune cell effects and that unlike the Trastuzumab binding site (98), the Pertuzumab binding site directly inhibited homodimerization (99).

Through steric hindrances by larger IgM molecules, the selection of epitopes directly at/near the active or dimerization sites allow for direct therapeutic inhibition as opposed to mere immune tagging for immune effector cells. When coupled with the use of multimeric or conjugated antibodies, better direct effects can be achieved (**Fig. 4A**). Such an approach is relevant not only in oncology where over-expressed biomarkers lead to proliferation, but also in infectious diseases to neutralize virus attachment and entry (100). With further understanding of viral spike proteins, epitopes that facilitate direct inhibition of potentially deadly viral pathological effects ought to be considered. Given the recent evidence showing the SARS-CoV-2 spike to also possess superantigenic characteristics that can cause hyperinflammation leading to multisystem inflammatory syndrome in children (MIS-C) and

 doi:10.20944/preprints202109.0501.v2

cytokine storms in adult COVID-19 (101), such regions can be targeted through epitope selection of the region for mitigating the clinical morbidity, providing a much needed paradigm shift in treating viral infections (102) with pathogenesis roots in superantigenic activity. Although antibodies may already bind to such superantigenic motifs sites, further in-depth consideration of such recognition sites (103) will facilitate epitope selection for direct viral pathogenicity blocking in addition to that of attachment and entry.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 November 2021 doi:10.20944/preprints202109.0501.v2

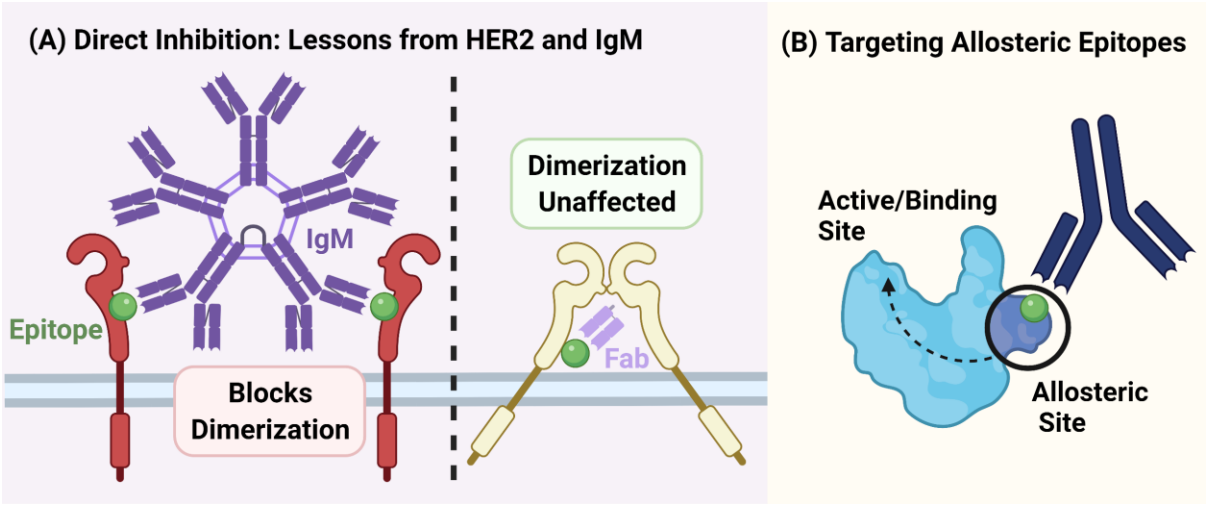

Figure 4 Choosing the epitopes for their potential functionality to (**A**) directly inhibit dimerization in the case of Pertuzumab IgM inhibiting Her2 dimerization better than its IgG counterpart due to its larger size (49). In such cases, an antibody fragment such as a Fab alone may not necessarily provide sufficient steric hindrance to block dimerization. In the event where direct blocking may not be suitable, the targeting of (**B**) allosteric epitopes to influence active or functional sites can also be explored to increase target areas in the same target antigen. Created with BioRender.com.

 doi:10.20944/preprints202109.0501.v2

Evident from oncology and HIV virology requiring multi-pronged interventions, where in the former, there is combinatorial synergistic use of Trastuzumab and Pertuzumab (47,104), and in the latter, combinatorial antiretroviral therapy (ART) (105), there is still much to consider in additional epitopes.

The addition of more targetable sites within the same antigen especially in the absence of other ideal epitopes for direct intervention can open the way for allosteric epitopes. With clear lessons in the use of NNRTI and recent evidence of a non-enzymatic subunit being potentially druggable in HIV (106), distal epitopes can be used to influence the enzyme/target's active sites by non-competitive antibodies (**Fig. 4B**). Allosteric sites may not necessary elicit obvious shape changes (107) and are thus often modulators than direct activators/inhibitors (see examples: IgE (76), FcεRI (108), IgA (59), microbial targets (109)). Such dampening of activity in oncology or virology can be what is precisely required at times to permit the necessary gain by the immune system. Furthermore, these allosteric sites in microbial and viral proteins could drive function-crippling mutations through drug resistance mechanisms at the drug binding site such as the non-cleavage site mutations contributing to drug resistance associated protein fitness compensation in HIV-1 Gag (110,111). These sites could also be epitopes for intrabodies to prevent drug-resistance or reduce protein fitness and be leveraged upon to dissociate already bound complexes by increasing dissociation (e.g. in the case of IgE-FcERI (108)).

With additional areas to target apart from active/functional sites, it may be easier to find common allosteric sites to allow the development of broad-spectrum targeting antibodies, and such widening of the search field may overcome constraints present in extracellular localization. While the allosteric analysis is currently most effectively performed computationally, it is important to be as inclusive of the entire protein structure/model as holistically possible to study the distal effects in a counterintuitive manner when compared to

 doi:10.20944/preprints202109.0501.v2

raising antibodies based on the reductionist approach in experiments of using purified recombinant proteins, regions or peptides. Though new databases such as the Post Translational Modification Structural Database (PTM-SD) (112,113) now include annotated posttranslational modification, the majority of crystal/ nuclear magnetic resonance (NMR) structures in structural databases without post-translational modifications can result in incongruency of findings, especially when using cell lines with varying glycosylation patterns (114).

Applications, Solutions, and Conclusion - Holistic Epitope selection

Sagacity in target and epitope can maximize the success for vaccine, therapeutics, and diagnostic development. Apart from deep understanding of the target biology, there are numerous factors from accessibility at the cellular and molecular level, the inducibility, occlusion, conservation at family and species levels, influence by distal allosteric sites, that can now be considered not only individually but together. Coupling such efforts with sagacious antibody development (6) that can involve swapping of elements (115), the very expensive developmental processes can be made more cost and time efficient.

 doi:10.20944/preprints202109.0501.v2

Financial Interests: The authors declare no competing financial interests.

Acknowledgements: The understanding that contributed to the insights were funded by Joint Council Office, Agency for Science, Technology, and Research, Singapore under Grant number JCO1334i00050. The staff are funded by the National Research Foundation (NRF) Singapore grant to Experimental Drug Development Centre (EDDC). Figures were created with BioRender.com.

Author contributions:

Conceptualization: SKEG

Visualization: JYY, SXP

Supervision: SKEG

Writing – original draft: SKEG

Writing – review & editing: SKEG, SXP, JYY

References

1. El Jaddaoui, I., Allali, M., Raoui, S., Sehli, S., Habib, N., Chaouni, B., Al Idrissi, N., Benslima, N., Maher, W., Benrahma, H. *et al.* (2021) A review on current diagnostic techniques for COVID-19. *Expert Review of Molecular Diagnostics*, **21**, 141-160.
2. Alpdagtas, S., Ilhan, E., Uysal, E., Sengor, M., Ustundag, C.B. and Gunduz, O. (2020) Evaluation of current diagnostic methods for COVID-19. *APL Bioeng*, **4**, 041506.
3. U.S. Food and Drug Administration. (2021) Genetic Variants of SARS-CoV-2 May Lead to False Negative Results with Molecular Tests for Detection of SARS-CoV-2 - Letter to Clinical Laboratory Staff and Health Care Providers.
4. U.S. Food and Drug Administration. (2021) SARS-CoV-2 Viral Mutations: Impact on COVID-19 Tests.
5. Eguia, R.T., Crawford, K.H.D., Stevens-Ayers, T., Kelnhofer-Millevolte, L., Greninger, A.L., Englund, J.A., Boeckh, M.J. and Bloom, J.D. (2021) A human coronavirus evolves antigenically to escape antibody immunity. *PLOS Pathogens*, **17**, e1009453.
6. Ling, W.-L., Lua, W.-H. and Gan, S.K.-E. (2020) Sagacity in antibody humanization for therapeutics, diagnostics and research purposes: considerations of antibody elements and their roles. *Antibody Therapeutics*, **3**, 71-79.
7. Krishnaswami, S., Austin, D., Della Pasqua, O., Gastonguay, M.R., Gobburu, J., van der Graaf, P.H., Ouellet, D., Tannenbaum, S. and Visser, S.A.G. (2020) MID3: Mission Impossible or Model-Informed Drug Discovery and Development? Point-Counterpoint Discussions on Key Challenges. *Clinical Pharmacology & Therapeutics*, **107**, 762-772.
8. Marshall, S., Madabushi, R., Manolis, E., Krudys, K., Staab, A., Dykstra, K. and Visser, S.A.G. (2019) Model-Informed Drug Discovery and Development: Current Industry Good Practice and Regulatory Expectations and Future Perspectives. *CPT: Pharmacometrics & Systems Pharmacology*, **8**, 87-96.
9. U.S. Food and Drug Administration. (2021) Model-Informed Drug Development Pilot Program.
10. Sou, T., Soukarieh, F., Williams, P., Stocks, M.J., Cámara, M. and Bergström, C.A.S. (2020) Model-Informed Drug Discovery and Development in Pulmonary Delivery: Biopharmaceutical Pharmacometric Modeling for Formulation Evaluation of Pulmonary Suspensions. *ACS Omega*, **5**, 25733-25746.
11. Ascoli, C.A. (2021) Could mutations of SARS-CoV-2 suppress diagnostic detection? *Nature Biotechnology*, **39**, 274-275.
12. Meuleman, T.J., Dunlop, J.I., Owsianka, A.M., van de Langemheen, H., Patel, A.H. and Liskamp, R.M.J. (2018) Immobilization by Surface Conjugation of Cyclic Peptides for Effective Mimicry of the HCV-Envelope E2 Protein as a Strategy toward Synthetic Vaccines. *Bioconjugate Chemistry*, **29**, 1091-1101.
13. Li, W., Joshi, M.D., Singhania, S., Ramsey, K.H. and Murthy, A.K. (2014) Peptide Vaccine: Progress and Challenges. *Vaccines*, **2**, 515-536.
14. Geysen, H.M., Rodda, S.J. and Mason, T.J. (1986) A priori delineation of a peptide which mimics a discontinuous antigenic determinant. *Molecular Immunology*, **23**, 709-715.
15. Tobias, J., Battin, C., De Sousa Linhares, A., Lebens, M., Baier, K., Ambroz, K., Drinić, M., Högler, S., Inic-Kanada, A., Garner-Spitzer, E. *et al.* (2020) A New Strategy Toward B Cell-Based Cancer Vaccines by Active Immunization With Mimotopes of Immune Checkpoint Inhibitors. *Frontiers in Immunology*, **11**.

16. Sanchez-Trincado, J.L., Gomez-Perosanz, M. and Reche, P.A. (2017) Fundamentals and Methods for T- and B-Cell Epitope Prediction. *Journal of Immunology Research*, **2017**, 2680160.
17. Hingorani, A.D., Kuan, V., Finan, C., Kruger, F.A., Gaulton, A., Chopade, S., Sofat, R., MacAllister, R.J., Overington, J.P., Hemingway, H. *et al.* (2019) Improving the odds of drug development success through human genomics: modeling study. *Sci Rep*, **9**, 18911.
18. Grignolo, A. and Pretorius, S. (2016) Phase III trial failures: costly, but preventable. *Applied Clinical Trials*.
19. Kolaskar, A.S. and Tongaonkar, P.C. (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Letters*, **276**, 172-174.
20. Emini, E.A., Hughes, J.V., Perlow, D.S. and Boger, J. (1985) Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *Journal of Virology*, **55**, 836-839.
21. Karplus, P.A. and Schulz, G.E. (1985) Prediction of chain flexibility in proteins. *Naturwissenschaften*, **72**, 212-213.
22. Parker, J.M.R., Guo, D. and Hodges, R.S. (1986) New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and x-ray-derived accessible sites. *Biochemistry*, **25**, 5425-5432.
23. Smith, C.C., Chai, S., Washington, A.R., Lee, S.J., Landori, E., Field, K., Garness, J., Bixby, L.M., Selitsky, S.R., Parker, J.S. *et al.* (2019) Machine-learning prediction of tumor immunogenicity in the selection of therapeutic epitopes. *Cancer Immunology Research*, **7**.
24. Denisova, G.F., Denisov, D.A. and Bramson, J.L. (2010) Applying bioinformatics for antibody epitope prediction using affinity-selected mimotopes relevance for vaccine design. *Immunome Research*, **6**, S6.
25. Phloyphisut, P., Pornputtapong, N., Sriswasdi, S. and Chuangsuwanich, E. (2019) MHCSeqNet: a deep neural network model for universal MHC binding prediction. *BMC Bioinformatics*, **20**, 270.
26. Noumi, T., Inoue, S., Fujita, H., Sadamitsu, K., Sakaguchi, M., Tenma, A. and Nakagami, H. (2021) Epitope Prediction of Antigen Protein Using Attention-based LSTM Network. *Journal of Information Processing*, **29**, 321-327.
27. Trenevska, I., Li, D. and Banham, A.H. (2017) Therapeutic Antibodies against Intracellular Tumor Antigens. *Frontiers in Immunology*, **8**.
28. Singh, K., Ejaz, W., Dutta, K. and Thayumanavan, S. (2019) Antibody Delivery for Intracellular Targets: Emergent Therapeutic Potential. *Bioconjugate Chemistry*, **30**, 1028-1041.
29. Slastnikova, T.A., Ulasov, A.V., Rosenkranz, A.A. and Sobolev, A.S. (2018) Targeted Intracellular Delivery of Antibodies: The State of the Art. *Frontiers in Pharmacology*, **9**.
30. Thura, M., Al-Aidaroos, A.Q., Gupta, A., Chee, C.E., Lee, S.C., Hui, K.M., Li, J., Guan, Y.K., Yong, W.P., So, J. *et al.* (2019) PRL3-zumab as an immunotherapy to inhibit tumors expressing PRL3 oncoprotein. *Nature Communications*, **10**, 2484.
31. Wintjens, R., Bifani, A.M. and Bifani, P. (2020) Impact of glycan cloud on the B-cell epitope prediction of SARS-CoV-2 Spike protein. *npj Vaccines*, **5**, 81.
32. Davis, D., Stephens, D.M., Willers, C. and Lachmann, P.J. (1990) Glycosylation Governs the Binding of Antipeptide Antibodies to Regions of Hypervariable Amino Acid Sequence within Recombinant gp120 of Human Immunodeficiency Virus Type 1. *Journal of General Virology*, **71**, 2889-2898.

33. Rieder, F.J.J., Biebl, J., Kastner, M.-T., Schneider, M., Jungbauer, C., Redlberger-Fritz, M., Britt, W.J., Kundi, M. and Steininger, C. (2016) Microbial Cryptotopes are Prominent Targets of B-cell Immunity. *Scientific Reports*, **6**, 31657.
34. Lindesmith, L.C., Mallory, M.L., Debbink, K., Donaldson, E.F., Brewer-Jensen, P.D., Swann, E.W., Sheahan, T.P., Graham, R.L., Beltramello, M., Corti, D. *et al.* (2018) Conformational Occlusion of Blockade Antibody Epitopes, a Novel Mechanism of GII.4 Human Norovirus Immune Evasion. *mSphere*, **3**, e00518-00517.
35. Adachi, Y., Tonouchi, K., Nithichanon, A., Kuraoka, M., Watanabe, A., Shinnakasu, R., Asanuma, H., Ainai, A., Ohmi, Y., Yamamoto, T. *et al.* (2019) Exposure of an occluded hemagglutinin epitope drives selection of a class of cross-protective influenza antibodies. *Nature Communications*, **10**, 3883.
36. Francica, J.R., Varela-Rohena, A., Medvec, A., Plesa, G., Riley, J.L. and Bates, P. (2010) Steric Shielding of Surface Epitopes and Impaired Immune Recognition Induced by the Ebola Virus Glycoprotein. *PLOS Pathogens*, **6**, e1001098.
37. Moscoso, C.G., Xing, L., Hui, J., Hu, J., Kalkhoran, M.B., Yenigun, O.M., Sun, Y., Paavolainen, L., Martin, L., Vahlne, A. *et al.* (2014) Trimeric HIV Env provides epitope occlusion mediated by hypervariable loops. *Scientific Reports*, **4**, 7025.
38. Warnock, M.G. and Goodacre, J.A. (1997) Cryptic T-cell epitopes and their role in the pathogenesis of autoimmune diseases. *Rheumatology*, **36**, 1144-1150.
39. Kang, H.-E., Bian, J., Kane, S.J., Kim, S., Selwyn, V., Crowell, J., Bartz, J.C. and Telling, G.C. (2020) Incomplete glycosylation during prion infection unmasks a prion protein epitope that facilitates prion detection and strain discrimination. *Journal of Biological Chemistry*, **295**, 10420-10433.
40. Starr, T.N., Czudnochowski, N., Liu, Z., Zatta, F., Park, Y.-J., Addetia, A., Pinto, D., Beltramello, M., Hernandez, P., Greaney, A.J. *et al.* (2021) SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape. *Nature*, **597**, 97-102.
41. Tortorici, M.A., Czudnochowski, N., Starr, T.N., Marzi, R., Walls, A.C., Zatta, F., Bowen, J.E., Jaconi, S., Di Iulio, J., Wang, Z. *et al.* (2021) Broad sarbecovirus neutralization by a human monoclonal antibody. *Nature*, **597**, 103-108.
42. Minor, P.D., Pipkin, P.A., Hockley, D., Schild, G.C. and Almond, J.W. (1984) Monoclonal antibodies which block cellular receptors of poliovirus. *Virus Research*, **1**, 203-212.
43. Ludwig, R.J., Vanhoorelbeke, K., Leypoldt, F., Kaya, Z., Bieber, K., McLachlan, S.M., Komorowski, L., Luo, J., Cabral-Marques, O., Hammers, C.M. *et al.* (2017) Mechanisms of Autoantibody-Induced Pathology. *Frontiers in Immunology*, **8**.
44. Weigel, J.A., Raymond, R.C., McGary, C., Singh, A. and Weigel, P.H. (2003) A Blocking Antibody to the Hyaluronan Receptor for Endocytosis (HARE) Inhibits Hyaluronan Clearance by Perfused Liver *. *Journal of Biological Chemistry*, **278**, 9808-9812.
45. Su, C.T.-T., Lua, W.-H., Poh, J.-J., Ling, W.-L., Yeo, J.Y. and Gan, S.K.-E. (2021) Molecular Insights of Nickel Binding to Therapeutic Antibodies as a Possible New Antibody Superantigen. *Frontiers in Immunology*, **12**.
46. Fuentes, G., Scaltriti, M., Baselga, J. and Verma, C.S. (2011) Synergy between trastuzumab and pertuzumab for human epidermal growth factor 2 (Her2) from colocalization: an in silicobased mechanism. *Breast Cancer Research*, **13**, R54.
47. Lua, W.-H., Gan, S.K.-E., Lane, D.P. and Verma, C.S. (2015) A search for synergy in the binding kinetics of Trastuzumab and Pertuzumab whole and F(ab) to Her2. *npj Breast Cancer*, **1**, 15012.
48. Matos, D.M. (2021) Steric hindrance: A practical (and frequently forgotten) problem in flow cytometry. *Cytometry Part B: Clinical Cytometry*, **100**, 397-401.

49. Samsudin, F., Yeo, J.Y., Gan, S.K.-E. and Bond, P.J. (2020) Not all therapeutic antibody isotypes are equal: the case of IgM versus IgG in Pertuzumab and Trastuzumab. *Chemical Science*, **11**, 2843-2854.
50. Lua, W.-H., Ling, W.-L., Yeo, J.Y., Poh, J.-J., Lane, D.P. and Gan, S.K.-E. (2018) The effects of Antibody Engineering CH and CL in Trastuzumab and Pertuzumab recombinant models: Impact on antibody production and antigen-binding. *Scientific Reports*, **8**, 718.
51. Yerabham, A. and Ho, M. (2021) A novel IgM intranasal intervention against SARS-CoV-2. *Antibody Therapeutics*, **4**, 171-174.
52. Duffy, K.T., Wharton, P.J., Johnson, J.D., New, L. and Holliman, R.E. (1989) Assessment of immunoglobulin-M immunosorbent agglutination assay (ISAGA) for detecting toxoplasma specific IgM. *Journal of Clinical Pathology*, **42**, 1291-1295.
53. Cambiaso, C.L., Galanti, L.M., Leautaud, P. and Masson, P.L. (1992) Latex agglutination assay of human immunoglobulin M antitoxoplasma antibodies which uses enzymatically treated antigen-coated particles. *J Clin Microbiol*, **30**, 882-888.
54. Ryu, W.-S. (2017) In Ryu, W.-S. (ed.), *Molecular Virology of Human Pathogenic Viruses*. Academic Press, Boston, pp. 47-62.
55. Payne, S. (2017) In Payne, S. (ed.), *Viruses*. Academic Press, pp. 37-52.
56. Mazor, Y., Hansen, A., Yang, C., Chowdhury, P.S., Wang, J., Stephens, G., Wu, H. and Dall'Acqua, W.F. (2015) Insights into the molecular basis of a bispecific antibody's target selectivity. *mAbs*, **7**, 461-469.
57. Haber, L., Olson, K., Kelly, M.P., Crawford, A., DiLillo, D.J., Tavaré, R., Ullman, E., Mao, S., Canova, L., Sineshchekova, O. *et al.* (2021) Generation of T-cell-redirecting bispecific antibodies with differentiated profiles of cytokine release and biodistribution by CD3 affinity tuning. *Scientific Reports*, **11**, 14397.
58. Einsele, H., Borghaei, H., Orlowski, R.Z., Subklewe, M., Roboz, G.J., Zugmaier, G., Kufer, P., Iskander, K. and Kantarjian, H.M. (2020) The BiTE (bispecific T-cell engager) platform: Development and future potential of a targeted immuno-oncology therapy across tumor types. *Cancer*, **126**, 3192-3201.
59. Su, C.T.-T., Lua, W.-H., Ling, W.-L. and Gan, S.K.-E. (2018) Allosteric Effects between the Antibody Constant and Variable Regions: A Study of IgA Fc Mutations on Antigen Binding. *Antibodies*, **7**, 20.
60. Kemp, T.J., Hildesheim, A., Safaeian, M., Dauner, J.G., Pan, Y., Porras, C., Schiller, J.T., Lowy, D.R., Herrero, R. and Pinto, L.A. (2011) HPV16/18 L1 VLP vaccine induces cross-neutralizing antibodies that may mediate cross-protection. *Vaccine*, **29**, 2011-2014.
61. Lu, R.-M., Hwang, Y.-C., Liu, I.J., Lee, C.-C., Tsai, H.-Z., Li, H.-J. and Wu, H.-C. (2020) Development of therapeutic antibodies for the treatment of diseases. *Journal of Biomedical Science*, **27**, 1.
62. Mullard, A. (2020) FDA approves antibody cocktail for Ebola virus. *Nat Rev Drug Discov*, **19**, 827.
63. Hayer, J., Jadeau, F., Deléage, G., Kay, A., Zoulim, F. and Combet, C. (2012) HBVdb: a knowledge database for Hepatitis B Virus. *Nucleic Acids Research*, **41**, D566-D570.
64. Kuiken, C., Yusim, K., Boykin, L. and Richardson, R. (2004) The Los Alamos hepatitis C sequence database. *Bioinformatics*, **21**, 379-384.
65. Kuiken, C., Thurmond, J., Dimitrijevic, M. and Yoon, H. (2011) The LANL hemorrhagic fever virus database, a new platform for analyzing biothreat viruses. *Nucleic Acids Research*, **40**, D587-D592.
66. Los Alamos National Laboratory. (2021) HIV Sequence Database.

67. Kuiken, C., Korber, B. and Shafer, R.W. (2003) HIV sequence databases. *AIDS Rev*, **5**, 52-61.
68. Kantor, R., Machekano, R., Gonzales, M.J., Dupnik, K., Schapiro, J.M. and Shafer, R.W. (2001) Human Immunodeficiency Virus Reverse Transcriptase and Protease Sequence Database: an expanded data model integrating natural language text and sequence analysis programs. *Nucleic Acids Research*, **29**, 296-299.
69. Shu, Y. and McCauley, J. (2017) GISAID: Global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance*, **22**, 30494.
70. Hadfield, J., Megill, C., Bell, S.M., Huddleston, J., Potter, B., Callender, C., Sagulenko, P., Bedford, T. and Neher, R.A. (2018) Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics*, **34**, 4121-4123.
71. Bottoms, C.A., White, T.A. and Tanner, J.J. (2006) Exploring structurally conserved solvent sites in protein families. *Proteins: Structure, Function, and Bioinformatics*, **64**, 404-421.
72. Madu, I.G., Roth, S.L., Belouzard, S. and Whittaker, G.R. (2009) Characterization of a Highly Conserved Domain within the Severe Acute Respiratory Syndrome Coronavirus Spike Protein S2 Domain with Characteristics of a Viral Fusion Peptide. *Journal of Virology*, **83**, 7411-7421.
73. Sundaravaradan, V., Hahn, T. and Ahmad, N. (2005) Conservation of functional domains and limited heterogeneity of HIV-1 reverse transcriptase gene following vertical transmission. *Retrovirology*, **2**, 36.
74. Okayama, Y., Matsumoto, H., Odajima, H., Takahagi, S., Hide, M. and Okubo, K. (2020) Roles of omalizumab in various allergic diseases. *Allergology International*, **69**, 167-177.
75. van Loghem, E., Aalberse, R.C. and Matsumoto, H. (1984) A Genetic Marker of Human IgE Heavy Chains, Em(1)1. *Vox Sanguinis*, **46**, 195-206.
76. Lua, W.-H., Su, C.T.-T., Yeo, J.Y., Poh, J.-J., Ling, W.-L., Phua, S.-X. and Gan, S.K.-E. (2019) Role of the IgE variable heavy chain in FcεRIα and superantigen binding in allergy and immunotherapy. *Journal of Allergy and Clinical Immunology*, **144**, 514-523.e515.
77. Wang, R., Chen, J., Gao, K. and Wei, G.-W. (2021) Vaccine-escape and fast-growing mutations in the United Kingdom, the United States, Singapore, Spain, India, and other COVID-19-devastated countries. *Genomics*, **113**, 2158-2170.
78. Pardi, N., Hogan, M.J., Porter, F.W. and Weissman, D. (2018) mRNA vaccines — a new era in vaccinology. *Nature Reviews Drug Discovery*, **17**, 261-279.
79. Weeks, S.A., Lee, C.A., Zhao, Y., Smidansky, E.D., August, A., Arnold, J.J. and Cameron, C.E. (2012) A polymerase mechanism-based strategy for viral attenuation and vaccine development. *J Bio Chem*, **287**, 31618-31622.
80. Chen, T.-C., Chang, H.-Y., Lin, P.-F., Chern, J.-H., Hsu, J.T.-A., Chang, C.-Y. and Shih, S.-R. (2009) Novel antiviral agent DTriP-22 targets RNA-dependent RNA polymerase of Enterovirus 71. *Antimicrob Agents Chemother*, **53**, 2740-2747.
81. Ozturk, K., Dow, M., Carlin, D.E., Bejar, R. and Carter, H. (2018) The Emerging Potential for Network Analysis to Inform Precision Cancer Medicine. *Journal of Molecular Biology*, **430**, 2875-2899.
82. Özcan ŞİmŞek, N.Ö., ÖzgÜr, A. and GÜrgen, F. (2019) Statistical representation models for mutation information within genomic data. *BMC Bioinformatics*, **20**, 324.
83. Yin, R., Luusua, E., Dabrowski, J., Zhang, Y. and Kwoh, C.K. (2020) Tempel: time-series mutation prediction of influenza A viruses via attention-based recurrent neural networks. *Bioinformatics*, **36**, 2697-2704.

84. Yin, R., Tran, V.H., Zhou, X., Zheng, J. and Kwoh, C.K. (2018) Predicting antigenic variants of H1N1 influenza virus based on epidemics and pandemics using a stacking model. *PLOS ONE*, **13**, e0207777.
85. Goh, W.W.B. and Wong, L. (2018) Why breast cancer signatures are no better than random signatures explained. *Drug Discovery Today*, **23**, 1818-1823.
86. Omichessan, H., Severi, G. and Perduca, V. (2019) Computational tools to detect signatures of mutational processes in DNA from tumours: A review and empirical comparison of performance. *PLOS ONE*, **14**, e0221235.
87. Yeo, J.Y. and Gan, S.K.-E. (2021) Peering Into Avian Influenza A(H5N8) for a Framework towards Pandemic Preparedness. *Preprints*.
88. Yeo, J.Y., Koh, D.W.-S., Yap, P., Goh, G.-R. and Gan, S.K.-E. (2021) Spontaneous Mutations in HIV-1 Gag, Protease, RT p66 in the First Replication Cycle and How They Appear: Insights from an In Vitro Assay on Mutation Rates and Types. *International Journal of Molecular Sciences*, **22**, 370.
89. Gold, B. (2017) Somatic mutations in cancer: Stochastic versus predictable. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, **814**, 37-46.
90. Chan, K.-F., Koukouravas, S., Yeo, J.Y., Koh, D.W.-S. and Gan, S.K.-E. (2020) Probability of change in life: Amino acid changes in single nucleotide substitutions. *Biosystems*, **193-194**, 104135.
91. Su, C.T.-T., Ling, W.-L., Lua, W.-H., Haw, Y.-X. and Gan, S.K.-E. (2016) Structural analyses of 2015-updated drug-resistant mutations in HIV-1 protease: an implication of protease inhibitor cross-resistance. *BMC Bioinformatics*, **17**, 500.
92. Chiang, Ron Z.-H., Gan, Samuel K.-E. and Su, Chinh T.-T. (2018) A computational study for rational HIV-1 non-nucleoside reverse transcriptase inhibitor selection and the discovery of novel allosteric pockets for inhibitor design. *Bioscience Reports*, **38**.
93. Summers, J. and Litwin, S. (2006) Examining The Theory of Error Catastrophe. *Journal of Virology*, **80**, 20-26.
94. Bull, J.J., Sanjuán, R. and Wilke, C.O. (2007) Theory of Lethal Mutagenesis for Viruses. *Journal of Virology*, **81**, 2930-2939.
95. Usach, I., Melis, V. and Peris, J.-E. (2013) Non-nucleoside reverse transcriptase inhibitors: a review on pharmacokinetics, pharmacodynamics, safety and tolerability. *Journal of the International AIDS Society*, **16**, 18567.
96. Perales, C. and Domingo, E. (2016) In Domingo, E. and Schuster, P. (eds.), *Quasispecies: From Theory to Experimental Systems*. Springer International Publishing, Cham, pp. 323-339.
97. Brennan, P.J., Kumogai, T., Berezov, A., Murali, R. and Greene, M.I. (2000) HER2/Neu: mechanisms of dimerization/oligomerization. *Oncogene*, **19**, 6093-6101.
98. Cho, H.-S., Mason, K., Ramyar, K.X., Stanley, A.M., Gabelli, S.B., Denney, D.W. and Leahy, D.J. (2003) Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature*, **421**, 756-760.
99. Franklin, M.C., Carey, K.D., Vajdos, F.F., Leahy, D.J., de Vos, A.M. and Sliwkowski, M.X. (2004) Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex. *Cancer Cell*, **5**, 317-328.
100. Corti, D. and Lanzavecchia, A. (2013) Broadly Neutralizing Antiviral Antibodies. *Annual Review of Immunology*, **31**, 705-742.
101. Cheng, M.H., Zhang, S., Porritt, R.A., Noval Rivas, M., Paschold, L., Willscher, E., Binder, M., Arditi, M. and Bahar, I. (2020) Superantigenic character of an insert unique to SARS-CoV-2 spike supported by skewed TCR repertoire in patients with hyperinflammation. *Proceedings of the National Academy of Sciences*, **117**, 25254-25262.

102. Scaglioni, V. and Soriano, E.R. (2020) Are superantigens the cause of cytokine storm and viral sepsis in severe COVID-19? Observations and hypothesis. *Scandinavian Journal of Immunology*, **92**, e12944.
103. Deacy, A.M., Gan, S.K.-E. and Derrick, J. (2021) Superantigen Recognition and Interactions: Functions, Mechanisms and Applications (Provisionally accepted). *Frontiers in Immunology*.
104. Maly, J.J. and Macrae, E.R. (2014) Pertuzumab in Combination with Trastuzumab and Chemotherapy in the Treatment of HER2-Positive Metastatic Breast Cancer: Safety, Efficacy, and Progression Free Survival. *Breast Cancer: Basic and Clinical Research*, **8**, BCBCR.S9032.
105. Bhatti, A.B., Usman, M. and Kandi, V. (2016) Current Scenario of HIV/AIDS, Treatment Options, and Major Challenges with Compliance to Antiretroviral Therapy. *Cureus*, **8**, e515-e515.
106. Chan, K.-F., Su, C.T.-T., Krah, A., Phua, S.-X., Yeo, J.Y., Ling, W.-L., Bond, P.J. and Gan, S.K.-E. (2020) An Alternative HIV-1 Non-Nucleoside Reverse Transcriptase Inhibition Mechanism: Targeting the p51 Subunit. *Molecules*, **25**, 5902.
107. Tsai, C.-J., del Sol, A. and Nussinov, R. (2008) Allostery: Absence of a Change in Shape Does Not Imply that Allostery Is Not at Play. *Journal of Molecular Biology*, **378**, 1-11.
108. Phua, S.-X., Chan, K.-F., Su, Chinh T.-T., Poh, J.-J. and Gan, Samuel K.-E. (2019) Perspective: The promises of a holistic view of proteins—impact on antibody engineering and drug discovery. *Bioscience Reports*, **39**.
109. Blaise, P., Clevenbergh, P., Vaira, D., Moutschen, M. and Dellamonica, P. (2002) HIV resistance to antiretroviral drugs: mechanisms, genotypic and phenotypic resistance testing in clinical practice. *Acta Clinica Belgica*, **57**, 191-201.
110. Su, C.T.-T., Kwoh, C.-K., Verma, C.S. and Gan, S.K.-E. (2018) Modeling the full length HIV-1 Gag polyprotein reveals the role of its p6 subunit in viral maturation and the effect of non-cleavage site mutations in protease drug resistance. *Journal of Biomolecular Structure and Dynamics*, **36**, 4366-4377.
111. Samsudin, F., Gan, S.K.-E. and Bond, P.J. (2021) The impact of Gag non-cleavage site mutations on HIV-1 viral fitness from integrative modelling and simulations. *Computational and Structural Biotechnology Journal*, **19**, 330-342.
112. Arakhamia, T., Lee, C.E., Carlomagno, Y., Duong, D.M., Kundinger, S.R., Wang, K., Williams, D., DeTure, M., Dickson, D.W., Cook, C.N. *et al.* (2020) Posttranslational Modifications Mediate the Structural Diversity of Tauopathy Strains. *Cell*, **180**, 633-644.e612.
113. Craveur, P., Rebehmed, J. and de Brevern, A.G. (2014) PTM-SD: a database of structurally resolved and annotated posttranslational modifications in proteins. *Database*, **2014**.
114. Böhm, E., Seyfried, B.K., Dockal, M., Graninger, M., Hasslacher, M., Neurath, M., Konetschny, C., Matthiessen, P., Mitterer, A. and Scheiflinger, F. (2015) Differences in N-glycosylation of recombinant human coagulation factor VII derived from BHK, CHO, and HEK293 cells. *BMC Biotechnology*, **15**, 87.
115. Ling, W.-L., Lua, W.-H., Poh, J.-J., Yeo, J.Y., Lane, D.P. and Gan, S.K.-E. (2018) Effect of VH–VL Families in Pertuzumab and Trastuzumab Recombinant Production, Her2 and FcγIIA Binding. *Frontiers in Immunology*, **9**.