



Contents lists available at ScienceDirect

## Advanced Drug Delivery Reviews

journal homepage: [www.elsevier.com/locate/adr](http://www.elsevier.com/locate/adr)

## Current and prospective computational approaches and challenges for developing COVID-19 vaccines



Woochang Hwang<sup>a,1</sup>, Winnie Lei<sup>a,d,1</sup>, Nicholas M Katritsis<sup>a,c,1</sup>, Méabh MacMahon<sup>a,b,1</sup>, Kathryn Chapman<sup>a</sup>, Namshik Han<sup>a,\*</sup>

<sup>a</sup> Milner Therapeutics Institute, University of Cambridge, Cambridge, UK

<sup>b</sup> Centre for Therapeutics Discovery, LifeArc, Stevenage, UK

<sup>c</sup> Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge, UK

<sup>d</sup> Department of Surgery, University of Cambridge, Cambridge, UK

### ARTICLE INFO

#### Article history:

Received 23 October 2020

Revised 1 February 2021

Accepted 3 February 2021

Available online 6 February 2021

### ABSTRACT

SARS-CoV-2, which causes COVID-19, was first identified in humans in late 2019 and is a coronavirus which is zoonotic in origin. As it spread around the world there has been an unprecedented effort in developing effective vaccines. Computational methods can be used to speed up the long and costly process of vaccine development. Antigen selection, epitope prediction, and toxicity and allergenicity prediction are areas in which computational tools have already been applied as part of reverse vaccinology for SARS-CoV-2 vaccine development. However, there is potential for computational methods to assist further. We review approaches which have been used and highlight additional bioinformatic approaches and PK modelling as *in silico* methods which may be useful for SARS-CoV-2 vaccine design but remain currently unexplored. As more novel viruses with pandemic potential are expected to arise in future, these techniques are not limited to application to SARS-CoV-2 but also useful to rapidly respond to novel emerging viruses.

Crown Copyright © 2021 Published by Elsevier B.V. All rights reserved.

### Contents

1. Introduction	250
2. Antigen selection	251
2.1. SARS-CoV-2 proteins targeted	251
2.2. Antigen prediction based on viral genomic sequences	251
2.2.1. Antigen prediction based on cellular immune response	251
3. Epitope prediction	253
3.1. B cell epitope prediction tools	253
3.1.1. Continuous B cell epitope prediction in SARS-CoV-2	253
3.1.2. Discontinuous B cell epitope prediction in SARS-CoV-2	253
3.2. T cell epitope prediction tools	253
3.2.1. CD4+ T cell epitope prediction in SARS-CoV-2	256
3.2.2. CD8+ T cell epitope prediction in SARS-CoV-2	256
3.2.3. HLA coverage	259
3.2.4. TCR binding prediction	260
3.3. Epitope prediction evaluation and challenges	260
4. Immunogenicity prediction	260
4.1. Mathematical modelling on SARS-CoV-2 infection	260
4.2. PK/PD studies in vaccine design	261
4.2.1. Timing and location in vaccine efficacy	261

\* Corresponding author.

E-mail address: [n.han@milner.cam.ac.uk](mailto:n.han@milner.cam.ac.uk) (N. Han).

<sup>1</sup> These authors contributed equally.

<https://doi.org/10.1016/j.addr.2021.02.004>

0169-409X/Crown Copyright © 2021 Published by Elsevier B.V. All rights reserved.

4.2.2.	Predicting antigen immunogenicity	262
4.2.3.	Predicting adjuvant immunogenicity	262
4.2.4.	Predicting antigen ADME properties	263
4.2.5.	Incorporating epitope binding affinities	264
4.2.6.	Integrated PBPK models of vaccines	264
4.3.	Perspectives and challenges	264
5.	Toxicity and allergenicity prediction	265
5.1.	Toxicity prediction	265
5.1.1.	Non-specific toxicity predictions for SARS-CoV-2 vaccine epitopes	265
5.1.2.	Haemolytic toxicity predictions for SARS-CoV-2 vaccine epitopes	265
5.1.3.	Vaccine component toxicity predictions with computational network analysis	266
5.2.	Allergenicity prediction	266
5.2.1.	IgE binding affinity predictions for SARS-CoV-2 vaccine epitopes	266
5.3.	Predicting vaccine-induced enhanced disease	266
6.	Discussion	267
	Acknowledgement	268
	Competing interests	268
	Appendix A. Supplementary data	268
	References	268

## 1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which causes Coronavirus disease 2019 (COVID-19), was first identified in Wuhan in December 2019 and by December 2020 had spread to 191 Countries or Regions and infected over 75 million people, causing more than 1.5 million deaths [1]. Effects of SARS-CoV-2 infection in patients with COVID-19 can range from asymptomatic to life-threatening [2], with most symptomatic patients having moderate symptoms.

This pandemic has impacted wider society around the world, leading to restrictions on travel and severely disrupting global economies and many health services. In the absence of effective treatments to prevent the development of acute respiratory distress syndrome (ARDS) in COVID-19 patients, a vaccine is the fastest and most promising way to stop the spread of SARS-CoV-2 and bring the pandemic under control. Like Severe Acute Respiratory Syndrome (SARS-CoV) and Middle East Respiratory Syndrome coronavirus (MERS-CoV), SARS-CoV-2 is a coronavirus with a zoonotic origin, and it is likely that further zoonotic viruses with pandemic potential will emerge, making additional crises of this kind probable in the future [3]. To respond effectively to novel emerging infectious diseases, it is necessary to refine and establish technologies that can develop effective, safe vaccines, rapidly.

Vaccines provide adaptive immunity to disease and are a form of preventative medicine. Generally, vaccines are made of an attenuated (live, but weakened) or inactivated (killed with heat or chemicals) pathogen, or a specific subcomponent or subcomponents of the pathogen which are antigenic (subunit and conjugate vaccines). The first vaccine, verified by Edward Jenner, was the use of cowpox to prevent smallpox infection [4]. Vaccines have improved since Jenner's day in step with new scientific discoveries and the technology used to develop them. For instance, Goodpasture's invention of methods to propagate viruses in chicken eggs, led to mass production of vaccines including for typhus [5,6]. More recently, use of recombinant DNA technology led to the subunit vaccine for hepatitis B [7]. The ability to engineer synthetic RNA aided the current development of mRNA vaccine approaches [8]. mRNA vaccines encode an antigen of interest for translation in the body where an immune response is generated against it. The

first mRNA vaccines approved in history, BNT162b2 and mRNA-1273, have been recently approved for SARS-CoV-2.

Reverse vaccinology is an application that aids the development of new vaccines, initiated with pathogen genomic sequencing [9,10]. It was first applied successfully to meningococcus B to predict alternative antigens to develop a vaccine [11,12]. After 4 decades of unsuccessful attempts with conventional vaccine design, reverse vaccinology facilitated successful vaccine development through overcoming the challenge of molecular mimicry between the meningococci capsular polysaccharide and human protein [13]. Reverse vaccinology can be used to select an antigen for a novel vaccine which can generate an immune response through epitope selection and screening vaccine candidates *in silico*, speeding up the process of vaccine design and reducing its cost. An antigen is a molecule capable of producing an immune response, while an epitope is a site on an antigen which is recognised by the immune system. There are two types of epitopes, B cell epitopes and T cell epitopes. B cell epitopes are recognised by antibodies, whereas T cell epitopes are recognised by T cell receptors. The full SARS-CoV-2 genome was released in January 2020 allowing reverse vaccinology approaches to be applied in the development of a vaccine for SARS-CoV-2 as the virus was spreading around the world.

Modern vaccine design aims to maximise efficacy while minimising potential serious adverse effects, such as inducing an allergic reaction or toxic effects causing damage to the organism. To this end, newer approaches have shifted away from early live-attenuated and inactivated whole-pathogen vaccines, towards purified antigens (subunits), as these come with improved safety profile [14]. Funk *et al.* [15] in a snapshot of the global race for a SARS-CoV-2 vaccine, report that "protein-based" (subunit) vaccines represent the largest platform of development. However, subunit vaccines lack sufficient immunostimulatory capabilities on their own and are typically paired with adjuvants, which are compounds or formulations capable of providing the inflammatory cues needed to elicit the desired immune response [14].

Surprisingly, the knowledge that is currently available on vaccine absorption, distribution, metabolism, and excretion (ADME) processes is very limited [16]. However, a few PBPK models of adjuvants do exist in the literature and they may have relevance to the COVID-19 vaccine development process.

**Table 1**  
COVID-19 severity studies.

Study	Omics	Moderate	Severe	Cell types	Data	DEG analysis	Ref
Overmyer et al.	Bulk RNA-seq	51	55	PBMC	Available	EBSeq	[35]
Jain et al.	Bulk RNA-seq	10	3	PBMC	Available	DESeq2	[245]
Liu et al.	scRNA-seq	3	6	PBMC	NA	NA	[248]
Xu et al.	scRNA-seq	5	8	PBMC	Available	Seurat	[249]
Silvin et al.	scRNA-seq	1	2	PBMC	NA	NA	[250]
Arunachalam et al.	Bulk RNA-seq	4	12	PBMC	Available	DESeq2	[251]

In this review, we focus on computational techniques applied at the very early stage of the vaccine development pipeline for SARS-CoV-2. We explore computational approaches to antigen selection, epitope prediction, adjuvant selection, toxicology and allergenicity prediction and immune response modelling in the context of COVID-19.

## 2. Antigen selection

### 2.1. SARS-CoV-2 proteins targeted

The SARS-CoV-2 is made up of four structural proteins E (envelope), M (membrane), N (nucleocapsid), and S (spike), and several non-structural proteins (nsps) [17]. The S protein is an attractive vaccination candidate because it is involved in viral entry and elicited an immune response in the case of SARS-CoV and MERS-CoV studies [18]. An early study searching for vaccine candidates for SARS-CoV-2 used *in silico* methods to compare the sequence of the N and S proteins of SARS-CoV-2 to B and T cell epitopes derived from SARS-CoV which have been experimentally determined [18]. BNT162b2 [19], mRNA-1273 [20] and AZD1222 [21] are three recently approved vaccines for SARS-CoV-2, all of which use the S protein. BNT162b2 and mRNA-1273 are mRNA vaccines, both of which generate a prefusion-stabilised S protein and AZD1222 is a replication-deficient viral vector vaccine which uses a chimpanzee adenovirus as a vector with the S protein.

Most *in silico* epitope prediction studies also focused on the spike protein, although several studies focused on multiple structural proteins [22–24]. One study focused exclusively on the E protein [25]. Ong *et al.* proposed a 'cocktail' vaccine using both structural and non-structural proteins [26].

### 2.2. Antigen prediction based on viral genomic sequences

Virus replication only happens within cells, so blocking viral entry prevents the virus proliferating within the host. The S protein of SARS-CoV-2 is the part of the virus that enables it to enter our cells. If a person has antibodies that can recognise the S protein, this stops the virus in its tracks. That is why most of the vaccine candidates are focusing on the S protein as an antigen. The focus on the S protein is logical, however mutations have already arisen which affect the function of this protein [27]. If further such mutations arise which affect vaccine efficacy, antigens based on other viral proteins will be required. Since such an outcome isn't unlikely, we need to identify alternative antigen candidates as a backup to respond in a timely manner if such problematic mutations arise. In addition to the S protein, other proteins, such as the N protein, M protein, nsps, open-reading frames (ORFs) and accessory factors, may have the potential to serve as antigens.

It is key to identify an antigen which can bind with the T or B cell receptors which are recognised and can induce an immune response. Vaxijen, a publicly available server for prediction of protective antigens, tumour antigens and subunit vaccines [28], is based on properties of principal amino acids, which are generated from protein sequences by autocross covariance transformation.

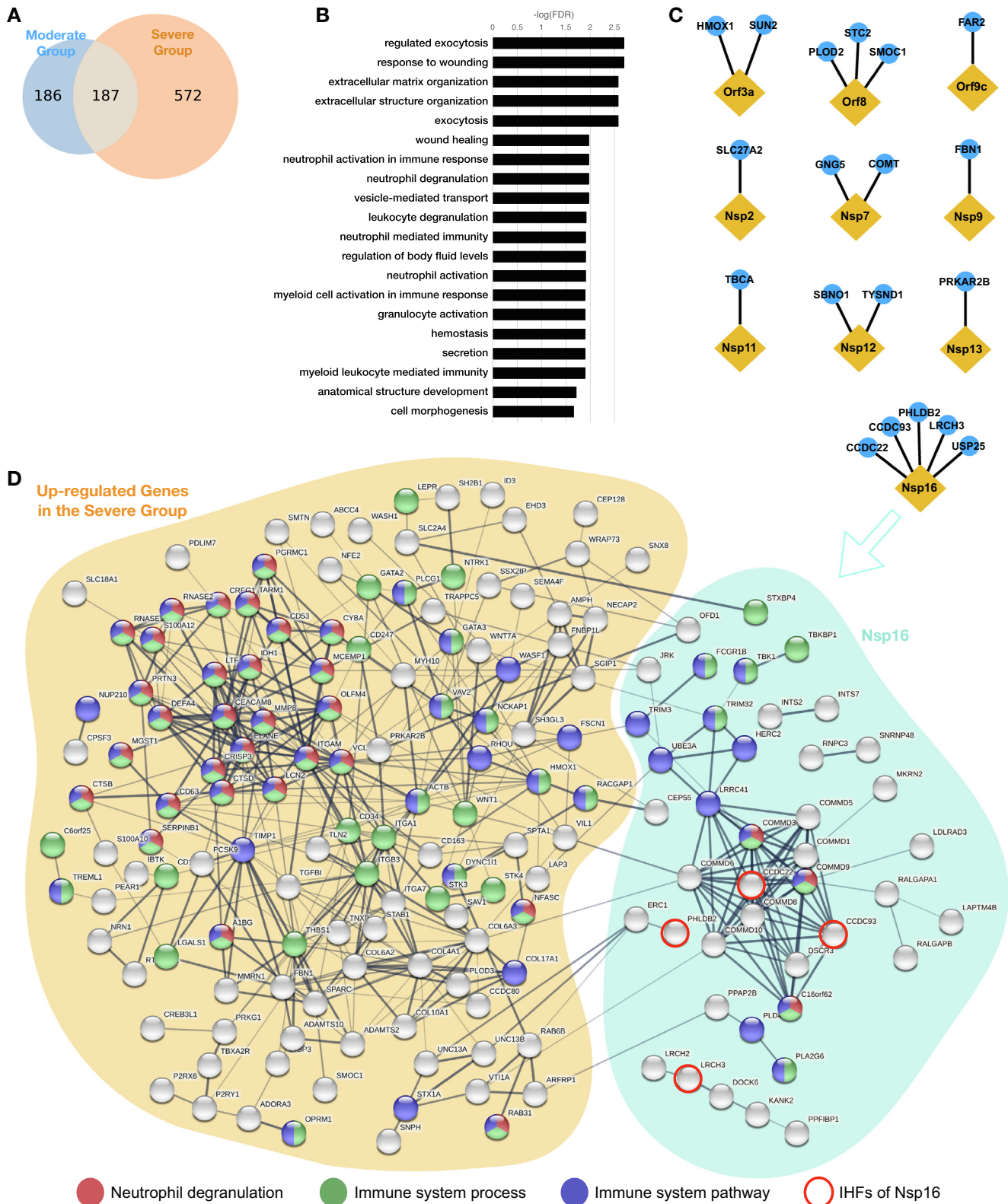
Vaxijen predicts protective antigens without relying on sequences of previously known antigens and has been used for this purpose by several papers investigating SARS-CoV-2 [22,25,29]. Vaxign-ML [30] is an analysis system based on reverse vaccinology which predicts protective antigens for vaccine targets using the machine learning (ML) algorithm extreme gradient boosting, trained on data from Protogen, a database of protective antigens [31].

#### 2.2.1. Antigen prediction based on cellular immune response

In searching for antigens beyond the S protein, alternative approaches not solely based on the genomic sequence of the virus could be considered. One such approach could be to search for antigens related to cellular immune response. Cellular immune response is activated for protection against pathogens which enter and inhabit the host cells during infection, as viruses including SARS-CoV-2 do. Although antigen properties which are important for this type of immunity are not yet well characterized, it is an area worth exploring for COVID-19 vaccine development [32]. Gene expression (transcriptomic) data from host and virus and virus-host interaction data could be combined to build a virus infection network to predict antigens based on their relation to host cellular immune response. SARS-CoV-2 proteins and their interactions with host factors were associated with imbalanced host immune responses, such as elevated proinflammatory cytokine levels [33]. Elevated proinflammatory cytokine levels can lead to progression of severe COVID-19 [34].

To understand which SARS-CoV-2 proteins are associated with imbalanced host immune responses, computational biology can be used to identify genes coding proteins associated with COVID-19 severity. Differential gene expression analysis using data from COVID-19 patient and non-COVID-19 infected controls has shown that neutrophil activation is significantly associated with COVID-19 status and severity [35]. We investigated this further by performing a *meta*-analysis on four transcriptomics datasets from COVID-19 severity studies (Table 1) and created a protein interactome network to test the hypothesis that other computational approaches may be useful in identifying alternative antigens (Fig. 1). We used EBSeq [37], an R package for gene and isoform differential expression analysis of RNA-seq data, to identify differentially expressed genes in patients with 1) moderate and 2) severe COVID-19 as defined by the World Health Organisation (WHO) standards [36]. By performing gene ontology (GO) enrichment analysis using g:Profiler, a webserver developed for this purpose [37], we identified significantly enriched biological functions of the differentially expressed genes, such as exocytosis, wound healing, and neutrophil degranulation, which are related to viral replication [38] and adaptive immune system [39].

To investigate which, if any, SARS-CoV-2 proteins are associated with an imbalanced immune response, protein interactome networks [40] for each SARS-CoV-2 protein can be created to capture protein interactions between each SARS-CoV-2 protein-host protein interactions (SHPs) and the up-regulated differentially expressed genes in the severe COVID-19 group. By using this technique, we found that the integration host factors of nsp16 are closely connected to the up-regulated differentially expressed genes



**Fig. 1.** Antigen selection from transcriptomics analysis and protein interactome network analysis. (A) Venn diagram to show the numbers of DEGs in the COVID-19 moderate group and the severe group. (B) Bar graph to show the top 20 highly enriched functions of the up-regulated DEGs in the severe group. (C) SARS-CoV-2 proteins and their SHPs that are differentially expressed genes in the severe group. (D) Protein interactome network between the up-regulated DEGs and SHPs of nsp16.

that play a key role in neutrophil degranulation and immune system (Fig. 1C and D). Direct interactions and the proteins of differentially expressed genes can be visualised using STRING [42], an online tool to create protein–protein interaction networks. Nsp16

has been suggested as a target for SARS-CoV and MERS-CoV vaccine because it encodes ribose 2'-O-methyltransferase and this methylation helps coronavirus avoid the activation of type I interferon-dependent innate immune response by viral RNA [41].



Taken together, we suggest this methodology shows promise in the identification of antigens based on cellular immune response.

### 3. Epitope prediction

#### 3.1. B cell epitope prediction tools

Upon infection, B cell immune response is initiated by recognising free-floating antigens through B cell receptors (BCR), a membrane-bound immunoglobulin. Such an event activates B cells to differentiate into memory B cells or antibodies-secreting plasma cells in mediating adaptive humoral immunity. Antibodies from plasma cells can function in neutralizing toxins, opsonising pathogens, and further activating CD4+ T cells by acting as an antigen-presenting cell (APC) [42]. Memory B cells can retain antigen-specific immune memory and different immunoglobulin memory B cells distinct functions [43]. Amongst them, the IgG+ sub-class is most heavily studied in vaccinology as they preferentially differentiate into plasma cells [43].

B cell epitopes are parts of an antigen, which may contain a solvent-exposed region and range from 5 to 20 amino acids (AA) in length. They are further divided into continuous (linear) or discontinuous (conformational) epitopes. Whilst continuous epitopes are made up of consecutive solvent-exposed residues, discontinuous epitopes are formed by solvent-exposed residues that may not be sequential. Consequently, continuous epitopes, but not discontinuous epitopes, can be recognized by B cells even when the antigen is denatured. Although an estimated 90% of B-cell epitopes are discontinuous, more tools are available and utilised in predicting continuous epitopes than discontinuous epitopes. This is also reflected in the case of designing the COVID-19 vaccine.

Traditional methods in mapping B cell epitopes vary drastically, including 3D structure study of the antigen–antibody complexes [44], antibody-binding peptide library screening [45], and functional assays with mutagenesis approach [46]. Although these epitope selection methods are more accurate than *in silico* prediction methods [47], epitope mapping experiments are restricted by their costs and feasibility. *In silico* epitope prediction is faster and can greatly improve the success rate by filtering promising candidates.

##### 3.1.1. Continuous B cell epitope prediction in SARS-CoV-2

Earlier established prediction tools for continuous B cell epitopes were based on smaller datasets and their AA propensity scales that characterise the epitopes physicochemical features. For example, Parker and Hodges have developed a hydrophilicity parameter using high-performance liquid chromatography data with smoothing values in a seven-residue window [48] which has been used singly in the SARS-CoV-2 epitope design study [49]. Other parameters including surface accessibility by Emini, flexibility prediction on the mobility of protein segments by Karplus and Schulz [50] and  $\beta$ -turn prediction in the Chou-Fasman method [51]. Rahman *et al.* [52] also utilised additional structural prediction models including the RaptorX Property server for 2° structures and the Phyre2 server for 3° structures [53]. The relatively poor performance of AA propensity scales prompted the developments of ML-based methods [54,55]. These ML-based tools were trained on feature vectors capturing the selected properties of experimentally validated B-cell and non-B-cell epitopes [56–61]. Specifically, some tools trained the data with Support-Vector Machine classifiers (SVM), whilst others preferred the uses of Random Forests algorithms (RF) or artificial neural networks (ANN). Their accuracies and applications to SARS-CoV-2 have been summarised in Table 2.

Multi-parametric algorithms on AA propensity scales have also been developed, such as PREDITOP (based on 4 propensity scales),

PEOPLE (based on 5 propensity scales), and BEEPro (based on 19 propensity scales) [55,62,63]. Despite the indication of increased accuracy, these tools are not as popular as the easily accessible IEDB server that acts as a single platform providing the methods by Emini (surface accessibility), Parker (hydrophilicity), Kolaskar and Tongaonkar's (AA frequencies), Karplus and Schulz (flexibility), and Chou and Fasman ( $\beta$ -turn propensity) as a multi-parametric filter, as seen in multi studies [64–67]. Some SARS-CoV-2 epitope design studies also integrated both the ML-based approaches and the propensity scales in predicting the epitope from their targets [68–73]. Similarly, some studies filtered epitopes that were mutually found in two or more ML-based tools [74–76]. It is unclear whether a specific combination is superior to others, further investigation maybe needed to provide more guidance on which tool should be integrated or not.

##### 3.1.2. Discontinuous B cell epitope prediction in SARS-CoV-2

In contrast, discontinuous B cell epitope prediction is more challenging than continuous B cell epitope prediction mainly due to two reasons: 1) limited 3D-structural information of antigen–antibody complex; 2) difficulties in isolating the discontinuous epitope for selective antibody production. Early computational tools rely on the identification of specific AA regions such as CEP (residue exposed to solvent) [77], DiscoTope1.2 (solvent accessibility and spatial information of AA) [78], and ElliPro (protruding region on antigen surfaces) [79]. In the latest version of DiscoTope2.0, a DiscoTope method was incorporated with spatial neighbourhood definition and half-sphere exposure and has achieved significant improvement in prediction accuracy (AUC from 0.791 to 0.824) [80]. Another approach classed as mimotope-based methods, identify conformational B cell epitope in proteins with known 3D structures. These include MIMOX [81], PEPITOPE [82], EpiSearch [83], MimoPro [84], and PEPMAPPER [85]. However, additional information including the 3D structure of the selected antigen and the antibody-affinity of the peptide is required for optimal prediction performances.

Advances in computational power also led to development of ML-based prediction tools for discontinuous B cell epitopes. Specifically, EPSVR used the Support Vector Regression (SVR) method to integrate six epitope physicochemical properties and has been applied [86]. In the same publication, the authors also combined EPSVR with DiscTope2.0 [78], PEPITO [87], SEPPA [88], Epitopia [89], and EPCES [90] to construct EPMeta, which was reported to have a significantly higher performance than PEPITO and DiscTope [86]. In contrast, BepiPRED-2.0, another version of the above-mentioned tools, has used a RF algorithm trained on the dataset of the antibody-antigen protein structures of conformational epitopes [60]. Although its reported accuracy for discontinuous B cell epitopes was fair, it is more commonly applied on linear epitope predictions in SARS-CoV-2 vaccine studies. Instead, most studies utilised DiscoTope2.0 or ElliPro in their discontinuous epitope predictions (see Table 2).

In addition to continuous/discontinuous B cell epitope prediction tools, an alternative tool, IgPred, enables the prediction of which immunoglobulin subclass the B cell epitope is capable of [91]. The tool was trained on over 14,000 epitopes using the SVM method and can be used to identify epitopes with preference for inducing IgG and IgA antibodies [92].

#### 3.2. T cell epitope prediction tools

Beside humoral immune response, circulating T follicular helper ( $T_{FH}$ ) cells and other T lymphocytes are important in inducing antibody-producing plasma cells and long-lived memory B cells [93]. A recent report also indicated that antibody level in COVID-19 patients is correlated with the level of circulating  $T_{FH}$  cells

**Table 2**  
Common B-cell epitope prediction tools.

Tool	Method	Accuracy	Utility in SARS-CoV-2 vaccine design publications
<b>Continuous B-cell epitope prediction tools</b>			
<b>Parker</b> (Accessible on IEDB)	Single AA propensity scale AA sequence hydrophilicity with smoothing values in a seven-residue window	61.21 AUC with epitopes of 16 AA length [Liu et al., 2020] [252]	[Rakib et al., 2020] [49]: applied to filter B-cell epitopes predicted by BepiPred, Emini, and Klaskar and Tongaonkar methods
<b>Kolaskar and Tongaonkar</b> (Accessible on IEDB)	Single AA propensity scale Based on physicochemical properties and frequencies of AA in experimentally determined B-cell epitopes	Self-reported optimal accuracy of 75% 55.76 AUC with epitopes length of 16 AA [Liu et al., 2020] [252]	[Rakib et al., 2020] [49]: identified 15 epitopes in N protein [Rahman et al., 2020] [52]: Identified 5 epitopes in the RBD regions of S, M, and E proteins [Abdulhameed Odhar et al., 2020] [253]: Identified 5B cell epitopes on chain A of S protein [Marino Gammazza et al., 2020] [254]: identified B cell epitope region in Replicase Polyprotein 1ab
<b>Emini</b> (Accessible on IEDB)	Single AA propensity scale Surface accessibility	60.76 AUC with epitopes length of 16 AA [Liu et al., 2020] [252]	[Rakib et al., 2020] [49]: Identified 9 epitopes in N protein [Abdulhameed Odhar et al., 2020] [253]: identified 3 epitopes on chain A of S protein [Yazdani et al., 2020] [255]: identified B-cell epitopy of predicted CTL epitopes in S, M, N, and E proteins [Alam et al., 2020] [256]: identified 6B-cell epitopes in S protein and transmembrane protein 199
<b>Karplus and Schulz</b>	Single AA propensity scale Flexibility Prediction on the mobility of protein segments.	Unavailable	[Rakib et al., 2020] [49]: applied to filter B-cell epitopes predicted by BepiPred, Emini, and Klaskar and Tongaonkar methods [Alam et al., 2020] [256]: identified 6B-cell epitopes in S protein and transmembrane protein 199
<b>Chou and Fasman</b>	Single AA propensity scale $\beta$ -turn prediction	52.70 AUC with epitopes length of 16 AA [Liu et al., 2020] [252]	[Rakib et al., 2020] [49]: applied to filter B-cell epitopes predicted by BepiPred, Emini, and Klaskar and Tongaonkar methods [Alam et al., 2020] [256]: identified 6B-cell epitopes in S protein and transmembrane protein 199
<b>PREDITOP</b>	Multiple AA propensity scales Based on AA hydrophilicity, accessibility, flexibility, and secondary structure property	Self-reported optimal accuracy of 60%	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>BeePro</b>	Multiple AA propensity scales Based on the antigenicity, hydrophilicity, hydrophobicity, accessible surface area, flexibility, interactivity, buriability, composition, polarity, volume, charge transfer and donor capability, hydrogen-bond donor capability, and secondary structure	Self-reported optimal accuracy of 99.29%	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>AAP</b>	ML-based SVM-based model trained on AA pair	Self-reported optimal accuracy 71.09%	[Tohidinia and Sefid 2020] [205]: initial filtering on S protein epitopes
<b>ABCpred</b>	ML-based ANN-based model trained on AA patterns	Self-reported optimal accuracy of 66.41%	[Behmard et al 2020] [131]: Identified nine 16-mer epitopes in S, E, M, and N proteins [Dai et al 2020] [76]: Identified 11 16-mer epitopes in N proteins [Rahman et al., 2020] [52]: Identified 29 peptides in S, M, and E proteins [Lon et al., 2020] [75]: 6 epitopes predicted in S, M, and E proteins [He et al., 2020] [132]: Predicting linear B-cell epitopy of 16 sequences of S protein [Vashi et al., 2020] [257]: 24 epitopes predicted with various AA length from S protein
<b>APCpred</b>	ML-based SVM-based model trained on anchoring pair composition	Self-reported optimal accuracy of 72.94%	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>BepiPred (2.0)</b> (Accessible on IEDB)	ML-based RF-based model trained on epitopes annotated from antibody-antigen protein structures	Self-reported optimal accuracy of 0.62 AUC	[Rahman et al., 2020] [52]: Predicted 14 epitopes of S, E, and M proteins [Rakib et al., 2020] [49]: 11 epitopes identified with various AA length in N protein [Ayyagari et al., 2020] [112]: 3 epitopes identified in M protein [He et al., 2020] [132]: 25 epitopes identified on S protein [Khairkhan et al., 2020] [92]: 8 epitopes identified on S and N proteins

Table 2 (continued)

Tool	Method	Accuracy	Utility in SARS-CoV-2 vaccine design publications
<b>LBEEP</b>	ML-based SVM and AdaBoost-RF based model trained on Dipeptide Deviation from Expected Mean (DDE) of epitopes	Self-reported optimal accuracy of 73%	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>Bayesb</b>	ML-based Bayes Feature Extraction	Self-reported optimal accuracy 68.50%	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>BCPRED</b>	ML-based SVM-based model with various string kernels that select representing sequence into length-fixed feature vectors	Self-reported optimal accuracy of 0.758 AUC	[Singh et al., 2020] [64]: Identified 12B-cell epitopes in S protein [Singh et al., 2020] [64]: identified 2 linear epitopes in the structural glycoproteins [Rahman et al., 2020] [52]: identified 4 epitopes shared with B-cells and T-cells [Chauhan et al., 2020] [258]: 3B cell epitopes identified in SARS-CoV-2 genome-associated proteins
<b>iBCE-EL</b>	ML-based Combination Of 6 ML algorithms trained on epitope sequences	Self-reported optimal accuracy of 72.9%	[Dar et al., 2020] [238]: Identification of linear B cell epitopes in selected HLA II class epitopes in S protein [Samad et al., 2020] [133]: identified 4B cell epitopes in S protein [Ahammad and Lira, 2020] [134]: identified 117 epitopes in S protein
<b>SVMTriP</b>	ML-based SVM-based model trained on length-fixed tripeptide composition vectors	Self-reported optimal accuracy of 0.702 AUC	[Banerjee et al., 2020] [259]: identification of B-cell epitopes overlapping with CTL and HTL epitopes in N, M, E, Orf6, Orf7a, and Orf10 proteins
<b>COBEpro</b>	ML-based SVM-based model predicting short peptide fragments followed by epitopic propensity prediction	Self-reported optimal accuracy of 0.829 AUC	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>EPMLR</b>	ML-based Multiple linear regression-based model trained on BEOracle dataset	Self-reported optimal accuracy of 0.728 AUC	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>LBtope</b>	ML-based Several models including SVM-based and nearest neighbor based trained on the modified AAP profile of epitopes	Self-reported optimal accuracy of 86%	[Anand et al., 2020] [260]: identified 7B cell epitopes in S, M, N, and Orf3a proteins
<b>Discontinuous B-cell epitope prediction tools</b>			
<b>CEP</b>	Algorithms using residue accessibility and spatial distance cut-off	Self-reported optimal accuracy of 75%	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>DiscoTope (2.0)</b> (Accessible on IEDB)	DiscoTope method incorporated with spatial neighbourhood and half-sphere exposure information	Self-reported optimal accuracy of 0.824 AUC	[Forni et al., 2020] [261]: identification of immunogenicity of antigenic variation in S and N proteins [Arshad Dar et al., 2020] [238]: identification of 77 epitopes on S protein [Ayyagari et al., 2020] [112]: identified 25B-epitopes on M protein [Grifoni et al., 2020] [109]: Predicted discontinuous epitope on 959 continuous epitopes on surface glycoprotein
<b>ElliPro</b> (Accessible on IEDB)	Thornon's method integrated with residue clustering algorithm, MODELLER program in predicting and visualizing the epitope structures.	Self-reported optimal accuracy of 91%	[Rajesh Anand et al., 2020] [260]: 10 conformational epitopes identified on S, Orf3a, M, and N proteins [Dar et al., 2020] [238]: identification of 148 epitopy residues on S protein [Singh et al., 2020] [64]: Scoring discontinuous epitope probability in 12 selected S proteins [Sarkar et al., 2020] [234]: Identified 4 potential regions in S, N, M, and E proteins [Wagas et al., 2020] [65]: identified 5 conformational epitopes in S, N, M, and E proteins
<b>PEPITO</b>	Incorporating AA propensity scale with side chain orientation and solvent accessibility information of epitope/non-epitope	Self-reported optimal AUC of 75.38	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>SEPPA</b>	Scoring based on comparison to local spatial information of surface residues of 82 antigen-antibody protein complexes	Self-reported optimal accuracy of 0.742 AUC	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>PepMapper</b>	Mimotope-based An ensemble of MimoPro and Pep-3D-search for epitope mapping	Self-reported optimal sensitivity of 1.00, specificity of 0.839, and precision of 0.256	Not utilised/cited in any SARS-CoV-2 vaccine design publications

(continued on next page)

Table 2 (continued)

Tool	Method	Accuracy	Utility in SARS-CoV-2 vaccine design publications
<b>Epitopia</b>	ML-based Naïve Bayes classifier trained on 3D antigen–antibody structures and epitope sequences	Self-reported optimal accuracy of 89.4%	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>EPSVR</b>	ML-based SVR-based model trained on residue epitope propensity, conservation score, side chain energy score, contact number, surface planarity score, and secondary structure composition	Self-reported optimal accuracy of 0.597 AUC	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>CBTOPE</b>	ML-based SVM-based model trained on the physicochemical and sequence featured of conformational B-cell epitopes	Self-reported optimal accuracy of 86.6%	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>BepiPred (2.0)</b> (Accessible on IEDB)	ML-based RF-based model trained on conformational epitopes annotated from antibody-antigen protein structures	Self-reported optimal accuracy of 0.62 AUC on conformational epitopes	Not utilised/cited in any SARS-CoV-2 vaccine design publications for discontinuous B-cell epitope prediction.

This table shows some of the common B-cell epitope prediction tools and their reported accuracy or evaluated accuracy, if available. Their utilities in SARS-CoV-2 vaccine design papers are summarised briefly by the result they have indicated and their application on which protein on SARS-CoV-2, if mentioned. As seen, some tools with relatively high accuracies are yet to be utilised for SARS-CoV-2 vaccine designs.

[94]. Another study also reported a population of SARS-CoV-2-specific T cells with a ‘stem-like memory phenotype’ is present in the antibody-seronegative asymptomatic and mild COVID-19 patients [95]. Many more reports have highlighted the critical role played by T cells in inducing effective SARS-CoV-2-specific immune response [96–98].

T cells can be divided into 2 major types: CD4+ T cell (helper T cell) and CD8+ T cell (cytotoxic T cell). Their activations rely on 3 signals: 1) Binding between the antigen present on the major histocompatibility complex (MHC) of the antigen-presenting cells (APCs) and the T cell receptor (TCR) on the T cell; 2) Binding of CD80 or CD86 on APC by CD4+ T cells/binding of CD70 or CD137 on APC by CD8+ T cells; 3) Released cytokines in determining the cell fate of CD4+ T cells (i.e., IL-12 for  $T_{H1}$  type, IL-4 for  $T_{H2}$  type, and IL-6 and IL-23 for  $T_{H17}$ ) [99]. Specifically, MHC I is recognised by CD8+ T cells and MHC II by CD4+ T cells. Hence, epitope mapping for T cell can be generally divided into 3 steps: 1) Antigen processing by APC; 2) Peptide-binding to MHC molecules; 3) Recognition by TCR. Here, we have listed some T cell epitope prediction tools that can potentially accelerate epitope selection for MHC multimer [100], lymphoproliferation [101], and ELISPOT assays [102].

### 3.2.1. CD4+ T cell epitope prediction in SARS-CoV-2

CD4+ T lymphocytes and their associated cells recognise exogenous antigens. When APCs encounter, exogenous antigens are taken up by the APCs through phagocytosis into the endo/lysosomal pathway. The antigens are then broken down by enzymes such as cathepsins and proteases, and present on the class II MHC molecule [103]. Current established CD4+ T cell epitope prediction methods focus on the MHC II binding prediction. MHC II binding peptides vary in length between 9 and 22 residues but only the peptide-binding core with 9 AA in length sits in the binding groove [104]. Compared to MHC I molecule, MHC II binding groove is open, allowing ends of the peptide to extend beyond the groove, and the pocket is shallower than MHC I. Such characteristics constitute the cause of a lower accuracy in peptide-MHC II binding prediction methods [105].

Table 3 shows a list of peptide-MHC II binding prediction tools, their accuracy, and their applications in SARS-CoV-2 vaccine design studies. Notably, NetMHCIIpan 4.0, available on IEDB, integrated large-scale MHC II ligand mass spectrometry (MS) data for training using the NNAlign\_MA method [106,107]. Some SARS-CoV-2 vaccine design studies utilised MixMHC2Pred alternatively. MixMHC2Pred is a predictor trained on the MS data with a motif

deconvolution algorithm of over 99,000 MHC-II peptides and has been shown to outperform the earlier version of NetMHCIIpan [108]. Another notable example is TepiTool, a pipeline for predicting MHC I and II epitopes using a combined number of tools [109] and have been used in several SARS-CoV-2 studies [29,76,80,110]

### 3.2.2. CD8+ T cell epitope prediction in SARS-CoV-2

CD8+ T lymphocytes are responsible for processing endogenous antigens. Endogenous antigens are cleaved by the proteasome into peptides and then transported by transporter-associated protein (TAP) to the endoplasmic reticulum for the association to MHC I molecules [103]. MHC I epitopes are usually shorter than MHC II epitopes, ranging from 9 to 11 AA in length, and sit in the MHC I groove delineated by two  $\alpha$ -helices [104]. Although most predictions only focused on the intermediate step of MHC I binding affinity, tools are also available in predicting the proteasomal cleavage site, TAP binding affinity, and TCR binding affinity, as mentioned above. Additionally, a platform such as NetCTLpan [111] provides end-to-end cytotoxic T cell (CTL) epitope predictions and has been utilised in several SARS-CoV-2 vaccine studies [74,112,113].

Endogenous antigens are known to have proteasome cleavage sites at their C-terminal for cleavage by immunoproteasomes [114]. The cleavage specificities of the immunoproteasome are determined by the residues located at the cleavage sites and neighbouring positions [115]. Recently, ML classifiers have been trained on cleavage and non-cleavage sites such as Pcleavage [116], PCPS [117] and NetChop [118]. Following proteasomal cleavage, TAP binding is also essential in ensuring transport to MHC I binding. TAP transporters prefer to translocate peptides of 8–11 AA in length, similar to the MHC I epitope lengths [119]. They also have higher affinity to hydrophobic residues at specific positions [120]. Based on the TAP motifs, consensus matrix, and their subsequent application to ML-techniques, several methods have been developed. A summary of all endogenous antigen processing prediction tools is beyond the scope of this review. Instead, we present, in Table 3, the accuracy of the listed tools and some examples of their application on SARS-CoV-2 vaccine design.

Notably, some tools listed for MHC II binder predictions can also be used for MHC I prediction due to the data used for their model training. For example, RANKPEP [121,122] provides position-specific scoring matrix for both MHC I and II binding peptides and MHCnuggets [123] is a neural network-based model trained on common and rare alleles trained on MHC binders. Some tools are specific for MHC I peptide predictions, e.g., NetMHCpan-4.1



**Table 3**  
Common T-cell epitope prediction tools.

Tool	Method	Accuracy	Utility in SARS-CoV-2 vaccine design publications
<b>MHC I/II binding prediction tools</b>			
<b>ARB</b> (Accessible on IEDB)	Average relating binding (ARB) matrix-based prediction of binding propensity to MHC II	Self-reported optimal accuracy of 83%	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>RANKPEP</b> (MHC I: <b>RANKPEP</b> )	Position specific scoring matrix-based predicting of MHC binding peptides	Self-reported optimal accuracy of 0.96 AUC for MHC II and 80% for MHC I	[He et al., 2020] [132]: used to predict MHC II epitopy in 28 epitopes from S protein [Yazdani et al., 2020] [255]: identified MHC-II immunogenicity of epitopes selected from S, E, N, and M proteins
<b>Tepitope</b> (Accessible on IEDB as Sturniolo)	A matrix-based method utilising pocket profiles of MHC binding peptides	Self-reported optimal sensitivity of 80%	[Dong et al., 2020] [69]: predicted MHC II epitopes in combination of NN-align, SMM-align, and MetMHCIIpan method of various SARS-CoV-2 proteins [Martin and Cheng, 2020] [262]: predicted MHC II epitopes in combination of NN-align, SMM-align, ComLib and MetMHCIIpan method of glycan protein on S protein [Srivastava et al., 2020] [236]: predicted MHC I epitopes of 11 ORF, structural and non-structural proteins in combination with NN-align, consensus, SMM-align, combinatorial library, and NetMHCIIpan
<b>MHCAttnNet</b>	A deep neural model trained on amino acid sequences of MHC epitopes	Self-reported optimal accuracy of 94.18%	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>MHCnuggets</b>	A long short-term memory network trained on common or rare alleles of MHC epitopes	Self-reported optimal accuracy of 0.924 AUC	[Campbell et al., 2020] [126]: Predicted MHC-I epitopes with 5 other tools for HLA supertype covering 90% of the population in 57 countries
<b>Combinatorial library</b> (Accessible on IEDB)	Heuristic method translating positional scanning combinatorial library data of MHC I and II epitope data into main and secondary anchor positions with preferred residues for MHC binding peptide predictions	Self-reported optimal accuracy of 0.935 AUC	[Martin and Cheng, 2020] [262]: predicted MHC II epitopes in combination of NN-align, SMM-align, Tepitope and MetMHCIIpan method of glycan protein on S protein [Srivastava et al., 2020] [236]: predicted MHC I epitopes of 11 ORF, structural and non-structural proteins in combination with NN-align, consensus, SMM-align, Tepitope, and NetMHCIIpan [Yazdani et al., 2020] [255]: identified MHC-II immunogenicity of epitopes selected from S, E, N, and M proteins with ANN and SMM, applied as consensus method [Dong et al., 2020] [69]: Used as the 'Combinatorial library' with other IEDB accessible methods for predicting MHC epitopes in all proteins
<b>MHC II binding prediction tools for CD4+ T-cells epitopes</b>			
<b>NetMHCIIpan 4.0</b> (available on IEDB)	Indirect method NNAIalign_MA method trained on MHC II ligand MS data	Self-reported optimal accuracy of 0.952 PPV	[Liu et al., 2020] [128]: predicted MHC class II peptides in SARS-CoV-2 samples of different ethnic backgrounds [Ayyagari et al., 2020] [112]: Predicted 27 MHC II alleles in M protein [Khairkhah et al., 2020] [92]: Predicted 20 MHC II epitopes in S, N, and M proteins [Requena et al., 2020] [263]: Predicted 30 MHC II alleles with consideration of HLA frequencies by country [Liu et al., 2020] [264]: Filtered 22 potential MHC II epitope candidates in S, M, and N proteins [Dar et al., 2020] [238]: pan-MHC II binding prediction of surface glycoprotein
<b>EpiDock</b>	A peptide library of X-ray structure of HLA-DP2 proteins with prediction performed using AutoDock [Morris et al., 2009] and Rosetta Dock [Lyskov and Gray 2008]	Self-reported optimal accuracy of 0.900 AUC	[Can et al., 2020] [233]: used for filtering MHC II epitopes with low IC50 docking to specific MHC II allele
<b>EpiTOP</b>	Quantitative matrix-based approach in predicting MHC II binding based on proteochemometrics	Self-reported optimal accuracy of 93%	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>HLA-DR4Pred</b>	SVM-based model in identifying HLA-DRB1*0401 binding peptides	Self-reported optimal accuracy of 86%	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>PREDIVAC</b>	Specificity-determining residue (SDR) based MHC II binding prediction integrated for 95% of MHC II allelic variants	Self-reported optimal accuracy of 0.872 AUC	[Dar et al., 2020] [238]: used to predict MHC II supertype allele interaction following prediction by NetMHCIIpan
<b>SMM-align</b> (Accessible on IEDB)	A stabilized matrix alignment method based on MHC II epitope amino acids preferences by Gibbs sampler and predicting using SVRMHC predictions	Self-reported optimal accuracy of 0.756 AUC	[Dong et al., 2020] [69]: predicted MHC II epitopes in combination of NN-align, Tepitope, and MetMHCIIpan method of various SARS-CoV-2 proteins [Martin and Cheng, 2020] [262]: predicted MHC II epitopes in combination of NN-align, Tepitope, ComLib and MetMHCIIpan method of glycan protein on S protein

(continued on next page)

Table 3 (continued)

Tool	Method	Accuracy	Utility in SARS-CoV-2 vaccine design publications
<b>NN-align</b> (Accessible on IEDB)	ANN-based trained on MHC II binding epitopes with binding core and affinity results available upon query	Self-reported average accuracy of 0.855 AUC	[Srivastava et al., 2020] [236]: predicted MHC I epitopes of 11 ORF, structural and non-structural proteins in combination with NN-align, consensus, Tepitope, combinatorial library, and NetMHCIpan [Campbell et al., 2020] [126]: MHC binder identification for epitope prediction targeting 90% populations of 57 countries [Behrman et al., 2020] [131]: utilised to identify MHC II binding epitopes in S and N proteins [Savafi et al., 2020] [74]: Predicted MHC II binding epitopes with SMM-align, MHCpred and NetMHCIpan 3.2 methods [Rehman et al., 2020] [67]: Used for predicting T helper cell epitopes in spike protein [Yazdani et al., 2020] [255]: identified MHC-II immunogenicity of epitopes selected from S, E, N, and M proteins [Singh et al., 2020] [64]: used to predict 15-mer T-helper cell epitope from S protein [Mukherjee et al., 2020] [110]: used to filter MHC II epitopes with lower than specified IC50 threshold
<b>Consensus</b> (Accessible on IEDB)	Combination of NN-align, SMM-align, and the combinatorial peptide scanning library methods on IEDB	Self-reported average accuracy of 0.89 ± 0.05 AUC	
<b>CD8+ T-cells epitopes Prediction Tools</b>			
<b>vPcleavage</b>	Proteasomal cleavage site prediction tool. A SVM-based classifier trained on amino acids of 7AA in length with cleavage site represented in sequence window	Self-reported optimal accuracy of 0.805 AUC	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>NetChop3.1</b> (Accessible on IEDB)	Proteasomal cleavage site prediction tool. Neural network-based trained on sequence encoded data	Self-reported optimal accuracy of 0.85 AUC	[Shomuradova et al., 2020] [265]: Predicted proteasomal cleavage score of C-terminal AA in S protein peptides [Safavi et al., 2020] [74]: Predicted proteasomal processing in T cell epitope together with MAPP
<b>PCPS</b>	Proteasomal cleavage site prediction tool. Provides three models for predicting immunoproteasome and proteasome cleavage site each. Developed through training on residue fragments	Reported sensitivity of 0.88 and specificity of 0.57 in Gomez-Perosanz et al., 2019[117]	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>TAPPred</b>	Proteasomal fragment-TAP binding prediction tool A cascade SVM-based method trained on quantitative matrix of TAP binding regions	Self-reported optimal accuracy of 0.81 Pearson's correlation coefficient	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>TAPHunter</b>	Proteasomal fragment-TAP binding prediction tool A SVM-based model trained on TAP peptide fragments and composition effects	Self-reported optimal accuracy of 88%	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>TAPreg</b>	Proteasomal fragment-TAP binding prediction tool SVM-based method trained on single residue positions and residue combinations	Self-reported optimal accuracy of 0.89 ± 0.03 Pearson's correlation coefficient	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>Peters et al., 2003</b> (Accessible on IEDB as part of MHC-I Processing Predictions)	Stabilized matrix method (SMM) used to score sequence position in distinguishing TAP-binding epitopes	Self-reported optimal accuracy of 0.89 AUC	[Srivastava et al., 2020] [236]: Used to predict Cytotoxic T cell epitopes in the 'MHC-I Processing Predictions' pipeline [Anand et al., 2020] [260]: Predicted 11 antigenic T cell epitopes from structural proteins [Banerjee et al., 2020] [259]: Predicted TAP binding affinity as part of the IEDB server
<b>MHCflurry2.0</b>	MHC I binding prediction tool A neural network-based model trained on MHC I allele with affinity and MS data	Self-reported optimal accuracy of 0.94 AUC	[Campbell et al., 2020] [126]: Predicted MHC-I epitopes with 5 other tools for HLA supertype covering 90% of the population in 57 countries
<b>NetCTLpan1.1</b>	MHC I binding prediction tool Integrated MHC peptide binding with NetMHCpan, with proteasome cleavage predicted with NetChop and TAP transport prediction by Peters et al., 2003	Self-reported optimal accuracy of 0.976 AUC	[Ayyagari et al., 2020] [112]: Identification MHC I epitopes on M proteins [Safavi et al., 2020] [74]: predicted MHC I epitopes on 6 non-structural proteins [Mishra et al., 2020] [266]: predicted MHC I epitopes with PickPocket 1.1 from structural and non-structural SARS-CoV-2 proteins
<b>NetMHCpan 4.1</b> (available on IEDB)	MHC I binding prediction tool NNAlign_MA method trained on MHC I ligand MS data	Self-reported optimal accuracy of 0.8291 PPV	[Campbell et al., 2020] [126]: Predicted MHC-I epitopes with 5 other tools for HLA supertype covering 90% of the population in 57 countries [Liu et al., 2020] [128]: MHC-I epitope prediction for augmentation to increase vaccine population coverage [Dearlove et al., 2020] [267]: 9-mer MHC-I epitope prediction in S protein [Liu et al., 2020][264]: Predicted MHC Class I epitope with MHCflurry in ensemble.
<b>DeepLigand</b>	MHC I binding prediction tool A semi-supervised model trained on sequence features correspond to secondary factors of MHC I binding epitopes	Self-reported optimal accuracy of 0.979 AUC	Not utilised/cited in any SARS-CoV-2 vaccine design publications

Table 3 (continued)

Tool	Method	Accuracy	Utility in SARS-CoV-2 vaccine design publications
<b>PickPocket1.1</b>	MHC I binding prediction tool A library of pocket-residues and corresponding binding specificities with consensus to NetMHCpan method	Self-reported optimal accuracy of 0.92 AUC	[Mishra, 2020] [113]: Predicted 921 cytotoxic T cell lymphocyte epitopes in 10 SARS-CoV-2 proteins with NetCTLpan1.1 [Yazdani et al., 2020] [255]: Predicted MHC-I epitopes on IEDB server together with other listed tools [Campbell et al., 2020] [126]: Predicted MHC-I epitopes with 5 other tools for HLA supertype covering 90% of the population in 57 countries
<b>HLA-CNN</b>	A convolutional neural network-based trained on HLA-Vec, an AA distributed representation of epitopes	Self-reported average accuracy of 0.836 AUC	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>HLA presentation coverage prediction</b>			
<b>EvalVax</b> (Unlinked/Robust)	A combinatorial optimization from OptiVax tools considering HLA allele frequencies Unlinked: assume independence between HLA loci Robust: considering linkage disequilibrium of HLA genes between loci with halotype frequencies for population coverage estimates	Unavailable	[Liu et al., 2020] [128]: Used for ensuring high population coverage in T cell epitopes
<b>ShinyNap</b>	A visualization tool for predicting antigen immunogenicity based on the HLA presentation	Unavailable	[Yamarkovich et al., 2020] [24]: used to identify 65 peptide sequence with potential coverage to 99.4% of population in the Bone Marrow Registry
<b>TCR-peptide binding prediction</b>			
<b>iPred</b>	Estimate baseline frequencies of TCR specificities by database of TCR sequence with known antigen specificities	Unavailable	[Shomuradova et al., 2020] [265]: Used for motif discovery for TCR repertoires in epitopes [Lee, 2020] [268]: used for predicting immunogenicity between the SARS-CoV-2 epitope and IEDB peptides
<b>ERGO</b>	NLP-based methods in predicting CD4+ or CD8+ TCR binders using large-scale TCR-peptide dictionaries	Self-reported optimal accuracy of 0.98 AUC	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>NetTCR</b>	A CNN-based sequence-based predictor trained on AA sequence of peptide and CDR3 region of TCR $\beta$ -chain	Self-reported optimal accuracy of 0.727 AUC	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>TcellMatch</b>	Multiple deep learning methods trained on data of MHC binding and TCR sequences from multimodal single-cell experiments	Self-reported optimal accuracy of 0.87 AUC	Not utilised/cited in any SARS-CoV-2 vaccine design publications
<b>IFNepitope</b>	SVM-based classifier applied on AA composition and length of IFN- $\gamma$ inducing MHC class II peptides	Self-reported optimal accuracy of 81.39%	[Behmard et al., 2020] [131]: predicted of IFN- $\gamma$ release in selected epitopes [Ayyagagari et al., 2020] [112]: predicted of IFN- $\gamma$ release in selected epitopes [He et al., 2020] [132]: predicted IFN- $\gamma$ inducing property of MHC epitopes [Samad et al., 2020] [133]: predicted IFN- $\gamma$ inducing property of MHC epitopes [Ahammad et al., 2020] [133]: predicted IFN- $\gamma$ inducing property of MHC epitopes
<b>IL4pred</b>	SVM-based classifier applied on IL4 inducing peptides for T <sub>H</sub> 2 helper cell epitope prediction	Self-reported optimal accuracy of 75.76%	[Behmard et al., 2020] [131]: predicted of IL-4 release in selected epitopes [Sarkar et al., 2020] [234]: predicted IL-4 inducing property of MHC epitopes [He et al., 2020] [132]: predicted IL-4 inducing property of MHC epitopes [Samad et al., 2020] [133]: predicted IL-4 inducing property of MHC epitopes [Ahammad et al., 2020] [134]: predicted IL-4 inducing property of MHC epitopes
<b>IL10pred</b>	RF-based model trained on motif and residues of IL-10 inducing peptides	Self-reported optimal accuracy of 0.88 AUC	[Behmard et al., 2020] [131]: predicted of IL-10 release in selected epitopes [Sarkar et al., 2020] [234]: predicted IL-10 inducing property of MHC epitopes [He et al., 2020] [132]: predicted IL-10 inducing property of MHC epitopes [Samad et al., 2020] [133]: predicted IL-10 inducing property of MHC epitopes [Ahammad et al., 2020] [134]: predicted IL-10 inducing property of MHC epitopes

This table shows some of the common T-cell epitope prediction tools and their reported accuracy or evaluated accuracy, if available. Their utilities in SARS-CoV-2 vaccine design papers are summarised briefly by the result they have indicated and their application on which protein on SARS-CoV-2, if mentioned. As seen, some tools with relatively high accuracies are yet to be utilised for SARS-CoV-2 vaccine designs.

[124]. Again, their utility in the SARS-CoV-2 study and reported accuracy are summarised in Table 3.

### 3.2.3. HLA coverage

With the aim of inducing herd immunity on a global scale against SARS-CoV-2, it is important that a vaccine candidate is effective across the human population. Campbell *et al.* prioritised

the HLA frequency data from Allele Frequency Net Database (AFND) [125] to ensure their epitope candidates can bind to greater than 90% of populations across 57 countries [126]. Other computational tools accounting HLA allele subtype coverage can estimate the immunogenicity of populations with different HLA subtype backgrounds. Yarmarkovich *et al.* applied ShinyNap [127] for scoring HLA presentation on a population scale for Cancer to

predicted epitopes for SARS-CoV-2, additionally they calculated their epitopes were not similar to endogenous human sequences [24]. Liu *et al.* created two tools for the optimisation (OptiVax) and evaluation (EvalVax) of population coverage of epitopes for peptide vaccines [128]. Additional information can be found in Table 3.

### 3.2.4. TCR binding prediction

Following binding prediction to MHC molecules, there are also several direct methods used to predict whether the MHC-bound peptides can be recognized by the T cells (see Table 3). For example, ERGO is a Natural Language Processing (NLP) based method in creating a large-scaled T cell receptor (TCR)-peptide dictionary for TCR-peptide binding prediction and is applicable for both CD4+ and CD8+ T cell epitopes [129]. Furthermore, cell-type specific TCR-peptide binding predictions are also available, such as IFNepitope for predicting epitope binding to IFN- $\gamma$ -induced T<sub>h</sub>1 cells [130]. The combined utility of IL4pred, IL10preds and IFNepitope are seen in several SARS-CoV-2 vaccine design studies to check the corresponding cytokine inducibility of their MHC epitope candidates [131–135].

### 3.3. Epitope prediction evaluation and challenges

Despite the exhaustive list in Table 2 and Table 3, there are still many more tools available for epitope prediction. Even though some models have similar or even better accuracy, only minorities are utilised in the SARS-CoV-2 epitope design, possibly due to the convenience introduced by the web servers. Furthermore, several efforts are seen in benchmarking various tools but none when used in combinations [60,135–138]. It is beyond our scope to conclude the best epitope prediction pipeline for the purpose of this review. However, we advise readers to refer to Sohail *et al.* [135] where the performance of each T cell epitope prediction method used in SARS-CoV-2 vaccine designs is reviewed. There is an additional concern in the novelty of epitopes predicted in studies using the same prediction methods. Alternatively, readers can refer to Prachar *et al.* [137] and Sohail *et al.* [135] to avoid utilising a suboptimal tool and additional guidance for other novel approaches.

SARS-CoV-2 is indicated to mutate  $9.8 \times 10^{-4}$  substitutions per site per year [139] and at least 15 variants have been reported throughout 2020 [140]. It is imperative that an epitope selected for vaccine development is conserved across mutations as well as across multiple populations. The IEDB conservancy tool [141] has been used to do this against SARS-CoV-2 epitope candidates [23,142]. Clustal omega, EBI's sequence alignment tool was also used [22,24]. Some studies have also screened for conservation between SARS-CoV-2 and other coronaviruses such as SARS-CoV and MERS-CoV [24,26].

In addition to the listed methodologies, Gupta *et al.* developed a platform of potential vaccine candidate or epitopes from SARS-CoV-2 proteins with predicted immunogenicity [91]. Similarly, Ahmed *et al.* (COVIDep) [18] and Wu *et al.* (COVIEDb) [143] presented databases of potential T and B cell epitopes for SARS-CoV-2. COVIEDb also provides SARS-CoV, and MERS-CoV with regularly updated validation results [143]. These resources could be used either for cross-validation or ensuring novelty of the epitope from other studies.

## 4. Immunogenicity prediction

The ongoing pandemic of COVID-19 has urged the scientific community to find answers, both in terms of repurposing existing drugs and expediting development of new vaccine candidates [15]. Currently, immunogenicity predictions of most SARS-CoV-2 vac-

cine studies focus on the static interaction between the antigen and the immune system component. These predictions, although useful, are unable to give specific knowledge about how the immune system, as a whole, acquires the pathogen-specific immune memory and fights off the pathogen. Such knowledge is only acquired during the later *in vivo* studies and may increase risks of failure, decrease productivity, and escalate costs. Additionally, limited knowledge about genomic instability and immunogenicity of genome-based (i.e., DNA, RNA and mRNA) vaccines encourage the development of novel strategies in studying the vaccine immune response dynamics [144].

Indeed, there is an increasing number of publications advocating the usage of relevant *in silico* modelling in vaccine design [145–148]. In their review, Brown *et al.* [147] pointed out that system-based approaches in biology and clinical medicine have been almost entirely data-driven, despite the limited insight these models give into the biological mechanisms involved. In contrast, mechanistic models are constrained by prior knowledge, possess greater predictive value under extrapolation, and can be iteratively improved by comparison against newly acquired data—a process that can lead to gradual model refinement. They mention three aims mechanistic modelling is well aligned with: 1) gaining deeper understanding of the protective immune responses mounted against pathogens, 2) building a mechanistic context, which can enable a more thorough interpretation of clinical data sets, and 3) designing better informed medical experiments and clinical trials, with the potential to lead to lower rates of therapeutic failures.

Rhodes *et al.* [146] focused instead on the dose-finding practices currently in place for new vaccines and proposed a novel *in silico* method which they termed Immunostimulation/ Immunodynamic (IS/ID) modelling. Typically, large dose ranges are first tested on small animal models, such as rats and mice, until a maximum plateau in response is reached. Allometric scaling assumes that physiological parameters follow a power-law relationship with biological measures, such as weight or height. Thus, allometric scaling is applied, even in the case of interspecies ones, allowing for effective human doses to be estimated. However, the applicability of allometric scaling on cross-species translation is challenging, as the immunological relationships involved are not well characterised. Additionally, dose escalation studies often rely on the long-standing assumption that the dose-host response relationship is saturating (sigmoidal). However, in some cases, a peaked dose–response curve has instead been observed. This exposes the potential risk of making sub-optimal, higher than needed, dosing decisions. The WHO has conducted retrospective dose ranging studies on vaccines in the past, where fractional doses were found to be equally, or in some cases more, immunogenic than the full licensed ones for diseases such as yellow fever, meningitis, and malaria [146].

The importance of informed dosing decisions cannot be overstated. In the case of Oxford/AstraZeneca's COVID-19 vaccine (AZD1222), an efficacy of 62.1% was reported for participants who received two standard doses, while in those who received a lower first dose, followed by a standard second dose, an increased efficacy of 90% was observed [21]. In this context, *in silico* immunogenicity prediction can potentially be used to better inform first-in-human dose selection, as studies transition from the preclinical to the clinical setting.

### 4.1. Mathematical modelling on SARS-CoV-2 infection

Many mathematical models of the immune response exist in literature today, where the complex immune system is simplified and defined in formalisms, the algorithmic or mathematical frameworks. These often rely on ordinary differential equations (ODEs) or partial differential equations (PDEs) or follow an agent-based



**Table 4**  
Common formalisms used in modelling immunogenicity.

Formalism	Advantages	Disadvantages	Application to SARS-CoV-2 Immune Response Studies
<b>Ordinary differential equations (ODEs)</b> [190–192,269]	Admit analytical solution; Can provide precise continuous-time dynamics of networks with multiple entities	Has a finite dimensional state vector; Unable to capture spatial dynamics; Relatively low scalability	[155–158,160,163]
<b>Partial differential equations (PDEs)</b> [270]	Admit analytical solution; Allow derivatives of unknown function for dynamics in time and space; Has an infinite dimensional state	More complex than ODEs and DDEs; Computationally demanding; Relatively low scalability	[160,161]
<b>Impulsive differential equations (IDEs)</b> [271]	Allow derivatives with sudden changes in states over continuous model; Can be multi-parametric;	Relatively low scalability; Most are fixed moments; Traditional IDEs do not allow non-instantaneous impulses	[163,272]
<b>Cellular automata</b> [273]	Able to exhibit complex system-level behaviour with qualitative results	Non-quantitative; Low scalability; Computationally demanding and time consuming; Difficult to transfer results into biological interpretation	[274]
<b>Agent-based modelling</b> [153]	Explores behaviour of individual entities; Some agents are adaptive in their behaviour; Depicts local interactions and environmental heterogeneity	Non-quantitative and entities of properties are discrete; Not suitable for intracellular networks	[159,161,275]

modelling (ABM) approach and are applied to both synthesizing the system through multi-scale or *trans*-scale integration. Some of the commonly used formalisms in simulating humoral and/or cell-mediated immunity upon infection have been summarised in Table 4. Additional examples of *in silico* modelling of immunogenicity have been discussed in Hande *et al.* [149], Charoentong *et al.* [150], Dobrovolny *et al.* [151], and Narang *et al.* [152].

There is no global best formalism for system immunology modelling and the choice of model is dependent on the study objective [149,152]. Modelling methodologies such as differential equations are more suitable to depict a homogeneous system whereas cellular automata and agent-based modelling are more commonly used in the simulation of a heterogeneous system [153,154]. Cellular automata and differential equations are also more commonly used to study the intracellular network as compared to the uses of agent-based modelling in studying a network composed of both innate and adaptive immunity. ODEs and PDEs are also more favourable over cellular automata and ABM if a numerical analysis is needed to validate the hypothesis of interest. For example, ODEs have been used in several simulations to study the within-host response to SARS-CoV-2 infections [155–158]. Alternatively, multi-formalism modelling is possible in overcoming the pitfalls of each formalism, and examples of applications to SARS-CoV-2 immune response studies are seen in Sahoo *et al.* [155], Getz *et al.* [159], Peter *et al.* [160], and Fain *et al.* [161]. For example, Hernandez-Vargas and Velasco-Hernandez [162] used target cell-limited models integrated with ODEs in exploring the T cell-induced immunity upon SARS-CoV-2 infection. The model also adapted mathematical terms in evaluating the consequence upon the introduction of a hypothetical drug. Although the approach is succinct, this paper identified some potential mechanisms regarding the timeline of symptoms-development and the benefits of prophylactic medication upon SARS-CoV-2 infection. Similarly, Getz *et al.* [159] has assembled a multidisciplinary team of scientists in developing a Physicell-based multi-scale model in simulating critical processes during SARS-CoV-2 infection. The framework is based on agent-based modelling in imitating the 2D and 3D tissue environments coupled with cell phenotypic parameters using ODEs.

In the study by Sahoo *et al.* [163], the model predicts that an intermediate strength innate immune response integrated with a

weaker adaptive immune response will require longer virus clearance time, as reflected by the age-dependent severity in COVID-19 in clinical settings [164]. In another study, their mathematical model suggests that the disease severity is increased for people who are exposed more often to SARS-CoV-2 [156]. This is consistent with a clinical study where those infected in a high transmission setting of Wuhan have a high odd ratio for severe COVID-19 outcomes as compared to low transmission settings [22]. The model in Fain *et al.* [161] also indicates that the infection duration can range up to 73 days, despite most clinical studies suggesting a median duration of 14–20 days. This is supported by several clinical findings with patients who shed the virus for more than 60 days post-hospitalization [156,165,166]. These findings suggest that although the models are simplified, they can provide additional comparable insights into the infection. Later in this review, we will discuss how mathematical models can be applied in studying the vaccine immune response.

#### 4.2. PK/PD studies in vaccine design

Mathematical modelling approaches have become an integral part of drug discovery and development in recent years and are used when studying drug PK and pharmacodynamics (PD). Surprisingly, there is only limited knowledge available on vaccine absorption, distribution, metabolism, and excretion (ADME) processes. Indeed, there is the long-lasting belief that PK studies are irrelevant in evaluating vaccine efficacy. This is also reflected on the current regulatory registration processes across the world, which require no experimental PK studies to be conducted for vaccine approval [16]. Here, the flawed premise of this notion is challenged by briefly revisiting how (adjuvanted) vaccines induce immunogenicity. Currently, PK modelling is rarely used in vaccine (antigen or adjuvant) immunogenicity prediction. The few PK models that exist in literature are presented and their relevance to the COVID-19 vaccine development process is discussed.

##### 4.2.1. Timing and location in vaccine efficacy

The immune response can be broken down into two general categories: innate and adaptive. When an invading pathogen is encountered, cellular effectors of the innate immune response, such as macrophages and dendritic cells (DC), are able to detect

the encounter, through activation of pattern recognition receptors (PRRs) that survey both the extracellular and intracellular space, responding to conserved pathogen-associated molecular patterns (PAMPs). This sets off a cascade of events, including antigen uptake by innate cells, such as immature DCs, that subsequently differentiate to mature antigen-presenting cells (APCs). Mature APCs that have processed the antigen to peptide fragments, present it on their surface in forms appropriate for inducing naïve T cell activation (priming). This interaction takes place within the lymph nodes (LNs), a pivotal meeting point between lymph-migrating APCs and naïve T and B cells— principal drivers of adaptive immunity— thus establishing a link between the innate and adaptive arms of the immune response. Indeed, the adaptive response requires innate signals for its activation [14,167,168].

The LNs are areas of high T and B cell concentration, however, these are compartmentalised into specific sites. T cells reside in the paracortex (T cell zone), deeper in the LNs, together with a large population of migratory antigen-presenting DCs. B cells are located in the follicles (comprising the B cell zone), areas where non-migratory follicular DCs (FDCs) are also found. Upon exposure to antigen presented by APCs, naïve CD4+ T cells (T helper cells or  $T_h$ ) are activated and differentiate into functionally distinct subpopulations. This process is influenced by the available levels of antigen and higher doses are associated with greater production of follicular helper T cells ( $T_{FH}$ ). B cell priming requires binding to cognate antigen in its native form and is less sensitive to antigen levels [167,168].

Following activation, B cells enter specialised regions of the follicles that are named germinal centres (GCs). GCs are further split into a light and a dark zone. In the dark zone, activated B cells undergo proliferation and somatic hypermutation of their antigen receptors; when a round of proliferation ends, GC B cells migrate to the light zone, where they acquire antigen from the FDCs, which is then processed and appropriately presented on their surface to  $T_{FH}$  cells. This is a cyclic process known as B cell affinity maturation and signals provided by the  $T_{FH}$  cells drive the fate of the B cells. Upon exiting the GCs, B cells either become antibody-producing cells (short-lived, termed plasmablasts, or long-lived, named plasma cells), or memory cells [167].

The aim of vaccination is the induction of immunological memory. Modern subunit vaccines lack the pathogenic features (PAMPs) required for sufficient activation of the innate immune response, which in turn affects the effectiveness of the downstream adaptive response. In this case, adjuvants are paired with purified antigen in order to provide the inflammatory cues that the former lack. As discussed, however, efficient antigen transport to the LNs is a critical component in the cascade of events eventually leading to the formation of memory cells. The use of adjuvants adds one more layer of complexity that needs to be considered during vaccine design. In their excellent review, Irvine *et al.* [167] point out that cell-mediated antigen/adjuvant transport to the draining LNs (dLNs) is a relatively inefficient process compared to direct lymphatic transport. Furthermore, adjuvants that fail to accumulate in the dLNs and reach the blood (lack of right location), pose the additional risk of inducing systemic toxicity. All in all, it becomes apparent that the biodistribution profiles of the antigen and adjuvant should be evaluated together. High antigen levels in the absence of inflammatory cues within the LNs or vice versa (lack of right timing) will result in suboptimal immune response [167].

#### 4.2.2. Predicting antigen immunogenicity

The introduction of IS/ID modelling, while in its early stages, aims to translate PK/PD modelling from drug development to vaccine design. Similar studies still remain scarce in the literature as far as vaccines are concerned. These models, while lacking strong mechanistic nature, provide concrete insights on the predictive

capabilities of PK/PD methods in vaccine design. Indeed, IS/ID modelling has already been used to facilitate animal model decision-making and inform first-in-human dose selection.

Rhodes *et al.* [169] use data from macaques to predict IFN- $\gamma$  responses in humans, following tuberculosis (TB) bacillus Calmette-Guérin (BCG) vaccination. Two different macaques species, rhesus and cynomolgus, are the primary non-human primate models for TB. However, it has been shown that under the same experimental conditions, the outcomes can be species-dependent and even within the same species, the country of origin of macaque might affect the level of protection against infection and response post-vaccination. The model adopts a compartmental structure and nonlinear mixed-effects modelling (NLMEM) are used to describe the IFN- $\gamma$  dynamics of two CD4+ T cell populations: transitional effector memory (TEM, short-lived) and resting “central” memory (CM, long-lived) cells. A proportion of TEM cells, whose rate of post-vaccination recruitment is time-dependent, undergo apoptosis, and the remaining transition to CM cells.

Notably, this is an over-simplified model, consisting of 5 parameters in total, out of which 3 are fitted to the experimental data. The model, separately fitted to macaque and human *ex vivo* IFN- $\gamma$  data, is shown to describe the empirical data well. It is then used, among others, to: 1) identify covariates that explain the within-population variation and 2) test which fitted macaque models best predict human IFN- $\gamma$  responses. Macaques were found to be best stratified by colony and humans by baseline BCG status, and these two were significantly associated regarding the BCG-induced IFN- $\gamma$  response. Indonesian cynomolgus macaques and Indian rhesus macaques best predicted the immune response of baseline BCG-naïve humans, while Mauritian cynomolgus macaques were found to do the same for baseline BCG-vaccinated ones.

In a subsequent publication, Rhodes *et al.* [169] use the same IS/ID modelling framework to inform TB vaccine dose decision-making in humans. The model is calibrated on IFN- $\gamma$  response data for the TB vaccine H56 adjuvanted with IC31<sup>®</sup> (H56 + IC31) in mice and humans, and H1 + IC31 data in humans only. The mouse data were stratified by dose group and the TEM to CM cell transition rate ( $\beta_{TEM}$ ) was the single parameter selected to differ among them. Subsequently, the  $\beta_{TEM}$ -dose curve was estimated for mice, assuming a peaked curve profile based on prior studies. This was used to extrapolate to doses ranging 0.01–50  $\mu$ g for H56 + IC31. In humans, the single available experimental point for 50  $\mu$ g H56/H1 + IC31 was used, together with an allometric scaling factor, to predict (human)  $\beta_{TEM}$  values for doses in the 0.1–500  $\mu$ g H56 + IC31 range, using the mouse-predicted curve as a starting point. This approach enabled the authors to predict that doses in the range of 0.8–8  $\mu$ g may be as, or more, immunogenic in humans as larger doses. This prediction was later independently supported by a Phase 1/2a H56 + IC31 dose-ranging clinical trial [146].

To put the importance of this into perspective, there is an ongoing scientific effort to find suitable animal models for COVID-19 vaccine development [170]. The implications of utilising similar approaches in vaccine development, range from accelerating the process to enabling the identification of animal models that are better suited for special human populations (such as the elderly or disease groups).

#### 4.2.3. Predicting adjuvant immunogenicity

Modern vaccine design aims to maximise efficacy while minimising potential serious adverse effects. To this end, newer approaches have shifted away from early live-attenuated and inactivated whole-pathogen vaccines, towards platforms with improved safety profiles, such as recombinant viral vectored, nucleic acid-based, and protein subunit vaccines.

Adjuvants play a multifaceted role in vaccine development; in addition to boosting the elicited immune response, the use of an

appropriate adjuvant can also induce a broadened antibody response with improved magnitude and functionality, due to a greater number of functional antibodies, antibodies with higher affinities, or both. Appropriate adjuvant selection can also guide towards desirable T cell responses. For example, Dong *et al* [33] propose that  $T_H1$ -type biased adjuvants may be preferable when compared to those inducing  $T_H2$ -type immune responses, since the latter have been connected with increased lung immunopathology in SARS-CoV-2 vaccines.

As discussed, the PK profiles of adjuvants can have profound impact on the immune response generated by vaccines that rely on them for inflammatory cues. Thus, *in silico* PK approaches could be used to assess adjuvant biodistribution profiles and, in turn, allow for informed immunogenicity predictions to be made. PK models tend to adopt a compartmental representation, but the compartments involved often lack physiological meaning. The concept of physiologically-based pharmacokinetic (PBPK) modelling differs from other PK modelling approaches in that each compartment corresponds to an organ or tissue, though lumped compartments are sometimes used for dimensionality reduction [171].

Such models come with many advantages due to their nature, which is inherently mechanistic. Among their typical applications, PBPK models are often used for translating existing understanding to novel settings, e.g., different species or special populations. Other important applications include, but are not limited to, informing dosing regimens and generating and testing hypotheses regarding the physiological processes that drive observed drug behaviour [172]. When approached from a different point of view, it becomes apparent that the majority of the aforementioned applications are, essentially, trying to answer the same question: how may intrinsic factors—age, gender, race, weight, height, pregnancy, and organ dysfunction— influence exposure (PK) and/or response (PD) to a specific drug? [173]

Achieving improved vaccine efficacy has significant advantages that become ever more relevant in the case of a pandemic. Reduction in the amount of antigen needed while maintaining the target antibody response would be expected to increase manufacturing capacity. Similarly, shortening the vaccination regimen— as in GlaxoSmithKline's (GSK's) Fendrix™ hepatitis B vaccine, where addition of the AS04® adjuvant enabled the reduction from a three to a two-dose regimen— could ease the logistical burden that distributing doses on a massive scale poses [174].

Despite their popularity in drug development, PBPK modelling has rarely been used in studying vaccine adjuvants. Tegenge and Mitkus [175] have developed a whole-body PBPK (WB-PBPK) model for squalene-containing adjuvants in human vaccines. In their initial publication, the model is used to make predictions for intramuscularly (IM) injected commercial MF59® adjuvant, a squalene-in-water (SQ/W) emulsion. In a subsequent study, the same generic WB-PBPK model is extended and applied to AS03®, which is a SQ/W adjuvant containing  $\alpha$ -tocopherol.

The original and subsequent models, adopting the typical PBPK structure and assumptions [171] are a collection of tissue-representing compartments, connected by a circulating blood system. Each tissue compartment is described as a mass-balance ODE under the assumptions of perfusion rate-limited and well-stirred conditions and defined by tissue volume and blood flow rate. These are physiological parameters, only specific to the species of interest, and widely available in the literature. Due to the inclusion of draining and distal LNs, the compartments are also described in terms of lymph flow rates, which are set at a fraction of the corresponding tissue blood flow. Niederalt *et al.* [176], making a similar assumption for lymphatic flow to Jones *et al.* [171], estimated the fraction by fitting to experimental concentration–time profiles. Similarly, in their original publication, Tegenge and Mitkus [175] assume *a priori* almost exclusive lymphatic transport of squalene

in part due to its highly lipophilic nature. In a later publication [177], an optimal fractional lymphatic transport value is estimated for a sheep PBPK model by fitting with available experimental data. This is then used to inform a human PBPK model, that is otherwise built on top of human-specific physiological parameters. This, while not ideal, is testament to the translational power of this type of modelling; the ability to extrapolate across species allows for predictions to be made, even under limited data availability.

Only passive transport (diffusion) is considered due to both blood and lymphatic flow. For that, tissue-plasma partition coefficients for squalene were estimated using mechanistic tissue composition-based equations for highly lipophilic compounds [175]. These methods, belonging to the broader family of *in vitro*-*in vivo* extrapolation (IVIVE) techniques, estimate tissue distribution from physicochemical and *in vitro* binding characteristics of the compound [171]. Inter-species extrapolation is therefore significantly simplified, since tissue-plasma ratios depend on the difference in tissue composition among species, physiological parameters that are generally available in literature [178]. Additionally, Tegenge and Mitkus [175] modelled emulsion cracking in deltoid muscle and dLNs, described by first-ordered kinetics fitted to *in vivo* data in squirrel monkeys. Squalene metabolism in adipose tissue and liver, as well as fecal excretion, were also considered following first-order kinetics and fitted to *in vivo* data in rats and humans. There is an important point to be made here; PBPK modelling, despite introducing complexity, maintains a clear separation among physiological and substance-specific parameters. The former tends to be easily accessible from various sources in literature, while, for the latter IVIVE, techniques exist that reduce the dependence on *in vivo* data and thus allow for easier extrapolations. Furthermore, in the event where enough experimental PK data exist, more advanced estimation methods can be applied, even in cases of highly dimensional parameter spaces [179].

While not often used, PBPK models like those presented above could be applied on different scenarios to improve adjuvant selection with minimal effort. This is due to the separation of the physiological parameters (that remain unchanged between models) from the substance-specific ones which simplifies model parameterisation, and the increased availability of IVIVE techniques that allow *in vitro* data, which is easier to generate, to be used. All in all, the few examples of PBPK models applied in vaccine design exist only for adjuvants. However, these models adopt a generic structure, in that they can be applied to different species and compounds with minimal changes or be extended to cover novel scenarios in a relatively straightforward manner. As such, while PBPK modelling has been traditionally utilised in drug development, a compelling case can be made that their applicability in vaccines might come with minimal changes and that vaccine design, in general, could benefit just as much by the advantages this technique has to offer.

#### 4.2.4. Predicting antigen ADME properties

The adjuvant-focused studies of Tegenge and Mitkus [175,177] incorporate dLNs and LNs compartments, assume tissue-level lymph flow as a fixed fraction of the corresponding tissue blood flow rates and local preferential lymphatic transfer. To date, there are several PBPK models that include a more involved representation of the lymphatic system [176,180–183]; such models are usually built to predict the biodistribution of monoclonal antibodies (mAbs) in animal models or humans. Also, these typically share a common framework, first introduced by Baxter *et al.* [184], who incorporated the two-pore formalism to describe transcapillary exchange of mAbs. Macromolecular compounds, proteins and peptides convect and diffuse across the barrier between plasma and interstitial space by passing through large and small pores. Fluid

recirculation also occurs, with small pores acting as filters, trapping large molecules in the interstitial space, while excess fluid flux into the latter is taken up by the lymphatic system [176,181]. Niederalt *et al.* [176] give a detailed overview of the model, but the hydrodynamic radius of the compound is the only drug-dependent parameter needed.

The two-pore model of extravasation can be used to describe the passive transport of both antigenic proteins or peptides, and endogenous IgG. Moreover, in contrast to the simplistic approach taken by Tegenge and Mitkus [175], where each compartment represents a single tissue, almost every model here splits tissues into several sub-compartments [176,181,182,185]. Among the two already mentioned, plasma and interstitial space, other sub-compartments can include blood cells (vascular space), cellular space, and endosomal space [176]. This detailed tissue representation requires no preferential lymphatic transfer assumption to be made, as lymph flow is instead incorporated in a mechanistic manner, while also enabling better localisation of active processes and reactions. Importantly, the additional parameterisation needed is physiologically-focused, e.g., tissue-plasma partition coefficients, therefore compound-independent, and can also be handled by IVIVE estimation techniques.

A notable example of a localised active process is incorporation of the neonatal Fc receptor (FcRn) binding [176,181,185]. The endosomal space represents the region within the vascular endothelial cells where catabolic clearance of IgG and albumin fusion proteins takes place. These proteins are taken up from plasma and interstitial space through pinocytosis; in the endosomes, they reversibly bind to FcRn with high affinity, due to the acidic environment. The protein-FcRn-complex is recycled back to plasma and interstitial space where it dissociates given the low affinity for FcRn binding due to the neutral environment. Any unbound proteins in the endosomal space degrade. The parameters involved in FcRn binding are species-specific [176,185]. Models that incorporate tissue-specific FcRn expression by assuming a proportional relationship among FcRn concentration and its tissue-specific mRNA expression also exist in literature [186]. This sub-model could be of great interest in the case of “albumin hitchhiking” [187,188], too, where antigens are modified with a lipophilic albumin-binding domain to achieve greater accumulation in LNs.

#### 4.2.5. Incorporating epitope binding affinities

The components discussed until now account for passive transport of both antigen (macromolecules) and adjuvant (assumed as small molecules) compounds, which includes lymphatic transfer to LNs. They do so by adopting a fully mechanistic representation, while requiring a minimal amount of compound-specific parameters. For a complete integrated approach, however, there is still need for an immune response-mediated antigen/adjuvant clearance and PD response component. A straightforward implementation would be to incorporate pre-existing immune response models, as the ones previously presented.

Quantitative systems pharmacology (QSP) techniques [189–192], in particular, are of much interest for this. Jafarnejad *et al.* [192] utilise, among others, the formalism given by Chen *et al.* [190]; here, we focus on the latter, as it presents a generalised framework, including applications [190], that is better aligned with the objectives of this review. This is a multiscale ODE model that at the whole-body level accounts for the *in vivo* disposition of the antigen. For the latter, an empirical PK model is utilised that could, in principle, be substituted for the more detailed PBPK approach discussed above.

Furthermore, mechanistic details are incorporated both at the subcellular and cellular level. The model considers the following immune cells: 1) DCs, chosen to represent all APCs, 2) CD4<sup>+</sup> T

helper cells, and 3) B cells, and includes activation and differentiation processes for all three. DC activation is modelled to be driven purely by lipopolysaccharide (LPS); after maturation, at the subcellular level, DCs uptake antigenic proteins, which are degraded into T epitope peptides within the endosomes, and then presented in MHC II complexes on their surface. This highly mechanistic description of antigen presentation allows for the integration of the number and MHC II binding affinities of T-epitopes, which can be obtained either *in vitro* or predicted *in silico*—as already discussed in this review. Indeed, the model is not only specifically built to allow for that, but also permits incorporation of subject-specific MHC II allele genotypes [189,190]. At the cellular level, naïve T cells that are activated through DC-presented antigen, proliferate and differentiate into either memory T cells or functional T helper cells. The former can be directly activated by mature DCs, while the latter facilitate downstream B cell activation. B cells then proliferate and differentiate into antibody-secreting short-lived and long-lived plasma cells, or memory B cells.

#### 4.2.6. Integrated PBPK models of vaccines

So far, techniques employed regarding antigen PK/PD (the IS/ID modelling approach) and adjuvant PBPK modelling have been explored. We focused on *in silico* studies that could be directly translated to COVID-19 vaccine design with minimal to no changes needed. As already discussed, the latter might have the potential to accelerate both preclinical and clinical studies. However, no integrated PBPK/PD modelling framework built around vaccine design exists to date; such approach could be thought to at least consist of: 1) a typical PBPK structure, expanded to include lymphoid compartments and transport, 2) passive compound transport and clearance processes, describing both adjuvant and antigen biodistribution, 3) active immune response-related transport and clearance of adjuvant and antigen compounds, and 4) an immune response PD component, that should be capable of predicting concentration–time profiles of markers relevant in preclinical and clinical research.

#### 4.3. Perspectives and challenges

Incorporating PBPK modelling of adjuvants in the vaccine design workflow of COVID-19 vaccines can provide significant benefits. First, in terms of accelerating the development process, partially through speeding up the transition from preclinical to first-in-human, Phase I, clinical studies. In their adjuvant models, Tegenge and Mitkus [175,177] extrapolate from limited *ex/in vivo* animal data in mice and sheep to humans. In similar fashion, Rhodes *et al.* [145] under limited available human data and more extensive, but still limited, mouse *ex vivo* data, use PK/PD modelling to predict the most immunogenic vaccine doses in humans. Taken together, these methods provide a significantly more refined approach to translating biodistribution profiles from animal models to humans than the currently-utilised allometric scaling approach. Future incorporation of a PB component could further strengthen the flexibility and predictive power of such approaches.

Improving dose translation capabilities are particularly important, as small animal models and non-human primates are used extensively in vaccine development; in regard to COVID-19, the WHO has been reporting periodically on the progress made in identifying suitable animal models, which are in part to be used to accelerate preclinical testing of vaccine candidates [170,193]. This is the same framework (PK/PD) that was also used to better inform animal model selection [169].

Second, in the presence of a validated human PBPK model, the framework can allow for extrapolation across human populations, potentially providing useful insights before, or during, Phase II and III trials are conducted. For example, in the presence of appropriate



*in vivo* PK data in healthy volunteers, such as in Phase I trials, estimation of the expected PK behaviour in specific populations, e.g. the elderly or people with underlying medical conditions or in underrepresented subjects, often women and BAME groups, could take place. Indeed, global evidence shows a higher COVID-19 burden with old age, male sex, obesity, and comorbidities. People from Black, Asian, and minority ethnic (BAME) groups in the UK and Black, Hispanic, and Native American groups in the USA are also at increased risk of COVID-19 complications and death [194]. In this case, computational methods capable of accounting for such covariates should provide predictions of increased importance. Notably, these are the same intrinsic factors discussed earlier, that PBPK modelling is built around of.

Such information might then be utilised for designing better-informed Phase II and III clinical trials and could potentially lead to *in silico*-augmented trials, in which both physical and virtual patients are combined [148]. Most importantly, given that appropriate biodistribution of both antigen and adjuvant is essential for vaccine efficacy, incorporating PBPK modelling in early vaccine development will increase the likelihood of progressing a successful vaccine.

In all vaccine development, acceleration of preclinical and clinical testing and reduction in clinical trials cost would be beneficial. These benefits are amplified under pandemic conditions that require rapid vaccine design and evaluation. However, while adjuvant PK data, sparse as may be, exist in literature, antigen related ADME processes are rarely evaluated. Indeed, antigen is administered in small amounts and widely spaced, making kinetic studies hard to implement [16]. A summary of PK parameters reported in vaccine development studies is given by Gómez-Mantilla *et al.* [16]. Then, the sparsity of such data can also be attributed, at least in part, to the absence of regulatory guidance mandating their existence for vaccine approval.

That, together with the high dimensionality of the parameter space (88 parameters in total [190]), as is often the case in PBPK/PD modelling, and mechanistic modelling in general, are expected to present a major challenge. However, many of the parameters needed are species-specific; Chen *et al.* [190] developed their model for both mouse and human. The ability to predict across species, in turn, should allow for easier adaptation into both preclinical and clinical settings. Also, by incorporation of MHC II allele frequency and T-epitope binding affinities, Chen *et al.* [190] were able to predict the immune response of 1,000 virtual subjects against adalimumab. If combined with the additional PBPK-provided subject-level PK information and integration of *in silico*-predicted T-epitope affinities of novel vaccine antigens, this could allow for early vaccine immunogenicity profile prediction in different human populations. Such tools could potentially prove useful in cases where rapid vaccine development is prioritised, like in the ongoing COVID-19 pandemic.

## 5. Toxicity and allergenicity prediction

### 5.1. Toxicity prediction

There are various adverse clinical effects associated to the currently authorized SARS-CoV-2 vaccines. Taken those approved by the UK authority, i.e., Comirnaty [19], mRNA-1273 [20], and AZD1222 [21], the common side effects include pain, swelling, and redness at the site of administration and reactogenicity symptoms including mild or moderate fever, myalgia or headache after 7 days post-injection. These adverse effects are potentially caused by inflammatory cells infiltration at the injection site or result from the host immune response evoked by the vaccine components. Most importantly, these side effects are tolerable and do not

possess lethal risks to the majority of the global population. In contrast, this may not be the case for some other vaccine candidates at the initial phase of vaccine design. It is possible that some promising candidates may present life-threatening toxicity and could not be picked up during *in vitro* or *in vivo* studies. More severe adverse events including trauma, macrophagic myofasciitis (MMF), enhanced diseases, and autoimmune diseases have been reported previously in toxicology studies of other vaccines [195–198]. These studies demonstrate the double-edged potential of vaccine-mediated immune-stimulation. Although the precise mechanisms of these severe adverse events are not well characterized, existing computational tools are available to eliminate vaccine candidates with high chance of evoking toxicities.

#### 5.1.1. Non-specific toxicity predictions for SARS-CoV-2 vaccine epitopes

Peptide toxicity predictions have been widely applied in facilitating the design of therapeutics with desired lack of toxicity [199]. Most prediction tools utilised ML approaches in classifying peptides based on their physicochemical, therapeutical, and toxicological properties [200–203]. Amongst the number of available tools, ToxinPred predicts peptide toxicity using a SVM classifier trained on SwissProt and TrEMBL peptide sequence data [201]. In ToxinPred, various features can be selected from the training data including AA composition, dipeptide composition, binary profile pattern, and motif-based profile for toxicity predictions. It has a high accuracy of 94% (using dipeptide composition) and its web interface allows analogue generation with minimum mutations to design a non-toxic version of the peptide. The precision and convenience make ToxinPred the commonest tool for toxicity prediction in the published SARS-CoV-2 vaccine design studies [22,204–206].

Alternative tools to ToxinPred are available. For example, ToxClassifier, a SVM-based classifier trained on their Tri-Blast Enhanced data, has the highest accuracy of 96% [202]. Unlike ToxinPred, ToxClassifier does not have a length limit of 30 AA. ToxClassifier also provides similar toxin or non-toxic sequence upon query submission. TOXIFY is a Recurrent Neural Networks-based classifier trained on UniProtKB protein sequences [203]. TOXIFY has a reported accuracy of 97.4% and shown superior to ToxClassifier in classifying non-toxic molecules. More recently, Jain and Kihara proposed NNTox, a neural network-based classifier trained on SwissProt data, for predicting 11 sub-classes of toxins. NNTox has an overall accuracy of 0.8 and can take DNA sequence as input. PredSTP is also available in predicting toxins based on the sequential tri-disulphide peptide from the Knottin database [207,208]. As far as we are aware from existing publications, despite the reported improved performance of these alternative tools, only ToxinPred [209] has been used for toxicity prediction. This contrasts to the combinatorial methodology used for epitope predictions in SARS-CoV-2 vaccine design studies. By applying multiple toxicity tools, the more rigorous screening can eradicate more candidates with a potential safety hazard at the beginning stage of vaccine development.

#### 5.1.2. Haemolytic toxicity predictions for SARS-CoV-2 vaccine epitopes

The above-mentioned tools provide prediction of all types of toxicity. They can be used to assist candidate-filtering in some of the conventional *in vitro* toxicity assays including lactate dehydrogenase (LDH) leakage, 3-(4,5-dimethyl-2-thiazolyl)-2, 5-diphenyl-2H-tetrazolium bromide (MTT), or adenosine triphosphate (ATP)-based assays. Other tools are available in indicating peptide's haemolytic toxicity, the ability to damage the red blood cells. In conventional studies, such activity is measured by the amount of haemoglobin released from membrane-compromised erythrocytes [210]. With the advances in computational power, haemolytic

activity can be predicted using tools such as HLPpred-Fuse [211], HemoPI [212], and HemoPImod [213].

Amongst the listed examples, HemoPI is constituted of HemoPI-1 (a dataset of 552 haemolytic peptides and 552 non-haemolytic peptides), HemoPI-2 (a set of peptides with low haemolytic potency,  $n = 462$ ), and HemoPI-3 (a dataset with high haemolytic potency,  $n = 1,623$ ) [212]. Users can select a SVM with/without motif-based models trained on all or specific dataset(s) according to their needs. For example, to filter a set of SARS-CoV-2 epitope candidates by their haemolytic potencies, the hybrid model trained on HemoPI-1 or HemoPI-2 should be chosen. HemoPI is available as a webserver and has been used in SARS-CoV-2 vaccine design studies [214,215]. Similarly, HLPpred-Fuse was trained with haemolytic peptides using six ML classifiers, including SVM and random forests, to select 54 probabilistic features (e.g., binary profile, conjoint triad, and grouped dipeptide and tripeptide composition etc.) for the final prediction model [211]. The result is given as the haemolytic peptide activity and probability values. More recently, Kumar *et al.* [213] developed HemoPImod, a classification model based on 4 classifiers trained with multiple peptide features, including its 3D descriptors. The reported accuracy of HemoPImod is 78.29% and require prediction of modified structure with PEPstrMOD [216]. In contrast to HemoPI, both HLPpred-Fuse and HemoPImod appear not to have been applied in SARS-CoV-2 vaccine design studies, possibly due to their time of launch.

### 5.1.3. Vaccine component toxicity predictions with computational network analysis

Another potentially toxic ingredient constituted in any vaccine is the adjuvant. As mentioned in the Immunogenicity section, adjuvants are included to evoke an effective immune response mostly through innate or humoral immunity. Of the authorized SARS-CoV-2 vaccines, alum has been used as an adjuvant in CoronaVac [217]. Other common adjuvants include minerals (oil), diethylaminoethyl (DEAE)-dextran, enterotoxin, CpG oligodeoxynucleotide (ODN), and cytokines. The toxicity of the adjuvant itself is well-studied and documented during its design stage. However, a predictive platform has not yet been developed which takes into account that all vaccine components in combination can evoke higher toxicity [218–220]. Here, we hypothesise that a computational network analysis-based method can predict the toxicity of multiple vaccine components based on their protein-interactors.

As indicated by Cheng *et al.* [221], network analysis on drug targets and disease proteins can be utilised in predicting efficacious combination therapies. In the study, the drug-target network was constructed based on the binding affinity data, inhibition constant/potency, dissociation constant, median effective concentration, or median inhibitory concentration from 6 databases (DrugBank [222], Therapeutic Target Database [223], PharmGKB-database [224], ChEMBL [225], BindingDB [226], and IUPHAR/BPS Guide to PHARMACOLOGY [227]). In the analysis, the adverse drug interactions were collected from DrugBank [222] and TWOSIDE [228] for adverse event-specific drugs. To study the potential side effects of a SARS-CoV-2 vaccine candidate, the binding affinity can be acquired through mass spectrometry-based proteomics with human cell *in vitro*. The data acquired can then be used to construct the protein interactome, in a similar method to Chent *et al.* [229]. Various network algorithms including Random Walk Restart and centralities can be used to extract key interacting proteins from the vaccine-protein interactome. The subsequent key proteins can deduct the potential side effects by consensus-filtering with SIDER database [230] or submission to toxicity-related protein database, such as DITOP [231]. Such approach assumes that the underlying mechanism for vaccine-induced toxicity would be similar to that of drug-induced toxicity. The method would capture cell-type specific toxicity which also presumes that

the intracellular vaccine-protein interaction would have a systematic effect. Alternatively, one could measure the interactome by applying proteomics on animal models administrated with the vaccine. Nevertheless, the network analysis method may not replace any *in vitro* or *in vivo* studies but would serve as an auxiliary study in understanding the mechanism of any potential toxicity arises from the vaccine candidate injection.

## 5.2. Allergenicity prediction

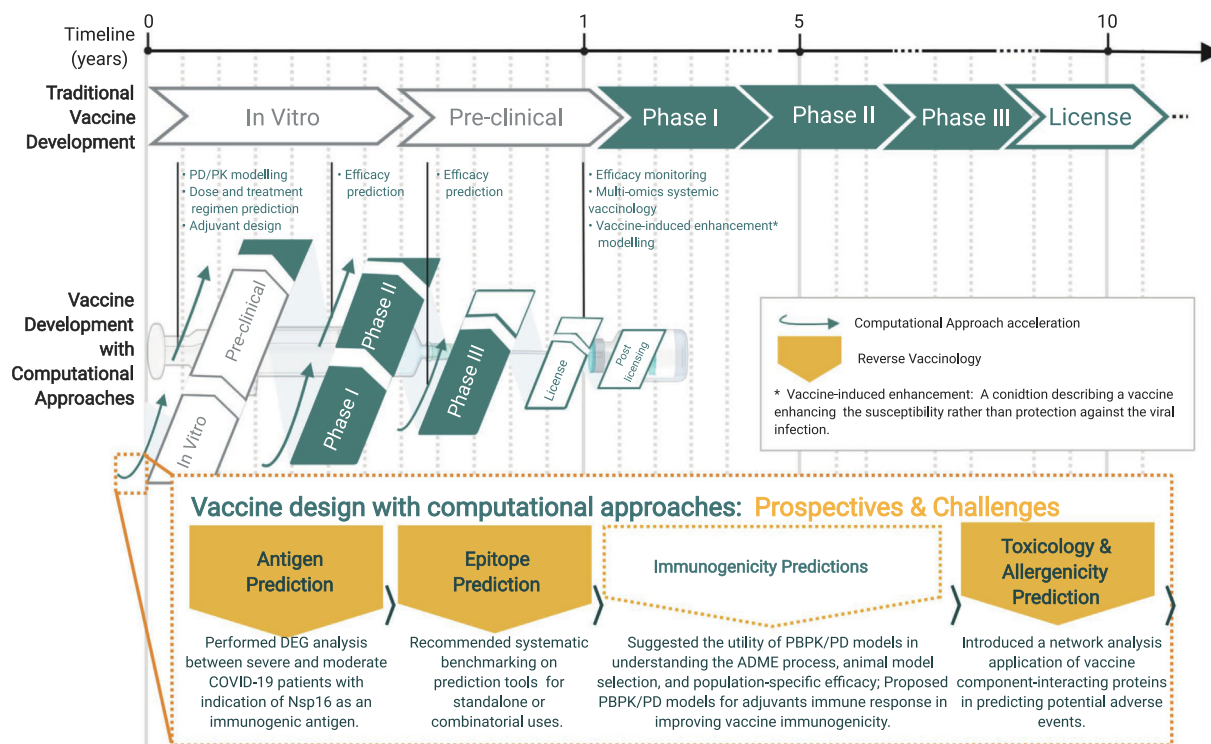
Allergenicity measures the ability of a substance in inducing hypersensitivity. There are four types of hypersensitivity. In type I hypersensitivity, increased production of IgE primes mast cells to release granules, histamine, and cytokines, thus causing the allergy symptoms. Type I hypersensitivity is responsible for manifestation in asthma, allergic rhinitis, food allergies, and anaphylactic reactions to substances. Hence, IgE binding affinity is frequently used for allergenicity predictions. For type II hypersensitivity, the immunoglobulins produced are IgG and IgM with subsequent complement system activation leading to cell damage or cell lysis. Similarly, in type III hypersensitivity, IgM, IgG, and IgA production increased for polymorphonuclear leukocytes chemotaxis and induce tissue damage, whilst type IV are associated with  $T_H1$ -mediated reactions where the symptoms are delayed, and macrophages or cytotoxic-T cells are activated in causing direct cellular damage.

### 5.2.1. IgE binding affinity predictions for SARS-CoV-2 vaccine epitopes

Similar to the technique used in epitope predictions, there are several ML-based classifiers available in predicting peptide binding affinity to IgE. Namely, AlgPred combines both SVM classifier trained on AA and dipeptide composition, motif-based method, IgE epitopes datasets, and BLAST search against allergen-representative peptides [56]. More recently, the latest version, AlgPred2.0 is trained on a larger dataset and has improved reported accuracy of 0.98 AUC [232]. AlgPred has been used in predicting the allergenicity of SARS-CoV-2 vaccine epitopes [204,233–236]. Another commonly used prediction tool is AllerTOP, a k-nearest neighbours classifier trained on 2,395 allergens [237]. AllerTOP was reported to have higher accuracy compared to AlgPred and is seen applied in Dar *et al.* [238], Behmard *et al.* [131], Ayyagari *et al.* [112], Dai *et al.* [76], and Das *et al.* [66]. Other popular prediction tools include AllerHunter [239], PREAL [240], ProAp [241], and AllergenFP [237]. It is beyond the scope of this review to benchmark these predictors. However, consensus-based filtering with multiple tools may decrease the chances of allergenicity amongst the epitope candidates, as employed in Sir-ohi *et al.* [204] and Behmard *et al.* [131]. Similarly, IgPred, a SVM-based B cell epitope prediction tool for predicting antibody-specific epitopes, can be used to eliminate candidates that tend to have a higher similarity to known-IgE epitopes [201].

### 5.3. Predicting vaccine-induced enhanced disease

Computational modelling of immunogenicity can also identify those who are least benefited by the COVID-19 vaccine candidate. Recent reports have suggested that the SARS-CoV-2 vaccine may be susceptible to inducing vaccine-induced enhanced disease—a condition reported being possibly caused by T helper 17 cells ( $T_H17$ -type immune response) or through antibody-mediated manner [242,243]. The immune enhancement can magnify SARS-CoV-2 infection and increase the severity of COVID-19 [243]. Although the condition is more commonly studied using *in vitro* and *in vivo* models [197,243], one may apply mathematical models in eliciting molecular and cellular conditions related to the enhanced disease. For example, ODEs have been applied for



**Fig. 2.** Summary of the advantages using computational approaches in vaccine development. Computational approaches can accelerate vaccine development at various stages. The top panel shows a traditional vaccine development process that requires at least 10 years of research and validation. In the centre, the diagram shows how different computational approaches (texts in teal) spur the vaccine development process. In the bottom panel, the timeline at the vaccine design stage has been enlarged. The box summarises the prospectives and challenges our review has proposed at each stage of our suggested computational-assisted vaccine design tools. Processes that have been associated as parts of reverse vaccinology are represented in filled-gold boxes.

predicting the likelihood of antibody-mediated enhancement in Influenza [197]. On the other hand, Carbo *et al.*'s method in predicting the probability of  $T_H17$ -type immunogenicity could be adapted. Through these methods, the immune enhancement-associated host factors can be identified and serve as a predictive biomarker for the vaccine in preventing disease exacerbation [244]. More importantly, the enhancement-associated properties of the vaccine can be identified and used to refine the vaccine design.

## 6. Discussion

Computational approaches can be employed to aid vaccine development in several ways and can accelerate the process (Fig. 2). While typical traditional vaccine development takes ~ 10 years, computational techniques could speed up this process and also reduce the associated costs. There has been a concerted, global effort to develop vaccines for COVID-19, which has recently seen three different vaccines approved with many more in the pipeline, and although computational approaches have been exploited in this effort, several techniques remain underutilised. In this review, we covered antigen prediction, epitope prediction, adjuvant selection and toxicology and allergenicity prediction, and the applicability of ADME modelling in vaccine design. Fig. 2 also includes other approaches that can stimulate the vaccine development process. This includes vaccine efficacy prediction during the clinical trial phases and multi-omics analysis to further our understanding of the disease and the vaccine.

Computational techniques for vaccine design have been used to select antigens, predict B and T cell epitopes, and evaluate these potential epitopes to ensure conservation across mutations in the virus and across the human population, and to predict immuno-

genicity and toxicity. There are other computational techniques which could aid vaccine design for COVID-19 which have so far been underutilised. Some research groups have studied the immune signature of moderate and severe COVID-19 patient cohorts to identify features which classify these cohorts [35,245,246]. The activities of the antigen-specific T cells, virus-neutralising antibodies, and the antigen presenting cells during SARS-CoV-2 infection can elicit the underlying protective mechanisms, thus these parameters that can influence vaccine epitope immunogenicity [162].

Attempts to develop COVID-19 vaccines have tended to focus on preventing viral entry into cells. For this reason, the S protein has been the focus for antigen selection. However, other proteins could be promising, including those which induce the adaptive immune system; our *meta*-analysis of four transcriptomics datasets suggests nsp16 warrants further investigation as a potential antigen for SARS-CoV-2, based on its interaction with proteins involved in host immune response. This may especially be interesting if any of the mutations which arise in the S protein render currently approved vaccines ineffective. In such a scenario, alternative antigen selection would be necessary to deal with the new variants.

Multiple groups have applied reverse vaccinology to predict epitopes for SARS-CoV-2 based on the AA structures. An issue with these approaches is that most studies use the same tools, mostly relying on web server software to predict epitopes. This over-reliance on specific tools/pipelines leads to a lack of diversity of approach which could mean promising epitopes are missed. Although newer, and indeed more accurate models have been developed, there has been a lack of uptake of these. There is also an urgent need for a systemic benchmarking of using different epitope prediction tools in combination. Similarly, many studies have



utilised structural databases for predicting epitopes with decreased risks of allergenicity and toxicology. However, there is currently no approach in predicting toxicity of all vaccine components in combination computationally, which may be achievable through computational network analysis.

PK studies are rarely carried out in vaccines. However, the importance of PK in vaccine safety and efficacy is well recognised [167]. This is especially relevant in subunit vaccines, the largest platform of COVID-19 vaccines at the moment, where the biodistribution of both adjuvant and antigen need to be well-timed together. However, modelling which has been applied looks at either antigen or adjuvant alone. Recently, many studies [145,179] advocate the use of *in silico* PK modelling in vaccine design. In the past, it has been shown it is possible to extrapolate from animal models to study adjuvant PK profiles in humans. There are also multiple immune response models studying human immunogenicity. PBPK/PD may not have been utilised in vaccine development, including COVID-19, yet, but has the potential to lower antigen doses through better informed adjuvant selection, accelerate the transition from preclinical animal models to human subjects, and make predictions for special populations that still remain underrepresented in modern clinical trials.

It is worth exploring these techniques and developing robust computational vaccine design strategies for COVID-19. If the unfortunate situation arises in which SARS-CoV-2 variants emerged with resistance to immunity induced by the approved vaccines [247], our proposed *in silico* approaches could accelerate the development of a vaccine against the new variant lineage. Specifically, antigen prediction based on cellular immune response could suggest alternative antigens for a new vaccine and the proposed immunogenicity and ADME prediction models could indicate if the appropriate subcomponents of the immune system would be activated to secure an effective immunisation.

Additionally, these strategies should be ready to be deployed against further zoonotic diseases as they emerge. The increase in contact between humans and animals due to the continued expansion of human society and the reduction of animal habitats is expected to result in similar viral infections emerging in the future. This is an era in which a system for fast and safe vaccine development is needed more than ever.

## Acknowledgement

The authors would like to thank Professor Gordon Dougan for his helpful advice and comments. W.H. and N.H. are funded by LifeArc.

## Competing interests

M.M. is an employee of LifeArc. K.C. is funded by pharmaceutical industry.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.addr.2021.02.004>.

## References

- [1] J.H.C.R. Center, COVID-19 Map, Johns Hopkins Coronavirus Resour, Cent, 2020, <https://coronavirus.jhu.edu/map.html>.
- [2] R.T. Gandhi, J.B. Lynch, C. del Rio, Mild or Moderate Covid-19, *N. Engl. J. Med.* 383 (2020), <https://doi.org/10.1056/NEJMcp2009249>.
- [3] D.M. Morens, A.S. Fauci, Emerging Pandemic Diseases: How We Got to COVID-19, *Cell.* 182 (2020) 1077–1092, <https://doi.org/10.1016/j.cell.2020.08.021>.

- [4] E. Jenner, *An Inquiry Into the Causes and Effects of the Variolae Vaccinae: A Disease Discovered in Some of the Western Counties of England, Particularly Gloucestershire, and Known by the Name of the Cow Pox*, author (1801).
- [5] S. Han, Clinical vaccine development, *Clin. Exp. Vaccine Res.* 4 (2015), <https://doi.org/10.7774/cevr.2015.4.1.46>.
- [6] M.R. Hilleman, Vaccines in historic evolution and perspective: a narrative of vaccine discoveries, *Vaccine.* 18 (2000), [https://doi.org/10.1016/S0264-410X\(99\)00434-X](https://doi.org/10.1016/S0264-410X(99)00434-X).
- [7] J.C. Adkins, A.J. Wagstaff, Recombinant hepatitis B vaccine: a review of its immunogenicity and protective efficacy against hepatitis B, *BioDrugs Clin. Immunother. Biopharm. Gene Ther.* 10 (1998) 137–158, <https://doi.org/10.2165/00063030-199810020-00005>.
- [8] N. Pardi, M.J. Hogan, F.W. Porter, D. Weissman, mRNA vaccines – a new era in vaccinology, *Nat. Rev. Drug Discov.* 17 (2018) 261–279, <https://doi.org/10.1038/nrd.2017.243>.
- [9] M. Mora, D. Veggi, L. Santini, M. Pizza, R. Rappuoli, Reverse vaccinology, *Drug Discov. Today.* 8 (2003), [https://doi.org/10.1016/S1359-6446\(03\)02689-8](https://doi.org/10.1016/S1359-6446(03)02689-8).
- [10] R. Rappuoli, Reverse vaccinology, *Curr. Opin. Microbiol.* 3 (2000), [https://doi.org/10.1016/S1369-5274\(00\)00119-3](https://doi.org/10.1016/S1369-5274(00)00119-3).
- [11] D. Maione, Identification of a Universal Group B Streptococcus Vaccine by Multiple Genome Screen, *Science* (80-. ) (2005) 309, <https://doi.org/10.1126/science.1109869>.
- [12] M. Pizza, Identification of Vaccine Candidates Against Serogroup B Meningococcus by Whole-Genome Sequencing, *Science* (80-. ) (2000) 287, <https://doi.org/10.1126/science.287.5459.1816>.
- [13] J. Finne, D. Bitter-Suermann, C. Goridis, U. Finne, An IgG monoclonal antibody to group B meningococci cross-reacts with developmentally regulated polysialic acid units of glycoproteins in neural and extraneural tissues, *J. Immunol.* 138 (1987) 4402–4407. <https://www.jimmunol.org/content/138/12/4402>.
- [14] A. Di Pasquale, S. Preiss, F. Tavares Da Silva, N. Garçon, Vaccine Adjuvants: from 1920 to 2015 and Beyond, *Vaccines* (Basel). 3 (2015) 320–343, <https://doi.org/10.3390/vaccines3020320>.
- [15] C.D. Funk, C. Laferrrière, A. Ardakani, A Snapshot of the Global Race for Vaccines Targeting SARS-CoV-2 and the COVID-19 Pandemic, *Front. Pharmacol.* 11 (2020) 937, <https://doi.org/10.3389/fphar.2020.00937>.
- [16] J.D. Gómez-Mantilla, I.F. Trocóniz, M.J. Garrido, ADME Processes in Vaccines and PK/PD Approaches for Vaccination Optimization, in, *Pharm. Sci. Encycl., American Cancer Society* (2015) 1–22.
- [17] D.E. Gordon, G.M. Jang, M. Bouhaddou, J. Xu, K. Obernier, K.M. White, M.J. O'Meara, V.V. Rezelj, J.Z. Guo, D.L. Swaney, T.A. Tummino, R. Hüttenhain, R.M. Kaake, A.L. Richards, B. Tutuncuoglu, H. Foussard, J. Batra, K. Haas, M. Modak, M. Kim, P. Haas, B.J. Polacco, H. Braberg, J.M. FABIUS, M. Eckhardt, M. Soucheray, M.J. Bennett, M. Cakir, M.J. McGregor, Q. Li, B. Meyer, F. Roesch, T. Vallet, A. Mac Kain, L. Miorin, E. Moreno, Z.Z.C. Naing, Y. Zhou, S. Peng, Y. Shi, Z. Zhang, W. Shen, I.T. Kirby, J.E. Melnyk, J.S. Chiorba, K. Lou, S.A. Dai, I. Barrio-Hernandez, D. Memon, C. Hernandez-Armenta, J. Lyu, C.J.P. Mathy, T. Perica, K.B. Pilla, S.J. Ganesan, D.J. Saltzberg, R. Rakesh, X. Liu, S.B. Rosenthal, L. Calviello, S. Venkataramanan, J. Liboy-Lugo, Y. Lin, X.-P. Huang, Y. Liu, S.A. Wankowicz, M. Bohn, M. Safari, F.S. Ugur, C. Koh, N.S. Savar, Q.D. Tran, D. Shengjiu, S.J. Fletcher, M.C. O'Neal, Y. Cai, J.C.J. Chang, D.J. Broadhurst, S. Klippsten, P.P. Sharp, N.A. Wenzell, D. Kuzuoglu-Ozturk, H.-Y. Wang, R. Trenker, J.M. Young, D.A. Caverio, J. Hiatt, T.L. Roth, U. Rathore, A. Subramanian, J. Noack, M. Hubert, R.M. Stroud, A.D. Frankel, O.S. Rosenberg, K.A. Verba, D.A. Agard, M. Ott, M. Emerman, N. Jura, M. von Zastrow, E. Verdini, A. Ashworth, O. Schwartz, C. d'Enfert, S. Mukherjee, M. Jacobson, H.S. Malik, D.G. Fujimori, T. Ideker, C.S. Craik, S.N. Floor, J.S. Fraser, J.D. Gross, A. Sali, B.L. Roth, D. Ruggiero, J. Taunton, T. Kortemme, P. Beltrao, M. Vignuzzi, A. García-Sastre, K.M. Shokat, B.K. Shoichet, N.J. Krogan, A SARS-CoV-2 protein interaction map reveals targets for drug repurposing, *Nature.* 583 (2020), <https://doi.org/10.1038/s41586-020-2286-9>.
- [18] S.F. Ahmed, A.A. Quadeer, M.R. McKay, Preliminary Identification of Potential Vaccine Targets for the COVID-19 Coronavirus (SARS-CoV-2) Based on SARS-CoV Immunological Studies, *Viruses.* 12 (2020), <https://doi.org/10.3390/v12030254>.
- [19] F.P. Polack, S.J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J.L. Perez, G. Pérez Marc, E.D. Moreira, C. Zerbini, R. Bailey, K.A. Swanson, S. Roychoudhury, K. Koury, P. Li, W.V. Kalina, D. Cooper, R.W. Frencik, L.L. Hammitt, Ö. Türeci, H. Nell, A. Schaefer, S. Ünal, D.B. Tresnan, S. Mather, P.R. Dormitzer, U. Şahin, K.U. Jansen, W.C. Gruber, Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine, *N. Engl. J. Med.* 383 (2020) 2603–2615, <https://doi.org/10.1056/NEJMoa2034577>.
- [20] L.R. Baden, H.M. El Sahly, B. Essink, K. Kotloff, S. Frey, R. Novak, D. Diemert, S. A. Spector, N. Roupheal, C.B. Creech, J. McGettigan, S. Khetan, N. Segall, J. Solis, A. Brosz, C. Fierro, H. Schwartz, K. Neuzil, L. Corey, P. Gilbert, H. Janes, D. Follmann, M. Marovich, J. Mascola, L. Polakowski, J. Ledgerwood, B.S. Graham, H. Bennett, R. Pajon, C. Knightly, B. Leav, W. Deng, H. Zhou, S. Han, M. Ivarsson, J. Miller, T. Zaks, Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine, *N. Engl. J. Med.* (2020), <https://doi.org/10.1056/NEJMoa2035389>.
- [21] M. Voysey, S.A.C. Clemens, S.A. Madhi, L.Y. Weckx, P.M. Folegatti, P.K. Aley, B. Angus, V.L. Baillie, S.L. Barnabas, Q.E. Bhorat, S. Bibi, C. Briner, P. Cicconi, A.M. Collins, R. Colijn-Jones, C.L. Cutland, T.C. Darton, K. Dheda, C.J.A. Duncan, K.R. W. Emary, K.J. Ewer, L. Fairlie, S.N. Faust, S. Feng, D.M. Ferreira, A. Finn, A.L. Goodman, C.M. Green, C.A. Green, P.T. Heath, C. Hill, H. Hill, I. Hirsch, S.H.C. Hodgson, A. Izu, S. Jackson, D. Jenkin, C.C.D. Joe, S. Kerridge, A. Koen, G.



- Kwatra, R. Lazarus, A.M. Lawrie, A. Lelliott, V. Libri, P.J. Lillie, R. Mallory, A.V.A. Mendes, E.P. Milan, A.M. Minassian, A. McGregor, H. Morrison, Y.F. Mujadidi, A. Nana, P.J. O'Reilly, S.D. Padayachee, A. Pittella, E. Plested, K.M. Pollock, M.N. Ramasamy, S. Rhead, A. V. Schwarzbald, N. Singh, A. Smith, R. Song, M.D. Snape, E. Sprinz, R.K. Sutherland, R. Tarrant, E.C. Thomson, M.E. Török, M. Toshner, D.P.J. Turner, J. Vekemans, T.L. Villafana, M.E.E. Watson, C.J. Williams, A.D. Douglas, A.V.S. Hill, T. Lambe, S.C. Gilbert, A.J. Pollard, O.C.V. T. Group, Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK, *Lancet*. (2020). [https://doi.org/10.1016/S0140-6736\(20\)32661-1](https://doi.org/10.1016/S0140-6736(20)32661-1).
- [22] I.X. Chen, K. Newcomer, K.E. Pauken, V.R. Juneja, K. Naxerova, M.W. Wu, M. Pinter, D.R. Sen, M. Singer, A.H. Sharpe, R.K. Jain, A bilateral tumor model identifies transcriptional programs associated with patient response to immune checkpoint blockade, *Proc. Natl. Acad. Sci.* 117 (2020) 23684–23694, <https://doi.org/10.1073/PNAS.2002806117>.
- [23] P.S. Porto, D. Anjos, N. Dábilla, S.G. da Fonseca, M. Souza, Immunoinformatic construction of an adenovirus-based modular vaccine platform and its application in the design of a SARS-CoV-2 vaccine, *Infect. Genet. Evol.* 85 (2020), <https://doi.org/10.1016/j.meegid.2020.104489> 104489.
- [24] M. Yarmarkovich, J.M. Warrington, A. Farrel, J.M. Maris, Identification of SARS-CoV-2 Vaccine Epitopes Predicted to Induce Long-Term Population-Scale Immunity, *Cell Reports Med.* 1 (2020), <https://doi.org/10.1016/j.xcrm.2020.100036> 100036.
- [25] M.I. Abdelmageed, A.H. Abdelmoneim, M.I. Mustafa, N.M. Elfadol, N.S. Murshed, S.W. Shantier, A.M. Makhawi, Design of a Multi-epitope-Based Peptide Vaccine against the E Protein of Human COVID-19: An Immunoinformatics Approach, *Biomed Res, Int*, 2020, <https://www.hindawi.com/journals/bmri/2020/2683286/>.
- [26] E. Ong, M.U. Wong, A. Huffman, Y. He, COVID-19 Coronavirus Vaccine Design Using Reverse Vaccinology and Machine Learning, *Front. Immunol.* 11 (2020), <https://doi.org/10.3389/fimmu.2020.01581>.
- [27] J.A. Plante, Y. Liu, J. Liu, H. Xia, B.A. Johnson, K.G. Lokugamage, X. Zhang, A.E. Muruato, J. Zou, C.R. Fontes-Garfias, D. Mirchandani, D. Scharton, J.P. Billello, Z. Ku, Z. An, B. Kalveram, A.N. Freiberg, V.D. Menachery, X. Xie, K.S. Plante, S.C. Weaver, P.-Y. Shi, Spike mutation D614G alters SARS-CoV-2 fitness, *Nature*. (2020), <https://doi.org/10.1038/s41586-020-2895-3>.
- [28] I.A. Doytchinova, D.R. Flower, Vaxijen: a server for prediction of protective antigens, tumour antigens and subunit vaccines, *BMC Bioinformatics.* 8 (2007) 4, <https://doi.org/10.1186/1471-2105-8-4>.
- [29] A. Naz, F. Shahid, T.T. Butt, F.M. Awan, A. Ali, A. Malik, Designing Multi-Epitope Vaccines to Combat Emerging Coronavirus Disease 2019 (COVID-19) by Employing Immuno-Informatics Approach, *Front. Immunol.* 11 (2020), <https://doi.org/10.3389/fimmu.2020.01663>.
- [30] E. Ong, H. Wang, M.U. Wong, M. Seetharaman, N. Valdez, Y. He, Vaxign-ML: supervised machine learning reverse vaccinology model for improved prediction of bacterial protective antigens, *Bioinformatics.* 36 (2020) 3185–3191, <https://doi.org/10.1093/bioinformatics/btaa119>.
- [31] B. Yang, S. Sayers, Z. Xiang, Y. He, Protegen: a web-based protective antigen database and analysis system, *Nucleic Acids Res.* 39 (2011) D1073–D1078, <https://doi.org/10.1093/nar/gkq944>.
- [32] C. Rollenhagen, M. Sörensen, K. Rizos, R. Hurvitz, D. Bumann, Antigen selection based on expression levels during infection facilitates vaccine development for an intracellular pathogen, *Proc. Natl. Acad. Sci. U. S. A.* 101 (2004) 8739 LP – 8744, <https://doi.org/10.1073/pnas.0401283101>.
- [33] Y. Dong, T. Dai, Y. Wei, L. Zhang, M. Zheng, F. Zhou, A systematic review of SARS-CoV-2 vaccine candidates, *Signal Transduct. Target. Ther.* 5 (2020), <https://doi.org/10.1038/s41392-020-00352-y>.
- [34] V.J. Costela-Ruiz, R. Illescas-Montes, J.M. Puerta-Puerta, C. Ruiz, L. Melguizo-Rodríguez, SARS-CoV-2 infection: The role of cytokines in COVID-19 disease, *Cytokine Growth Factor Rev.* 54 (2020), <https://doi.org/10.1016/j.cytogfr.2020.06.001>.
- [35] K.A. Overmyer, E. Shishkova, I.J. Miller, J. Balnis, N. Bernstein, T.M. Peters-clarke, J.G. Meyer, Q. Quan, K. Laura, E.A. Trujillo, Y. He, A. Chopra, H.C. Chieng, M.A. Judson, B. Paulson, D.R. Brademan, Y. Zhu, R. Lia, V. Linke, L.A. Drake, A.P. Adam, B.S. Schwartz, H.A. Singer, S. Swanson, D.F. Mosher, R. Stewart, J.J. Coon, A. Jaitovich, Large-scale Multi-omics Analysis of COVID-19 Severity, *Cell Syst.* (2020), <https://doi.org/10.1016/j.cels.2020.10.003>.
- [36] WHO, Clinical management of COVID-19, 2020. <https://doi.org/WHO/2019-nCoV/clinical/2020.5>.
- [37] U. Raudvere, L. Kolberg, I. Kuzmin, T. Arak, P. Adler, H. Peterson, J. Vilo, g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update), *Nucleic Acids Res.* 47 (2019) W191–W198, <https://doi.org/10.1093/nar/gkz369>.
- [38] M. Marsh, A. Pelchen-Matthews, *Endocytosis in Viral Replication, Traffic.* 1 (2000) 525–532, [10.1034/j.1600-0854.2000.010701.x](https://doi.org/10.1034/j.1600-0854.2000.010701.x).
- [39] N. Strbo, N. Yin, O. Stojadinovic, Innate and Adaptive Immune Responses in Wound Epithelialization, *Adv. Wound Care.* 3 (2013) 492–501, <https://doi.org/10.1089/wound.2012.0435>.
- [40] D. Szklarczyk, A.L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N.T. Doncheva, J.H. Morris, P. Bork, L.J. Jensen, C. von Mering, STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets, *Nucleic Acids Res.* 47 (2019) D607–D613, <https://doi.org/10.1093/nar/gky1131>.
- [41] Y.-D. Li, W.-Y. Chi, J.-H. Su, L. Ferrall, C.-F. Hung, T.-C. Wu, Coronavirus vaccine development: from SARS and MERS to COVID-19, *J. Biomed. Sci.* 27 (2020) 104, <https://doi.org/10.1186/s12929-020-00695-2>.
- [42] S. Hong, Z. Zhang, H. Liu, M. Tian, X. Zhu, Z. Zhang, W. Wang, X. Zhou, F. Zhang, Q. Ge, B. Zhu, H. Tang, Z. Hua, B. Hou, B Cells Are the Dominant Antigen-Presenting Cells that Activate Naive CD4+ T Cells upon Immunization with a Virus-Derived Nanoparticle Antigen, *Immunity.* 49 (2018), <https://doi.org/10.1016/j.immuni.2018.08.012>.
- [43] T. Kurosaki, K. Kometsani, W. Ise, Memory B cells, *Nat. Rev. Immunol.* 15 (2015), <https://doi.org/10.1038/nri3802>.
- [44] F. Long, R.H. Fong, S.K. Austin, Z. Chen, T. Klose, A. Fokine, Y. Liu, J. Porta, G. Sapparapu, W. Akahata, B.J. Doranz, J.E. Crowe, M.S. Diamond, M.G. Rossmann, Cryo-EM structures elucidate neutralizing mechanisms of anti-chikungunya human monoclonal antibodies with therapeutic activity, *Proc. Natl. Acad. Sci.* 112 (2015), <https://doi.org/10.1073/pnas.1515558112>.
- [45] T.A. Ahmad, A.E. Eweida, S.A. Sheweita, B-cell epitope mapping for the design of vaccines and effective diagnostics, *Trials Vaccinol.* 5 (2016), <https://doi.org/10.1016/j.trivac.2016.04.003>.
- [46] Y.C. Infante, A. Pupo, G. Rojas, A combinatorial mutagenesis approach for functional epitope mapping on phage-displayed target antigen, *MAbs.* 6 (2014), <https://doi.org/10.4161/mabs.28395>.
- [47] C. Lundegaard, O. Lund, M. Nielsen, Predictions versus high-throughput experiments in T-cell epitope discovery: competition or synergy?, *Expert Rev Vaccines.* 11 (2012), <https://doi.org/10.1586/erv.11.160>.
- [48] J.M.R. Parker, D. Guo, R.S. Hodges, 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 (1986), <https://doi.org/10.1021/bi00367a013>.
- [49] A. Rakib, S.A. Sami, M.A. Islam, S. Ahmed, F.B. Faiz, B.H. Khanam, K.K.S. Marma, M. Rahman, M.M.N. Uddin, F. Nainu, T. Bin Emran, J. Simal-Gandara, Epitope-Based Immunoinformatics Approach on Nucleocapsid Protein of Severe Acute Respiratory Syndrome-Coronavirus-2, *Molecules.* 25 (2020), <https://doi.org/10.3390/molecules25215088>.
- [50] P.A. Karplus, G.E. Schulz, Prediction of chain flexibility in proteins, *Naturwissenschaften.* 72 (1985), <https://doi.org/10.1007/BF01195768>.
- [51] P.Y. Chou, G.D. Fasman, Conformational parameters for amino acids in helical,  $\beta$ -sheet, and random coil regions calculated from proteins, *Biochemistry.* 13 (1974), <https://doi.org/10.1021/bi00699a001>.
- [52] M.S. Rahman, M.N. Hoque, M.R. Islam, S. Akter, A.S.M. Rubayet-UI-Alam, M.A. Siddique, O. Saha, M.M. Rahaman, M. Sultana, K.A. Crandall, M.A. Hossain, Epitope-based chimeric peptide vaccine design against S, M and E proteins of SARS-CoV-2 etiologic agent of global pandemic COVID-19: an in silico approach, *PeerJ.* 8 (2020), <https://doi.org/10.7717/peerj.9572>.
- [53] M. Källberg, G. Margaryan, S. Wang, J. Ma, J. Xu, RaptorX server: A Resource for Template-Based Protein Structure Modeling, in (2014), [https://doi.org/10.1007/978-1-4939-0366-5\\_2](https://doi.org/10.1007/978-1-4939-0366-5_2).
- [54] M.J. Blythe, D.R. Flower, Benchmarking B cell epitope prediction: Underperformance of existing methods, *Protein Sci.* 14 (2009), <https://doi.org/10.1110/ps.041059505>.
- [55] S. Lin, C.-W. Cheng, E. Su, Prediction of B-cell epitopes using evolutionary information and propensity scales, *BMC Bioinformatics.* 14 (2013), <https://doi.org/10.1186/1471-2105-14-52-510>.
- [56] S. Saha, G.P.S. Raghava, Prediction of continuous B-cell epitopes in an antigen using recurrent neural network, *Proteins Struct. Funct. Bioinforma.* 65 (2006) 40–48, <https://doi.org/10.1002/prot.21078>.
- [57] Y. El-Manzalawy, D. Dobbs, V. Honavar, Predicting linear B-cell epitopes using string kernels, *J. Mol. Recognit.* 21 (2008), <https://doi.org/10.1002/jmr.893>.
- [58] H. Singh, H.R. Ansari, G.P.S. Raghava, Improved Method for Linear B-Cell Epitope Prediction Using Antigen's Primary Sequence, *PLoS One.* 8 (2013), <https://doi.org/10.1371/journal.pone.0062216>.
- [59] W. Shen, Y. Cao, L. Cha, X. Zhang, X. Ying, W. Zhang, K. Ge, W. Li, L. Zhong, Predicting linear B-cell epitopes using amino acid anchoring pair composition, *BioData Min.* 8 (2015), <https://doi.org/10.1186/s13040-015-0047-3>.
- [60] M.C. Jespersen, B. Peters, M. Nielsen, P. Marcattili, BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes, *Nucleic Acids Res.* 45 (2017) W24–W29, <https://doi.org/10.1093/nar/gkx346>.
- [61] V. Saravanan, N. Gautham, Harnessing Computational Biology for Exact Linear B-Cell Epitope Prediction: A Novel Amino Acid Composition-Based Feature Descriptor, *Omi. A J. Integr. Biol.* 19 (2015), <https://doi.org/10.1089/omi.2015.0095>.
- [62] A.J.P. Alix, Predictive estimation of protein linear epitopes by using the program PEOPLE, *Vaccine.* 18 (1999), [https://doi.org/10.1016/S0264-410X\(99\)00329-1](https://doi.org/10.1016/S0264-410X(99)00329-1).
- [63] J.L. Pellequer, E. Westhof, PREDITOP: A program for antigenicity prediction, *J. Mol. Graph.* 11 (1993), [https://doi.org/10.1016/0263-7855\(93\)80074-2](https://doi.org/10.1016/0263-7855(93)80074-2).
- [64] A. Singh, M. Thakur, L.K. Sharma, K. Chandra, Designing a multi-epitope peptide based vaccine against SARS-CoV-2, *Sci. Rep.* 10 (2020), <https://doi.org/10.1038/s41598-020-73371-y>.
- [65] M. Waqas, A. Haider, M. Sufyan, S. Siraj, S.A. Sehgal, Determine the Potential Epitope Based Peptide Vaccine Against Novel SARS-CoV-2 Targeting Structural Proteins Using Immunoinformatics Approaches, *Front. Mol. Biosci.* 7 (2020), <https://doi.org/10.3389/fmolb.2020.00227>.
- [66] B.K. Das, D. Chakraborty, Epitope-Based Potential Vaccine Candidate for Humoral and Cell-Mediated Immunity to Combat Severe Acute Respiratory

- Syndrome Coronavirus 2 Pandemic, *J. Phys. Chem. Lett.* 11 (2020), <https://doi.org/10.1021/acs.jpcclett.0c02846>.
- [67] H.M. Rehman, M.U. Mirza, M.A. Ahmad, M. Saleem, M. Froeyen, S. Ahmad, R. Gul, H.A. Alghamdi, M.S. Aslam, M. Sajjad, M.A. Bhinder, A Putative Prophylactic Solution for COVID-19: Development of Novel Multi-epitope Vaccine Candidate against SARS-CoV-2 by Comprehensive Immunoinformatic and Molecular Modelling Approach, *Biology (Basel)*. 9 (2020), <https://doi.org/10.3390/biology9090296>.
- [68] R. Jakhar, S.K. Gakhar, An Immunoinformatics Study to Predict Epitopes in the Envelope Protein of SARS-CoV-2, *Can. J. Infect. Dis. Med. Microbiol.* 2020 (2020), <https://doi.org/10.1155/2020/7079356>.
- [69] R. Dong, Z. Chu, F. Yu, Y. Zha, Contriving Multi-Epitope Subunit of Vaccine for COVID-19: Immunoinformatics Approaches, *Front. Immunol.* 11 (2020), <https://doi.org/10.3389/fimmu.2020.01784>.
- [70] M.R.N. Akhand, K.F. Azim, S.F. Hoque, M.A. Moli, B. Das Joy, H. Akter, I.K. Afif, N. Ahmed, M. Hasan, Genome based evolutionary lineage of SARS-CoV-2 towards the development of novel chimeric vaccine, *Infect. Genet. Evol.* 85 (2020), <https://doi.org/10.1016/j.meegid.2020.104517>.
- [71] M. Enayatkhani, M. Hasaniazad, S. Faezi, H. Gouklani, P. Davoodian, N. Ahmadi, M.A. Einakian, A. Karmostaji, K. Ahmadi, Reverse vaccinology approach to design a novel multi-epitope vaccine candidate against COVID-19: an *in silico* study, *J. Biomol. Struct. Dyn.* (2020), <https://doi.org/10.1080/07391102.2020.1756411>.
- [72] W. Li, L. Li, T. Sun, Y. He, G. Liu, Z. Xiao, Y. Fan, J. Zhang, Spike protein-based epitopes predicted against SARS-CoV-2 through literature mining, *Med. Nov. Technol. Devices*. 8 (2020), <https://doi.org/10.1016/j.medntd.2020.100048>.
- [73] S.O. Abd Albagi, M.Y. Al-Nour, M. Elhag, A. Tageldein Idris Abdelihalim, E. Mosa Haroun, M.E. Adam Essa, M. Abubaker, H. Deka, A. Ghosh, M.A. Hassan, A multiple peptides vaccine against COVID-19 designed from the nucleocapsid phosphoprotein (N) and Spike Glycoprotein (S) via the immunoinformatics approach, *Informatics Med. Unlocked*. 21 (2020), <https://doi.org/10.1016/j.imu.2020.100476>.
- [74] A. Safavi, A. Kefayat, E. Mahdevar, A. Abiri, F. Ghahremani, Exploring the out of sight antigens of SARS-CoV-2 to design a candidate multi-epitope vaccine by utilizing immunoinformatics approaches, *Vaccine*. 38 (2020), <https://doi.org/10.1016/j.vaccine.2020.10.016>.
- [75] J.R. Lon, Y. Bai, B. Zhong, F. Cai, H. Du, Prediction and evolution of B cell epitopes of surface protein in SARS-CoV-2, *Virology*. 17 (2020), <https://doi.org/10.1186/s12985-020-01437-4>.
- [76] Y. Dai, H. Chen, S. Zhuang, X. Feng, Y. Fang, H. Tang, R. Dai, L. Tang, J. Liu, T. Ma, G. Zhong, Immunodominant regions prediction of nucleocapsid protein for SARS-CoV-2 early diagnosis: a bioinformatics and immunoinformatics study, *Pathog. Glob. Health*. 114 (2020), <https://doi.org/10.1080/20477724.2020.1838190>.
- [77] U. Kulkarni-Kale, S. Bhosle, A.S. Kolaskar, CEP: a conformational epitope prediction server, *Nucleic Acids Res.* 33 (2005), <https://doi.org/10.1093/nar/gki460>.
- [78] P. Haste Andersen, M. Nielsen, O. Lund, Prediction of residues in discontinuous B-cell epitopes using protein 3D structures, *Protein Sci.* 15 (2006), <https://doi.org/10.1110/ps.062405906>.
- [79] J. Ponomarenko, H.-H. Bui, W. Li, N. Füsseder, P.E. Bourne, A. Sette, B. Peters, ElliPro: a new structure-based tool for the prediction of antibody epitopes, *BMC Bioinformatics*. 9 (2008), <https://doi.org/10.1186/1471-2105-9-514>.
- [80] J.V. Kringleum, C. Lundegaard, O. Lund, M. Nielsen, Reliable B Cell Epitope Predictions: Impacts of Method Development and Improved Benchmarking, *PLoS Comput. Biol.* 8 (2012), <https://doi.org/10.1371/journal.pcbi.1002829>.
- [81] J. Huang, A. Gutteridge, W. Honda, M. Kanehisa, MIMOX: a web tool for phage display based epitope mapping, *BMC Bioinformatics*. 7 (2006), <https://doi.org/10.1186/1471-2105-7-451>.
- [82] I. Mayrose, O. Penn, E. Erez, N.D. Rubinstein, T. Shlomi, N.T. Freund, E.M. Bublil, E. Ruppim, R. Sharan, J.M. Gershoni, E. Martz, T. Pupko, Pepitope: epitope mapping from affinity-selected peptides, *Bioinformatics*. 23 (2007), <https://doi.org/10.1093/bioinformatics/btm493>.
- [83] S.S. Negi, B. Braun, Automated Detection of Conformational Epitopes Using Phage Display Peptide Sequences, *Bioinform. Biol. Insights*. 3 (2009), <https://doi.org/10.4137/BBI.S2745>.
- [84] W. Chen, P. Sun, Y. Lu, W.W. Guo, Y. Huang, Z. Ma, MimoPro: a more efficient Web-based tool for epitope prediction using phage display libraries, *BMC Bioinformatics*. 12 (2011), <https://doi.org/10.1186/1471-2105-12-199>.
- [85] W. Chen, W.W. Guo, Y. Huang, Z. Ma, PepMapper: A Collaborative Web Tool for Mapping Epitopes from Affinity-Selected Peptides, *PLoS One*. 7 (2012), <https://doi.org/10.1371/journal.pone.0037869>.
- [86] S. Liang, D. Zheng, D.M. Standley, B. Yao, M. Zacharias, C. Zhang, EPSVR and EPMeta: prediction of antigenic epitopes using support vector regression and multiple server results, *BMC Bioinformatics*. 11 (2010), <https://doi.org/10.1186/1471-2105-11-381>.
- [87] M.J. Sweredoski, P. Baldi, PEPITO: improved discontinuous B-cell epitope prediction using multiple distance thresholds and half sphere exposure, *Bioinformatics*. 24 (2008), <https://doi.org/10.1093/bioinformatics/btn199>.
- [88] J. Sun, D. Wu, T. Xu, X. Wang, X. Xu, L. Tao, Y.X. Li, Z.W. Cao, SEPPA: a computational server for spatial epitope prediction of protein antigens, *Nucleic Acids Res.* 37 (2009), <https://doi.org/10.1093/nar/gkp417>.
- [89] N.D. Rubinstein, I. Mayrose, E. Martz, T. Pupko, Epitopia: a web-server for predicting B-cell epitopes, *BMC Bioinformatics*. 10 (2009), <https://doi.org/10.1186/1471-2105-10-287>.
- [90] S. Liang, D. Zheng, C. Zhang, M. Zacharias, Prediction of antigenic epitopes on protein surfaces by consensus scoring, *BMC Bioinformatics*. 10 (2009), <https://doi.org/10.1186/1471-2105-10-302>.
- [91] A.K. Gupta, M.S. Khan, S. Choudhury, A. Mukhopadhyay, Sakshi, A. Rastogi, A. Thakur, P. Kumari, M. Kaur, Shalu, C. Saini, V. Sapehia, Barkha, P.K. Patel, K.T. Bhamare, M. Kumar, CoronaVR: A Computational Resource and Analysis of Epitopes and Therapeutics for Severe Acute Respiratory Syndrome Coronavirus-2, *Front. Microbiol.* 11 (2020), <https://doi.org/10.3389/fmicb.2020.01858>.
- [92] N. Khairkhan, M.R. Aghasadeghi, A. Namvar, A. Bolhassani, Design of novel multi-epitope constructs-based peptide vaccine against the structural S, N and M proteins of human COVID-19 using immunoinformatics analysis, *PLoS One*. 15 (2020), <https://doi.org/10.1371/journal.pone.0240577>.
- [93] A.C. Karlsson, M. Humbert, M. Buggert, The known unknowns of T cell immunity to COVID-19, *Sci. Immunol.* 5 (2020), <https://doi.org/10.1126/sciimmunol.abe8063>.
- [94] D. Mathew, J.R. Giles, A.E. Baxter, D.A. Oldridge, A.R. Greenplate, J.E. Wu, C. Alanio, L. Kuri-Cervantes, M.B. Pampena, K. D'Andrea, S. Manne, Z. Chen, Y.J. Huang, J.P. Reilly, A.R. Weisman, C.A.G. Ittner, O. Kuthuru, J. Dougherty, K. Nzingha, N. Han, J. Kim, A. Pattekar, E.C. Goodwin, E.M. Anderson, M.E. Weirick, S. Gouma, C.P. Arevalo, M.J. Bolton, F. Chen, S.F. Lacey, H. Ramage, S. Cherry, S.E. Hensley, S.A. Apostolidis, A.C. Huang, L.A. Vella, M.R. Betts, N.J. Meyer, E.J. Wherry, Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications, *Science* 80- (2020) ). 369, <https://doi.org/10.1126/science.abc8511>.
- [95] P. de Candia, F. Prattichizzo, S. Garavelli, G. Matarese, T. Cells, Warriors of SARS-CoV-2 Infection, *Trends Immunol.* 42 (2021), <https://doi.org/10.1016/j.it.2020.11.002>.
- [96] J.A. Juno, H.-X. Tan, W.S. Lee, A. Reynaldi, H.G. Kelly, K. Wragg, R. Esterbauer, H.E. Kent, C.J. Batten, F.L. Mordant, N.A. Gherardin, P. Pymm, M.H. Dietrich, N. E. Scott, W.-H. Tham, D.I. Godfrey, K. Subbarao, M.P. Davenport, S.J. Kent, A.K. Wheatley, Humoral and circulating follicular helper T cell responses in recovered patients with COVID-19, *Nat. Med.* 26 (2020), <https://doi.org/10.1038/s41591-020-0995-0>.
- [97] T. Sekine, A. Perez-Potti, O. Rivera-Ballesteros, K. Strålin, J.-B. Gorin, A. Olsson, S. Llewellyn-Lacey, H. Kamal, G. Bogdanovic, S. Muschiol, D.J. Wullmann, T. Kammann, J. Enggård, T. Parrot, E. Folkesson, O. Rooyackers, L.I. Eriksson, J.-I. Henter, A. Sönnnerborg, T. Allander, J. Albert, M. Nielsen, J. Klingström, S. Gredmark-Russ, N.K. Björkström, J.K. Sandberg, D.A. Price, H.-G. Ljunggren, S. Aleman, M. Buggert, M. Akber, L. Berglin, H. Bergsten, S. Brighenti, D. Brownlie, M. Butrym, B. Chambers, P. Chen, M.C. Jeannin, J. Grip, A.C. Gomez, L. Dillner, I.D. Lozano, M. Dzidic, M.F. Tullberg, A. Färnert, H. Glans, A. Haroun-Izquierdo, E. Henriksson, L. Hertwig, S. Kalsum, E. Kokkinou, E. Kvedaraitte, M. Loreti, M. Lourda, K. Maleki, K.-J. Malmberg, N. Marquardt, C. Maucourant, J. Michaelsson, J. Mjösberg, K. Moll, J. Muva, J. Mårtensson, P. Naucière, A. Norrby-Teglund, L.P. Medina, B. Persson, L. Radler, E. Ringqvist, J.T. Sandberg, E. Söhlberg, T. Soini, M. Svensson, J. Tynell, R. Varnaite, A. Von Kries, C. Unge, Robust T Cell Immunity in Convalescent Individuals with Asymptomatic or Mild COVID-19, *Cell*. 183 (2020), <https://doi.org/10.1016/j.cell.2020.08.017>.
- [98] Y. Peng, A.J. Mentzer, G. Liu, X. Yao, Z. Yin, D. Dong, W. Dejnirattisai, T. Rostron, P. Supasa, C. Liu, C. López-Camacho, J. Slon-Compos, Y. Zhao, D.I. Stuart, G.C. Paesen, J.M. Grimes, A.A. Antson, O.W. Bayfield, D.E.D.P. Hawkins, D.-S. Ker, B. Wang, L. Turtle, K. Subramaniam, P. Thomson, P. Zhang, C. Dold, J. Ratcliff, P. Simmonds, T. de Silva, P. Sopp, D. Wellington, U. Rajapaksa, Y.-L. Chen, M. Salio, G. Napolitani, W. Paes, P. Borrow, B.M. Kessler, J.W. Fry, N.F. Schwabe, M.G. Semple, J.K. Baillie, S.C. Moore, P.J.M. Openshaw, M.A. Ansari, S. Dunachie, E. Barnes, J. Frater, G. Kerr, P. Goulder, T. Lockett, R. Levin, Y. Zhang, R. Jing, L.-P. Ho, R.J. Cornall, C.P. Conlon, P. Klenerman, G.R. Screaton, J. Mongkolsapaya, A. McMichael, J.C. Knight, G. Ogg, T. Dong, Broad and strong memory CD4+ and CD8+ T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19, *Nat. Immunol.* 21 (2020), <https://doi.org/10.1038/s41590-020-0782-6>.
- [99] J.E. Smith-Garvin, G.A. Koretzky, M.S. Jordan, T Cell Activation, *Annu. Rev. Immunol.* 27 (2009), <https://doi.org/10.1146/annurev.immunol.021908.132706>.
- [100] R. Casalegno-Garduño, A. Schmitt, J. Yao, X. Wang, X. Xu, M. Freund, M. Schmitt, Multimer technologies for detection and adoptive transfer of antigen-specific T cells, *Cancer Immunol. Immunother.* 59 (2010), <https://doi.org/10.1007/s00262-009-0778-4>.
- [101] R. Aggarwal, R. Shukla, S. Jameel, S. Agrawal, P. Puri, V.K. Gupta, A.P. Patil, S. Naik, T-cell epitope mapping of ORF2 and ORF3 proteins of human hepatitis E virus, *J. Viral Hepat.* 14 (2007), <https://doi.org/10.1111/j.1365-2893.2006.00796.x>.
- [102] P.V. Lehmann, M. Suwansaard, T. Zhang, D.R. Roen, G.A. Kirchenbaum, A.Y. Karulin, A. Lehmann, P.A. Reche, Comprehensive Evaluation of the Expressed CD8+ T Cell Epitope Space Using High-Throughput Epitope Mapping, *Front. Immunol.* 10 (2019), <https://doi.org/10.3389/fimmu.2019.00655>.
- [103] C. Watts, The endosome-lysosome pathway and information generation in the immune system, *Biochim. Biophys. Acta - Proteins Proteomics*. 1824 (2012), <https://doi.org/10.1016/j.bbapap.2011.07.006>.
- [104] L.J. Stern, D.C. Wiley, Antigenic peptide binding by class I and class II histocompatibility proteins, *Structure*. 2 (1994), [https://doi.org/10.1016/S0969-2126\(00\)00026-5](https://doi.org/10.1016/S0969-2126(00)00026-5).
- [105] J.L. Sanchez-Trincado, M. Gomez-Perosanz, P.A. Reche, Fundamentals and Methods for T- and B-Cell Epitope Prediction, *J. Immunol. Res.* 2017 (2017), <https://doi.org/10.1155/2017/2680160>.



- [106] B. Reynisson, C. Barra, S. Kaabinejadian, W.H. Hildebrand, B. Peters, M. Nielsen, Improved Prediction of MHC II Antigen Presentation through Integration and Motif Deconvolution of Mass Spectrometry MHC Eluted Ligand Data, *J. Proteome Res.* 19 (2020) 2304–2315, <https://doi.org/10.1021/acs.jproteome.9b00874>.
- [107] B. Alvarez, B. Reynisson, C. Barra, S. Buus, N. Ternette, T. Connelley, M. Andreatta, M. Nielsen, NNAlign\_MA; MHC Peptidome Deconvolution for Accurate MHC Binding Motif Characterization and Improved T-cell Epitope Predictions, *Mol. Cell. Proteomics*. 18 (2019), <https://doi.org/10.1074/mcp.TIR119.001658>.
- [108] J. Racle, J. Michaux, G.A. Rockinger, M. Arnaud, S. Bobisse, C. Chong, P. Guillaume, C. Coukos, A. Harari, C. Jandus, M. Bassani-Sternberg, D. Gfeller, Robust prediction of HLA class II epitopes by deep motif deconvolution of immunopeptidomes, *Nat. Biotechnol.* 37 (2019), <https://doi.org/10.1038/s41587-019-0289-6>.
- [109] A. Grifoni, J. Sidney, Y. Zhang, R.H. Scheuermann, B. Peters, A. Sette, A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2, *Cell Host Microbe*. 27 (2020), <https://doi.org/10.1016/j.chom.2020.03.002>.
- [110] S. Mukherjee, D. Tworowski, R. Detroja, S.B. Mukherjee, M. Frenkel-Morgenstern, Immunoinformatics and Structural Analysis for Identification of Immunodominant Epitopes in SARS-CoV-2 as Potential Vaccine Targets, *Vaccines*. 8 (2020), <https://doi.org/10.3390/vaccines8020290>.
- [111] T. Stranzl, M.V. Larsen, C. Lundegaard, M. Nielsen, NetCTLpan: pan-specific MHC class I pathway epitope predictions, *Immunogenetics*. 62 (2010), <https://doi.org/10.1007/s00251-010-0441-4>.
- [112] V.S. Ayyagari, V. T. C., A.P. K., K. Srirama, Design of a multi-epitope-based vaccine targeting M-protein of SARS-CoV2: an immunoinformatics approach, *J. Biomol. Struct. Dyn.* (2020). <https://doi.org/10.1080/07391102.2020.1850357>.
- [113] S. Mishra, Designing of cytotoxic and helper T cell epitope map provides insights into the highly contagious nature of the pandemic novel coronavirus SARS-CoV-2, *R. Soc. Open Sci.* 7 (2020), <https://doi.org/10.1098/rsos.201141>.
- [114] S. Uebel, R. Tampé, Specificity of the proteasome and the TAP transporter, *Curr. Opin. Immunol.* 11 (1999), [https://doi.org/10.1016/S0952-7915\(99\)80034-X](https://doi.org/10.1016/S0952-7915(99)80034-X).
- [115] H.-G. Holzthütter, C. Frömmel, P.-M. Kloetzel, A theoretical approach towards the identification of cleavage-determining amino acid motifs of the 20s proteasome 1 Edited by R. Huber, *J. Mol. Biol.* 286 (1999). <https://doi.org/10.1006/jmbi.1998.2530>.
- [116] M. Bhasin, G.P.S. Raghava, Pcleavage: an SVM based method for prediction of constitutive proteasome and immunoproteasome cleavage sites in antigenic sequences, *Nucleic Acids Res.* 33 (2005), <https://doi.org/10.1093/nar/gki587>.
- [117] M. Gomez-Perosanz, A. Ras-Carmona, P.A. Reche, PCPS: A Web Server to Predict Proteasomal Cleavage Sites, in (2020), [https://doi.org/10.1007/978-1-0716-0389-5\\_23](https://doi.org/10.1007/978-1-0716-0389-5_23).
- [118] M. Nielsen, C. Lundegaard, O. Lund, C. Keşmir, The role of the proteasome in generating cytotoxic T-cell epitopes: insights obtained from improved predictions of proteasomal cleavage, *Immunogenetics*. 57 (2005), <https://doi.org/10.1007/s00251-005-0781-7>.
- [119] T.N. Schumacher, D.V. Kantesaria, M.T. Heemels, P.G. Ashton-Rickardt, J.C. Shepherd, K. Fruh, Y. Yang, P.A. Peterson, S. Tonegawa, H.L. Ploegh, Peptide length and sequence specificity of the mouse TAP1/TAP2 translocator, *J. Exp. Med.* 179 (1994), <https://doi.org/10.1084/jem.179.2.533>.
- [120] P.M. van Endert, D. Riganelli, G. Greco, K. Fleischhauer, J. Sidney, A. Sette, J.F. Bach, The peptide-binding motif for the human transporter associated with antigen processing, *J. Exp. Med.* 182 (1995), <https://doi.org/10.1084/jem.182.6.1883>.
- [121] P. Reche, J.-P. Glutting, H. Zhang, E. Reinherz, Enhancement to the RANKPEP resource for the prediction of peptide binding to MHC molecules using profiles, *Immunogenetics*. 56 (2004), <https://doi.org/10.1007/s00251-004-0709-7>.
- [122] P.A. Reche, J.-P. Glutting, E.L. Reinherz, Prediction of MHC class I binding peptides using profile motifs, *Hum. Immunol.* 63 (2002), [https://doi.org/10.1016/S0198-8859\(02\)00432-9](https://doi.org/10.1016/S0198-8859(02)00432-9).
- [123] X.M. Shao, R. Bhattacharya, J. Huang, I.K.A. Sivakumar, C. Tokheim, L. Zheng, D. Hirsch, B. Kaminow, A. Omdahl, M. Bonsack, A.B. Riemer, V.E. Velculescu, V. Anagnostou, K.A. Pagel, R. Karchin, High-Throughput Prediction of MHC Class I and II Neoantigens with MHCnuggets, *Cancer Immunol. Res.* 8 (2020), <https://doi.org/10.1158/2326-6066.CIR-19-0464>.
- [124] B. Reynisson, B. Alvarez, S. Paul, B. Peters, M. Nielsen, NetMHCIpan-4.1 and NetMHCIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data, *Nucleic Acids Res.* 48 (2020), <https://doi.org/10.1093/nar/gkaa379>.
- [125] F.F. Gonzalez-Galarza, A. McCabe, E.J.M. dos Santos, J. Jones, L. Takeshita, N.D. Ortega-Rivera, G.M. Del Cid-Pavon, K. Ramsbottom, G. Ghattaoraya, A. Alfrevic, D. Middleton, A.R. Jones, Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools, *Nucleic Acids Res.* (2019), <https://doi.org/10.1093/nar/gkz1029>.
- [126] K.M. Campbell, G. Steiner, D.K. Wells, A. Ribas, A. Kalbasi, Prioritization of SARS-CoV-2 epitopes using a pan-HLA and global population inference approach, *BioRxiv*. (2020).
- [127] M. Yarmarkovich, A. Farrel, A. Sison, M. di Marco, P. Raman, J.L. Parris, D. Monos, H. Lee, S. Stevanovic, J.M. Maris, Immunogenicity and Immune Silence in Human Cancer, *Front. Immunol.* 11 (2020), <https://doi.org/10.3389/fimmu.2020.00069>.
- [128] G. Liu, B. Carter, D.K. Gifford, Predicted Cellular Immunity Population Coverage Gaps for SARS-CoV-2 Subunit Vaccines and Their Augmentation by Compact Peptide Sets, *Cell Syst.* (2020), <https://doi.org/10.1016/j.cels.2020.11.010>.
- [129] I. Springer, H. Besser, N. Tickotsky-Moskovitz, S. Dvorkin, Y. Louzoun, Prediction of Specific TCR-Peptide Binding From Large Dictionaries of TCR-Peptide Pairs, *Front. Immunol.* 11 (2020), <https://doi.org/10.3389/fimmu.2020.01803>.
- [130] S.K. Dhanda, P. Vir, G.P.S. Raghava, Designing of interferon-gamma inducing MHC class-II binders, *Biol. Direct*. 8 (2013), <https://doi.org/10.1186/1745-6150-8-30>.
- [131] E. Behmard, B. Soleymani, A. Najafi, E. Barzegari, Immunoinformatic design of a COVID-19 subunit vaccine using entire structural immunogenic epitopes of SARS-CoV-2, *Sci. Rep.* 10 (2020), <https://doi.org/10.1038/s41598-020-77547-4>.
- [132] J. He, F. Huang, J. Zhang, Q. Chen, Z. Zheng, Q. Zhou, D. Chen, J. Li, J. Chen, Vaccine design based on 16 epitopes of SARS-CoV-2 spike protein, *J. Med. Virol.* (2020), <https://doi.org/10.1002/jmv.26596>.
- [133] A. Samad, F. Ahammad, Z. Nain, R. Alam, R.R. Imon, M. Hasan, M.S. Rahman, Designing a multi-epitope vaccine against SARS-CoV-2: an immunoinformatics approach, *J. Biomol. Struct. Dyn.* (2020), <https://doi.org/10.1080/07391102.2020.1792347>.
- [134] I. Ahammad, S.S. Lira, Designing a novel mRNA vaccine against SARS-CoV-2: An immunoinformatics approach, *Int. J. Biol. Macromol.* 162 (2020), <https://doi.org/10.1016/j.ijbiomac.2020.06.213>.
- [135] M.S. Sohail, S.F. Ahmed, A.A. Quadeer, M.R. McKay, In silico T cell epitope identification for SARS-CoV-2: Progress and perspectives, *Adv. Drug Deliv. Rev.* 171 (2021) 29–47, <https://doi.org/10.1016/j.addr.2021.01.007>.
- [136] S. Mei, F. Li, A. Leier, T.T. Marquez-Lago, K. Giam, N.P. Croft, T. Akutsu, A.I. Smith, J. Li, J. Rossjohn, A.W. Purcell, J. Song, A comprehensive review and performance evaluation of bioinformatics tools for HLA class I peptide-binding prediction, *Brief. Bioinform.* 21 (2020), <https://doi.org/10.1093/bib/bb2051>.
- [137] M. Prachar, S. Justesen, D.B. Steen-Jensen, S. Thorgrimsen, E. Jurgons, O. Winther, F.O. Bagger, Identification and validation of 174 COVID-19 vaccine candidate epitopes reveals low performance of common epitope prediction tools, *Sci. Rep.* 10 (2020), <https://doi.org/10.1038/s41598-020-77466-4>.
- [138] S.E.C. Caolili, Benchmarking B-Cell Epitope Prediction for the Design of Peptide-Based Vaccines: Problems and Prospects, *J. Biomed. Biotechnol.* 2010 (2010), <https://doi.org/10.1155/2010/910524>.
- [139] L. van Dorp, D. Richard, C.C.S. Tan, L.P. Shaw, M. Acman, F. Balloux, No evidence for increased transmissibility from recurrent mutations in SARS-CoV-2, *Nat. Commun.* 11 (2020), <https://doi.org/10.1038/s41467-020-19818-2>.
- [140] T. Koyama, D. Platt, L. Parida, Variant analysis of SARS-CoV-2 genomes, *Bull. World Health Organ.* 98 (2020), <https://doi.org/10.2471/BLT.20.253591>.
- [141] H.-H. Bui, J. Sidney, W. Li, N. Fusseder, A. Sette, Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines, *BMC Bioinformatics*. 8 (2007) 361, <https://doi.org/10.1186/1471-2105-8-361>.
- [142] S. Ismail, S. Ahmad, S.S. Azam, Immunoinformatics characterization of SARS-CoV-2 spike glycoprotein for prioritization of epitope based multivalent peptide vaccine, *J. Mol. Liq.* 314 (2020), <https://doi.org/10.1016/j.molliq.2020.113612>.
- [143] J. Wu, W. Chen, J. Zhou, W. Zhao, Y. Sun, H. Zhu, P. Yao, S. Chen, J. Jiang, Z. Zhou, COVIEdb: A Database for Potential Immune Epitopes of Coronaviruses (2020).
- [144] C. Zhang, G. Maruggi, H. Shan, J. Li, Advances in mRNA Vaccines for Infectious Diseases, *Front. Immunol.* 10 (2019) 594, <https://doi.org/10.3389/fimmu.2019.00594>.
- [145] S.J. Rhodes, J. Guedj, H.A. Fletcher, T. Lindenstrøm, T.J. Scriba, T.G. Evans, G.M. Knight, R.G. White, Using vaccine Immunostimulation/Immunodynamic modelling methods to inform vaccine dose decision-making, *Npj Vaccines*. 3 (2018) 36, <https://doi.org/10.1038/s41541-018-0075-3>.
- [146] S.J. Rhodes, G.M. Knight, D.E. Kirschner, R.G. White, T.G. Evans, Dose finding for new vaccines: The role for immunostimulation/immunodynamic modelling, *J. Theor. Biol.* 465 (2019) 51–55, <https://doi.org/10.1016/j.jtbi.2019.01.017>.
- [147] L.V. Brown, E.A. Gaffney, J. Wagg, M.C. Coles, Applications of mechanistic modelling to clinical and experimental immunology: an emerging technology to accelerate immunotherapeutic discovery and development, *Clin. Exp. Immunol.* 193 (2018) 284–292, <https://doi.org/10.1111/cei.13182>.
- [148] F. Pappalardo, G. Russo, F.M. Tshinanu, M. Viceconti, In silico clinical trials: concepts and early adoptions, *Brief. Bioinform.* 20 (2019) 1699–1708, <https://doi.org/10.1093/bib/bby043>.
- [149] A. Handel, N.L. La Gruta, P.G. Thomas, Simulation modelling for immunologists, *Nat. Rev. Immunol.* 20 (2020) 186–195, <https://doi.org/10.1038/s41577-019-0235-3>.
- [150] P. Charoentong, M. Angelova, M. Efreanova, R. Gallasch, H. Hackl, J. Galon, Z. Trajanoski, Bioinformatics for cancer immunology and immunotherapy, *Cancer Immunol. Immunother.* 61 (2012) 1885–1903, <https://doi.org/10.1007/s00262-012-1354-x>.
- [151] H.M. Dobrovolsky, M.B. Reddy, M.A. Kamal, C.R. Rayner, C.A.A. Beauchemin, Assessing Mathematical Models of Influenza Infections Using Features of the

- Immune Response e57088 e57088 *PLoS One*. 8 (2013), <https://doi.org/10.1371/journal.pone.0057088>.
- [152] V. Narang, J. Decraene, S.-Y. Wong, B.S. Aiswarya, A.R. Wasem, S.R. Leong, A. Gouaillard, Systems immunology: a survey of modeling formalisms, applications and simulation tools, *Immunol. Res.* 53 (2012) 251–265, <https://doi.org/10.1007/s12026-012-8305-7>.
- [153] F. Chiacchio, M. Pennisi, G. Russo, S. Motta, F. Pappalardo, Agent-Based Modeling of the Immune System: NetLogo, a Promising Framework, *Biomed Res. Int.* 2014 (2014) 1–6, <https://doi.org/10.1155/2014/907171>.
- [154] P.S. Kim, D. Levy, P.P. Lee, Modeling and Simulation of the Immune System as a Self-Regulating Network, in (2009) 79–109, [https://doi.org/10.1016/S0076-6879\(09\)67004-X](https://doi.org/10.1016/S0076-6879(09)67004-X).
- [155] S. Sahoo, K. Hari, S. Jhunjhunwala, M.K. Jolly, Mechanistic modeling of the SARS-CoV-2 and immune system interplay unravels design principles for diverse clinicopathological outcomes, *BioRxiv*. (2020).
- [156] C. Li, J. Xu, J. Liu, Y. Zhou, The within-host viral kinetics of SARS-CoV-2, *Math. Biosci. Eng.* 17 (2020), <https://doi.org/10.3934/mbe.2020159>.
- [157] P. Abuin, A. Anderson, A. Ferramosca, E.A. Hernandez-Vargas, A.H. Gonzalez, Characterization of SARS-CoV-2 dynamics in the host, *Annu. Rev. Control.* 50 (2020), <https://doi.org/10.1016/j.arcontrol.2020.09.008>.
- [158] T.J. Segó, J.O. Aponte-Serrano, J. Ferrari Gianlupi, S.R. Heaps, K. Breithaupt, L. Bruschi, J. Crawshaw, J.M. Osborne, E.M. Quardokus, R.K. Plummer, J.A. Glazier, A modular framework for multiscale, multicellular, spatiotemporal modeling of acute primary viral infection and immune response in epithelial tissues and its application to drug therapy timing and effectiveness, *PLoS Comput. Biol.* 16 (2020), <https://doi.org/10.1371/journal.pcbi.1008451>.
- [159] M. Getz, Y. Wang, G. An, A. Becker, C. Cockrell, N. Collier, M. Craig, C.L. Davis, J. Faeder, A.N.F. Versypt, J.F. Gianlupi, J.A. Glazier, S. Hamis, R. Heiland, T. Hillen, D. Hou, M.A. Islam, A. Jenner, F. Kurtoglu, B. Liu, F. Macfarlane, P. Maygrunder, P.A. Morel, A. Narayanan, J. Ozik, E. Pienaar, P. Rangamani, J. E. Shoemaker, A.M. Smith, P. Macklin, Rapid community-driven development of a SARS-CoV-2 tissue simulator pp. 02.019075 *BioRxiv*. 2020 (04) (2020), <https://doi.org/10.1101/2020.04.02.019075>.
- [160] S. Peter, P. Dittrich, B. Ibrahim, Structure and Hierarchy of SARS-CoV-2 Infection Dynamics Models Revealed by Reaction Network Analysis, *Viruses*. 13 (2020), <https://doi.org/10.3390/v13010014>.
- [161] B. Fain, H.M. Dobrovoly, Initial Inoculum and the Severity of COVID-19: A Mathematical Modeling Study of the Dose-Response of SARS-CoV-2 Infections, *Epidemiologia*. 1 (2020), <https://doi.org/10.3390/epidemiologia1010003>.
- [162] E.A. Hernandez-Vargas, J.X. Velasco-Hernandez, In-host Mathematical Modelling of COVID-19 in Humans, *Annu. Rev. Control*, 2020, <https://doi.org/10.1016/j.arcontrol.2020.09.006>.
- [163] S. Sahoo, S. Jhunjhunwala, M.K. Jolly, The Good, The Bad and The Ugly: A Mathematical Model Investigates the Differing Outcomes Among COVID-19 Patients, *J. Indian Inst. Sci.* 100 (2020), <https://doi.org/10.1007/s41745-020-00205-1>.
- [164] T.S. Fung, D.X. Liu, *Human Coronavirus : Host-Pathogen Interaction* (2019) 529–560.
- [165] Y. Liu, L.-M. Yan, L. Wan, T.-X. Xiang, A. Le, J.-M. Liu, M. Peiris, L.L.M. Poon, W. Zhang, Viral dynamics in mild and severe cases of COVID-19, *Lancet Infect. Dis.* 20 (2020), [https://doi.org/10.1016/S1473-3099\(20\)30232-2](https://doi.org/10.1016/S1473-3099(20)30232-2).
- [166] S.Y. Park, S.G. Yun, J.W. Shin, B.Y. Lee, H.-J. Son, S. Lee, E. Lee, T.H. Kim, Persistent severe acute respiratory syndrome coronavirus 2 detection after resolution of coronavirus disease 2019-associated symptoms/signs, *Korean J. Intern. Med.* 35 (2020), <https://doi.org/10.3904/kjim.2020.203>.
- [167] D.J. Irvine, A. Aung, M. Silva, Controlling timing and location in vaccines, *Adv. Drug Deliv. Rev.* 158 (2020) 91–115. S0169409X2030065X. <https://doi.org/10.1016/j.addr.2020.06.019>.
- [168] D.D. Chaplin, Overview of the immune response, *J. Allergy Clin. Immunol.* 125 (2010) S3–S23, <https://doi.org/10.1016/j.jaci.2009.12.980>.
- [169] S.J. Rhodes, C. Sarfas, G.M. Knight, A. White, A.A. Pathan, H. McShane, T.G. Evans, H. Fletcher, S. Sharpe, R.G. White, Using Data from Macaques To Predict Gamma Interferon Responses after *Mycobacterium bovis* BCG Vaccination in Humans: a Proof-of-Concept Study of Immunostimulation/Immunodynamic Modeling Methods, *Clin. Vaccine Immunol.* 24 (e00525–16) (2017) e00525–e616, <https://doi.org/10.1128/CI.00525-16>.
- [170] WHO, R&D Blueprint and COVID-19, (n.d.). <https://www.who.int/teams/blueprint/covid-19>.
- [171] H. Jones, K. Rowland-Yeo, Basic Concepts in Physiologically Based Pharmacokinetic Modeling in Drug Discovery and Development, *CPT Pharmacometrics Syst. Pharmacol.* 2 (2013), <https://doi.org/10.1038/psp.2013.41> e63.
- [172] L. Kuepfer, C. Niederalt, T. Wendl, J.-F. Schlender, S. Willmann, J. Lippert, M. Block, T. Eissing, D. Teutonico, Applied Concepts in PBPK Modeling: How to Build a PBPK/PD Model, *CPT Pharmacometrics Syst. Pharmacol.* 5 (2016) 516–531, <https://doi.org/10.1002/psp4.12134>.
- [173] S.-M. Huang, M. Rowland, The Role of Physiologically Based Pharmacokinetic Modeling in Regulatory Review, *Clin. Pharmacol. Ther.* 91 (2012) 542–549, <https://doi.org/10.1038/clpt.2011.320>.
- [174] S.G. Reed, M.T. Orr, C.B. Fox, Key roles of adjuvants in modern vaccines, *Nat. Med.* 19 (2013) 1597–1608, <https://doi.org/10.1038/nm.3409>.
- [175] M.A. Tegenge, R.J. Mitkus, A physiologically-based pharmacokinetic (PBPK) model of squalene-containing adjuvant in human vaccines, *J. Pharmacokinet. Pharmacodyn.* 40 (2013) 545–556, <https://doi.org/10.1007/s10928-013-9328-y>.
- [176] C. Niederalt, L. Kuepfer, J. Solodenko, T. Eissing, H.-U. Siegmund, M. Block, S. Willmann, J. Lippert, A generic whole body physiologically based pharmacokinetic model for therapeutic proteins in PK-Sim, *J. Pharmacokinet. Pharmacodyn.* 45 (2018) 235–257, <https://doi.org/10.1007/s10928-017-9559-4>.
- [177] M.A. Tegenge, R.J. Mitkus, A first-generation physiologically based pharmacokinetic (PBPK) model of alpha-tocopherol in human influenza vaccine adjuvant, *Regul. Toxicol. Pharmacol.* 71 (2015) 353–364, <https://doi.org/10.1016/j.yrtph.2015.02.005>.
- [178] P. Poulin, S. Haddad, Advancing prediction of tissue distribution and volume of distribution of highly lipophilic compounds from a simplified tissue-composition-based model as a mechanistic animal alternative method, *J. Pharm. Sci.* 101 (2012) 12.
- [179] M. Krauss, K. Tappe, A. Schuppert, L. Kuepfer, L. Goerlitz, Bayesian Population Physiologically-Based Pharmacokinetic (PBPK) Approach for a Physiologically Realistic Characterization of Interindividual Variability in Clinically Relevant Populations, *PLoS One*. 22 (2015).
- [180] S. Basu, Y.T. (Kayla) Lien, V. Vozmediano, J.-F. Schlender, T. Eissing, S. Schmidt, C. Niederalt, Physiologically Based Pharmacokinetic Modeling of Monoclonal Antibodies in Pediatric Populations Using PK-Sim, *Front. Pharmacol.* 11 (2020) 868, <https://doi.org/10.3389/fphar.2020.00868>.
- [181] G.Z. Ferl, A.M. Wu, J.J. DiStefano, A Predictive Model of Therapeutic Monoclonal Antibody Dynamics and Regulation by the Neonatal Fc Receptor (FcRn), *Ann. Biomed. Eng.* 33 (2005) 1640–1652, <https://doi.org/10.1007/s10439-005-7410-3>.
- [182] E. Offman, A.N. Edginton, A PBPK workflow for first-in-human dose selection of a subcutaneously administered pegylated peptide, *J. Pharmacokinet. Pharmacodyn.* 42 (2015) 135–150, <https://doi.org/10.1007/s10928-015-9406-4>.
- [183] E. Offman, C. Phipps, A.N. Edginton, Population physiologically-based pharmacokinetic model incorporating lymphatic uptake for a subcutaneously administered pegylated peptide, *Silico Pharmacol.* 4 (2016) 3, <https://doi.org/10.1186/s40203-016-0018-5>.
- [184] L.T. Baxter, H. Zhu, D.G. Mackensen, Physiologically Based Pharmacokinetic Model for Specific and Nonspecific Monoclonal Antibodies and Fragments in Normal Tissues and Human Tumor Xenografts in Nude Mice, (n.d.) 13.
- [185] D.K. Shah, A.M. Betts, Towards a platform PBPK model to characterize the plasma and tissue disposition of monoclonal antibodies in preclinical species and human, *J. Pharmacokinet. Pharmacodyn.* 39 (2012) 67–86, <https://doi.org/10.1007/s10928-011-9232-2>.
- [186] Y. Chen, J.P. Balthasar, Evaluation of a Catenary PBPK Model for Predicting the In Vivo Disposition of mAbs Engineered for High-Affinity Binding to FcRn, *AAPS J.* 14 (2012) 850–859, <https://doi.org/10.1208/s12248-012-9395-9>.
- [187] H. Liu, K.D. Moynihan, Y. Zheng, G.L. Szeto, A.V. Li, B. Huang, D.S. Van Egeren, C. Park, D.J. Irvine, Structure-based programming of lymph-node targeting in molecular vaccines, *Nature*. 507 (2014) 519–522, <https://doi.org/10.1038/nature12978>.
- [188] M.D. Shin, S. Shukla, Y.H. Chung, V. Beiss, S.K. Chan, O.A. Ortega-Rivera, D.M. Wirth, A. Chen, M. Sack, J.K. Pokorski, N.F. Steinmetz, COVID-19 vaccine development and a potential nanomaterial path forward, *Nat. Nanotechnol.* 15 (2020) 646–655, <https://doi.org/10.1038/s41565-020-0737-y>.
- [189] X. Chen, T. Hickling, P. Vicini, A. Mechanistic, Multiscale Mathematical Model of Immunogenicity for Therapeutic Proteins: Part 2—Model Applications, *CPT Pharmacometrics Syst. Pharmacol.* 3 (2014) 134, <https://doi.org/10.1038/psp.2014.31>.
- [190] X. Chen, T.P. Hickling, P. Vicini, A. Mechanistic, Multiscale Mathematical Model of Immunogenicity for Therapeutic Proteins: Part 1—Theoretical Model, *CPT Pharmacometrics Syst. Pharmacol.* 3 (2014), <https://doi.org/10.1038/psp.2014.30> e133.
- [191] R.J. Sové, M. Jafarnejad, C. Zhao, H. Wang, H. Ma, A.S. Popel, QSP-IO: A Quantitative Systems Pharmacology Toolbox for Mechanistic Multiscale Modeling for Immuno-Oncology Applications, *CPT Pharmacometrics Syst. Pharmacol.* 9 (2020) 484–497, <https://doi.org/10.1002/psp4.12546>.
- [192] M. Jafarnejad, C. Gong, E. Gabrielson, I.H. Bartelink, P. Vicini, B. Wang, R. Narwal, L. Roskos, A.S. Popel, A Computational Model of Neoadjuvant PD-1 Inhibition in Non-Small Cell Lung Cancer, *AAPS J.* 21 (2019) 79, <https://doi.org/10.1208/s12248-019-0350-x>.
- [193] T. Thanh Le, Z. Andreadakis, A. Kumar, R. Gómez Román, S. Tollefsen, M. Saville, S. Mayhew, The COVID-19 vaccine development landscape, *Nat. Rev. Drug Discov.* 19 (2020) 305–306, <https://doi.org/10.1038/d41573-020-00073-5>.
- [194] S. Treweek, N.G. Frouhi, K.M.V. Narayan, K. Khunti, COVID-19 and ethnicity: who will research results apply to?, *Lancet*. 395 (2020) 1955–1957, [https://doi.org/10.1016/S0140-6736\(20\)31380-5](https://doi.org/10.1016/S0140-6736(20)31380-5).
- [195] C. Thuilliez, L. Dorso, P. Howroyd, S. Gould, F. Chanut, R. Burnett, Histopathological lesions following intramuscular administration of saline in laboratory rodents and rabbits, *Exp. Toxicol. Pathol.* 61 (2009), <https://doi.org/10.1016/j.etp.2008.07.003>.
- [196] R.K. Gherardi, M. Coquet, P. Chérin, F.-J. Authier, P. Laforêt, L. Bélec, D. Figarella-Branger, J.-M. Mussini, J.-F. Pellissier, M. Fardeau, Macrophagic myofasciitis: an emerging entity, *Lancet*. 352 (1998), [https://doi.org/10.1016/S0140-6736\(98\)02326-5](https://doi.org/10.1016/S0140-6736(98)02326-5).
- [197] P.L. Acosta, M.T. Caballero, F.P. Polack, Brief History and Characterization of Enhanced Respiratory Syncytial Virus Disease, *Clin. Vaccine Immunol.* 23 (2016) 189–195, <https://doi.org/10.1128/CI.00609-15>.



- [198] M. Vadalà, D. Poddighe, C. Laurino, B. Palmieri, Vaccination and autoimmune diseases: is prevention of adverse health effects on the horizon?, *EPMA J.* 8 (2017), <https://doi.org/10.1007/s13167-017-0101-y>.
- [199] S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, G.P.S. Raghava, Peptide Toxicity Prediction, in (2015), [https://doi.org/10.1007/978-1-4939-2285-7\\_7](https://doi.org/10.1007/978-1-4939-2285-7_7).
- [200] S. Vishnoi, H. Matre, P. Garg, S.K. Pandey, Artificial intelligence and machine learning for protein toxicity prediction using proteomics data, *Chem. Biol. Drug Des.* 96 (2020), <https://doi.org/10.1111/cbdd.13701>.
- [201] S. Gupta, H.R. Ansari, A. Gautam, G.P.S. Raghava, Identification of B-cell epitopes in an antigen for inducing specific class of antibodies, *Biol. Direct.* 8 (2013), <https://doi.org/10.1186/1745-6150-8-27>.
- [202] R. Gacesa, D.J. Barlow, P.F. Long, Machine learning can differentiate venom toxins from other proteins having non-toxic physiological functions, *PeerJ Comput. Sci.* 2 (2016), <https://doi.org/10.7717/peerj-cs.90>.
- [203] T.J. Cole, M.S. Brewer, TOXIFY: a deep learning approach to classify animal venom proteins, *PeerJ.* 7 (2019), <https://doi.org/10.7717/peerj.7200>.
- [204] P.R. Sirohi, J. Gupta, P. Somvanshi, V.K. Prajapati, A. Grover, Multiple epitope-based vaccine prediction against SARS-CoV-2 spike glycoprotein, *J. Biomol. Struct. Dyn.* (2020), <https://doi.org/10.1080/07391102.2020.1846626>.
- [205] M. Tohidinia, F. Sefid, Identification B and T-Cell epitopes and functional exposed amino acids of S protein as a potential vaccine candidate against SARS-CoV-2/COVID-19, *Microb. Pathog.* 148 (2020), <https://doi.org/10.1016/j.micpath.2020.104459>.
- [206] R. Ojha, N. Gupta, B. Naik, S. Singh, V.K. Verma, D. Prusty, V.K. Prajapati, High throughput and comprehensive approach to develop multi-epitope vaccine against minacious COVID-19, *Eur. J. Pharm. Sci.* 151 (2020), <https://doi.org/10.1016/j.ejps.2020.105375>.
- [207] S.M.A. Islam, T. Sajed, C.M. Kearney, E.J. Baker, PredSTP: a highly accurate SVM based model to predict sequential cystine stabilized peptides, *BMC Bioinformatics.* 16 (2015), <https://doi.org/10.1186/s12859-015-0633-x>.
- [208] G. Postic, J. Gracy, C. Périn, L. Chiche, J.-C. Gelly, KNOTTIN: the database of inhibitor cystine knot scaffold after 10 years, toward a systematic structure modeling, *Nucleic Acids Res.* 46 (2018), <https://doi.org/10.1093/nar/gkx1084>.
- [209] S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, O.S.D.D. Consortium, G.P.S. Raghava, In Silico Approach for Predicting Toxicity of Peptides and Proteins, *PLoS One.* 8 (2013), <https://doi.org/10.1371/journal.pone.0073957>.
- [210] C. Kalcheim, R.S. Goldstein, Segmentation of sensory and sympathetic ganglia: interactions between neural crest and somite cells, *J. Physiol.* (1991).
- [211] M.M. Hasan, N. Schaduangrat, S. Basith, G. Lee, W. Shoombuatong, B. Manavalan, HLLpred-Fuse: improved and robust prediction of hemolytic peptide and its activity by fusing multiple feature representation, *Bioinformatics.* 36 (2020), <https://doi.org/10.1093/bioinformatics/btaa160>.
- [212] K. Chaudhary, R. Kumar, S. Singh, A. Tuknait, A. Gautam, D. Mathur, P. Anand, G.C. Varshney, G.P.S. Raghava, A Web Server and Mobile App for Computing Hemolytic Potency of Peptides, *Sci. Rep.* 6 (2016), <https://doi.org/10.1038/srep22843>.
- [213] V. Kumar, R. Kumar, P. Agrawal, S. Patiyal, G.P.S. Raghava, A Method for Predicting Hemolytic Potency of Chemically Modified Peptides From Its Structure, *Front. Pharmacol.* 11 (2020), <https://doi.org/10.3389/fphar.2020.00054>.
- [214] D. Barh, S. Tiwari, B. Silva Andrade, M. Giovanetti, E. Almeida Costa, R. Kumavath, P. Ghosh, A. Góes-Neto, L. Carlos Junior Alcantara, V. Azevedo, Potential chimeric peptides to block the SARS-CoV-2 spike receptor-binding domain F1000Research (2020) 9.
- [215] A. Devi, N.S.N. Chaitanya, Designing of peptide aptamer targeting the receptor-binding domain of spike protein of SARS-CoV-2: an in silico study, *Mol. Divers.* (2021), <https://doi.org/10.1007/s11030-020-10171-6>.
- [216] S. Singh, H. Singh, A. Tuknait, K. Chaudhary, B. Singh, S. Kumaran, G.P.S. Raghava, PEPstrMOD: structure prediction of peptides containing natural, non-natural and modified residues, *Biol. Direct.* 10 (2015), <https://doi.org/10.1186/s13062-015-0103-4>.
- [217] C. Zhang, Z. Wu, J.-W. Li, H. Zhao, G.-Q. Wang, Cytokine release syndrome in severe COVID-19: interleukin-6 receptor antagonist tocilizumab may be the key to reduce mortality, *Int. J. Antimicrob. Agents.* 55 (2020), <https://doi.org/10.1016/j.ijantimicag.2020.105954>.
- [218] D.A. Sarianni, U. Hansen, Considering the cumulative risk of mixtures of chemicals – A challenge for policy makers, *Environ. Heal.* 11 (2012), <https://doi.org/10.1186/1476-069X-11-S1-S18>.
- [219] N. Cedergreen, Quantifying Synergy: A Systematic Review of Mixture Toxicity Studies within Environmental Toxicology, *PLoS One.* 9 (2014), <https://doi.org/10.1371/journal.pone.0096580>.
- [220] K.C. Bulusu, R. Guha, D.J. Mason, R.P.I. Lewis, E. Muratov, Y. Kalantar Motamed, M. Cokol, A. Bender, Modelling of compound combination effects and applications to efficacy and toxicity: state-of-the-art, challenges and perspectives, *Drug Discov. Today.* 21 (2016), <https://doi.org/10.1016/j.drudis.2015.09.003>.
- [221] F. Cheng, I.A. Kovács, A.-L. Barabási, Network-based prediction of drug combinations, *Nat. Commun.* 10 (2019), <https://doi.org/10.1038/s41467-019-09186-x>.
- [222] D.S. Wishart, Y.D. Feunang, A.C. Guo, E.J. Lo, A. Marcu, J.R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, N. Assempour, I. Iynkkaran, Y. Liu, A. Maciejewski, N. Gale, A. Wilson, L. Chin, R. Cummings, D. Le, A. Pon, C. Knox, M. Wilson, DrugBank 5.0: a major update to the DrugBank database for 2018, *Nucleic Acids Res.* 46 (2018) D1074–D1082, <https://doi.org/10.1093/nar/gkx1037>.
- [223] Y.H. Li, C.Y. Yu, X.X. Li, P. Zhang, J. Tang, Q. Yang, T. Fu, X. Zhang, X. Cui, G. Tu, Y. Zhang, S. Li, F. Yang, Q. Sun, C. Qin, X. Zeng, Z. Chen, Y.Z. Chen, F. Zhu, Therapeutic target database update 2018: enriched resource for facilitating bench-to-clinic research of targeted therapeutics, *Nucleic Acids Res.* 46 (2018) D1121–D1127, <https://doi.org/10.1093/nar/gkx1076>.
- [224] M. Whirl-Carrillo, E.M. McDonagh, J.M. Hebert, L. Gong, K. Sangkuhl, C.F. Thorn, R.B. Altman, T.E. Klein, Pharmacogenomics Knowledge for Personalized Medicine, *Clin. Pharmacol. Ther.* 92 (2012) 414–417, <https://doi.org/10.1038/clpt.2012.96>.
- [225] D. Mendez, A. Gaulton, A.P. Bento, J. Chambers, M. De Veij, E. Félix, M.P. Magariños, J.F. Mosquera, P. Mutowo, M. Nowotka, M. Gordillo-Marañón, F. Hunter, L. Junco, G. Mugumbate, M. Rodriguez-Lopez, F. Atkinson, N. Bosc, C.J. Radoux, A. Segura-Cabrera, A. Hersey, A.R. Leach, ChEMBL: Towards direct deposition of bioassay data, *Nucleic Acids Res.* 47 (2019) D930–D940, <https://doi.org/10.1093/nar/gky1075>.
- [226] M.K. Gilson, T. Liu, M. Baitaluk, G. Nicola, L. Hwang, J. Chong, BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology, *Nucleic Acids Res.* 44 (2016) D1045–D1053, <https://doi.org/10.1093/nar/gkv1072>.
- [227] J.F. Armstrong, E. Faccenda, S.D. Harding, A.J. Pawson, C. Southan, J.L. Sharman, B. Campo, D.R. Cavanagh, S.P.H. Alexander, A.P. Davenport, M. Spedding, J.A. Davies, NC-IUPHAR, The IUPHAR, BPS Guide to PHARMACOLOGY in 2020: extending immunopharmacology content and introducing the IUPHAR/MMV Guide to MALARIA PHARMACOLOGY, *Nucleic Acids Res.* 48 (2020) D1006–D1021, <https://doi.org/10.1093/nar/gkz951>.
- [228] N.P. Tatonetti, P.P. Ye, R. Daneshjou, R.B. Altman, Data-Driven Prediction of Drug Effects and Interactions, *Sci. Transl. Med.* 4 (2012) 125ra31 LP-125ra31, <https://doi.org/10.1126/scitranslmed.3003377>.
- [229] F. Cheng, I.A. Kovács, A.L. Barabási, Network-based prediction of drug combinations, *Nat. Commun.* (2019), <https://doi.org/10.1038/s41467-019-09186-x>.
- [230] M. Kuhn, I. Letunic, L.J. Jensen, P. Bork, The SIDER database of drugs and side effects, *Nucleic Acids Res.* 44 (2016), <https://doi.org/10.1093/nar/gkv1075>.
- [231] J.-X. Zhang, W.-J. Huang, J.-H. Zeng, W.-H. Huang, Y. Wang, R. Zhao, B.-C. Han, Q.-F. Liu, Y.-Z. Chen, Z.-L. Ji, DITOP: drug-induced toxicity related protein database, *Bioinformatics.* 23 (2007), <https://doi.org/10.1093/bioinformatics/btm139>.
- [232] O. Sharma, A.A. Sultan, H. Ding, C.R. Triggler, A Review of the Progress and Challenges of Developing a Vaccine for COVID-19, *Front. Immunol.* 11 (2020), <https://doi.org/10.3389/fimmu.2020.585354>.
- [233] H. Can, A.E. Köseoğlu, S. Erkunt Alak, M. Güvendi, M. Döşkaya, M. Karakavuk, A.Y. Gürüz, C. Ün, In silico discovery of antigenic proteins and epitopes of SARS-CoV-2 for the development of a vaccine or a diagnostic approach for COVID-19, *Sci. Rep.* 10 (2020), <https://doi.org/10.1038/s41598-020-79645-9>.
- [234] B. Sarkar, M.A. Ullah, Y. Araf, M.S. Rahman, Engineering a novel subunit vaccine against SARS-CoV-2 by exploring immunoinformatics approach, *Informatics Med. Unlocked.* 21 (2020), <https://doi.org/10.1016/j.imu.2020.100478>.
- [235] S. Khan, B. Shaker, S. Ahmad, S.W. Abbasi, M. Arshad, A. Haleem, S. Ismail, A. Zaib, W. Sajjad, Towards a novel peptide vaccine for Middle East respiratory syndrome coronavirus and its possible use against pandemic COVID-19, *J. Mol. Liq.* (2020), <https://doi.org/10.1016/j.molliq.2020.114706>.
- [236] S. Srivastava, S. Verma, M. Kamthania, D. Agarwal, A.K. Saxena, M. Kolbe, S. Singh, A. Kotnis, B. Rath, S.A. Nayar, H.-J. Shin, K. Vashisht, K.C. Pandey, Computationally validated SARS-CoV-2 CTL and HTL Multi-Patch vaccines, designed by reverse epitomics approach, show potential to cover large ethnically distributed human population worldwide, *J. Biomol. Struct. Dyn.* (2020), <https://doi.org/10.1080/07391102.2020.1838329>.
- [237] I. Dimitrov, L. Naneva, I. Doytchinova, I. Bangov, AllergenFP: allergenicity prediction by descriptor fingerprints, *Bioinformatics.* 30 (2014), <https://doi.org/10.1093/bioinformatics/btt619>.
- [238] H.A. Dar, Y. Waheed, M.H. Najmi, S. Ismail, H.F. Hetta, A. Ali, K. Muhammad, Multi-epitope Subunit Vaccine Design against COVID-19 Based on the Spike Protein of SARS-CoV-2: An In Silico Analysis, *J. Immunol. Res.* 2020 (2020), <https://doi.org/10.1155/2020/8893483>.
- [239] H.C. Muh, J.C. Tong, M.T. Tammi, AllerHunter: A SVM-Pairwise System for Assessment of Allergenicity and Allergic Cross-Reactivity in Proteins, *PLoS One.* 4 (2009), <https://doi.org/10.1371/journal.pone.0005861>.
- [240] J. Wang, D. Zhang, J. Li, PREAL: prediction of allergenic protein by maximum Relevance Minimum Redundancy (mRMR) feature selection, *BMC Syst. Biol.* 7 (2013), <https://doi.org/10.1186/1752-0509-7-55-59>.
- [241] J. Wang, Y. Yu, Y. Zhao, D. Zhang, J. Li, Evaluation and integration of existing methods for computational prediction of allergens, *BMC Bioinformatics.* 14 (2013), <https://doi.org/10.1186/1471-2105-14-S4-S1>.
- [242] P.J. Hotez, D.B. Corry, M.E. Bottazzi, COVID-19 vaccine design: the Janus face of immune enhancement, *Nat. Rev. Immunol.* 20 (2020) 347–348, <https://doi.org/10.1038/s41577-020-0323-4>.
- [243] R. de Alwis, S. Chen, E.S. Gan, E.E. Ooi, Impact of immune enhancement on Covid-19 polyclonal hyperimmune globulin therapy and vaccine development, *EBioMedicine.* 55 (2020), <https://doi.org/10.1016/j.ebiom.2020.102768>.
- [244] A. Carbo, R. Hontecillas, T. Andrew, K. Eden, Y. Mei, S. Hoops, J. Bassaganya-Riera, Computational modeling of heterogeneity and function of CD4+ T cells, *Front. Cell Dev. Biol.* 2 (2014), <https://doi.org/10.3389/fcell.2014.00031>.
- [245] R. Jain, S. Ramaswamy, D. Harilal, M. Uddin, T. Loney, N. Nowotny, H. Alsuwaidi, R. Varghese, Z. Deesi, A. Alkhajeh, H. Khansaheb, A. Alsheikh-Ali, A.

- Abou Tayoun, Host transcriptomic profiling of COVID-19 patients with mild, moderate, and severe clinical outcomes, *Comput. Struct. Biotechnol. J.* 19 (2021) 153–160.
- [246] J. Xu, P. Zhang, Y. Huang, L. Bekris, J. Lathia, Multimodal single-cell / nucleus RNA-sequencing data analysis uncovers molecular networks between disease-associated microglia and astrocytes with implications for drug repurposing in Alzheimer's disease, 2020.
- [247] T.C. Williams, W.A. Burgers, SARS-CoV-2 evolution and vaccines: cause for concern?, *Lancet Respir. Med.* (2021), [https://doi.org/10.1016/S2213-2600\(21\)00075-8](https://doi.org/10.1016/S2213-2600(21)00075-8).
- [248] T. Liu, P. Jia, B. Fang, Z. Zhao, Differential Expression of Viral Transcripts From Single-Cell RNA Sequencing of Moderate and Severe COVID-19 Patients and Its Implications for Case Severity, *Front. Microbiol.* 11 (2020) 2568. <https://www.frontiersin.org/article/10.3389/fmicb.2020.603509>.
- [249] G. Xu, F. Qi, H. Li, Q. Yang, H. Wang, X. Wang, X. Liu, J. Zhao, X. Liao, Y. Liu, L. Liu, S. Zhang, Z. Zhang, The differential immune responses to COVID-19 in peripheral and lung revealed by single-cell RNA sequencing, *Cell Discov.* 6 (2020) 73. <https://doi.org/10.1038/s41421-020-00225-2>.
- [250] A. Silvín, N. Chapuis, G. Dunsmore, A.-G. Goubet, A. Dubuisson, L. Derosa, C. Almiré, C. Hénon, O. Kosmider, N. Droin, P. Rameau, C. Catelain, A. Alfaro, C. Dussiau, C. Friedrich, E. Sourdeau, N. Marin, T.-A. Szwebel, D. Cantin, L. Mouthon, D. Borderie, M. Deloger, D. Bredele, S. Mouraud, D. Drubay, M. Andrieu, A.-S. Lhonneur, V. Saada, A. Stoclin, C. Willekens, F. Pommeret, F. Griscelli, L.G. Ng, Z. Zhang, P. Bost, I. Amit, F. Barlesi, A. Marabelle, F. Pène, B. Gachot, F. André, L. Zitvogel, F. Ginhoux, M. Fontenay, E. Solary, Elevated Calprotectin and Abnormal Myeloid Cell Subsets Discriminate Severe from Mild COVID-19, *Cell.* 182 (2020) 1401–1418.e18, <https://doi.org/10.1016/j.cell.2020.08.002>.
- [251] P.S. Arunachalam, F. Wimmers, C.K.P. Mok, R.A.P.M. Perera, M. Scott, T. Hagan, N. Sigal, Y. Feng, L. Bristow, O. Tak-Yin Tsang, D. Wagh, J. Coller, K.L. Pellegrini, D. Kazmin, G. Alaaeddine, W.S. Leung, J.M.C. Chan, T.S.H. Chik, C.Y.C. Choi, C. Huerta, M. Paine McCullough, H. Lv, E. Anderson, S. Edupuganti, A.A. Upadhyay, S.E. Bosinger, H.T. Maecker, P. Khatri, N. Roupheal, M. Peiris, B. Pulendran, Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans, *Science* (80–). 369 (2020) 1210 LP – 1220. <https://doi.org/10.1126/science.abc6261>.
- [252] T. Liu, K. Shi, W. Li, Deep learning methods improve linear B-cell epitope prediction, *BioData Min.* 13 (2020), <https://doi.org/10.1186/s13040-020-00211-0>.
- [253] H.A. Odhar, Towards the design of epitope candidates for Coronavirus 2, *Bioinformatics.* 16 (2020), <https://doi.org/10.6026/97320630016375>.
- [254] A. Marino Gammazza, S. Légaré, G. Lo Bosco, A. Fucarino, F. Angileri, E. Conway de Macario, A.J.L. Macario, F. Cappello, Human molecular chaperones share with SARS-CoV-2 antigenic epitopes potentially capable of eliciting autoimmunity against endothelial cells: possible role of molecular mimicry in COVID-19, *Cell Stress Chaperones.* 25 (2020), <https://doi.org/10.1007/s12192-020-01148-3>.
- [255] Z. Yazdani, A. Rafiei, M. Yazdani, R. Valadan, <p>Design an Efficient Multi-Epitope Peptide Vaccine Candidate Against SARS-CoV-2: An in silico Analysis</p>, *Infect. Drug Resist.* 13 (2020), <https://doi.org/10.2147/IDR.S264573>.
- [256] A. Alam, A. Khan, N. Imam, M.F. Siddiqui, M. Waseem, M.Z. Malik, R. Ishrat, Design of an epitope-based peptide vaccine against the SARS-CoV-2: a vaccine-informatics approach, *Brief. Bioinform.* (2020), <https://doi.org/10.1093/bib/bbaa340>.
- [257] Y. Vashi, V. Jagrit, S. Kumar, Understanding the B and T cell epitopes of spike protein of severe acute respiratory syndrome coronavirus-2: A computational way to predict the immunogens, *Infect. Genet. Evol.* 84 (2020), <https://doi.org/10.1016/j.meegid.2020.104382>.
- [258] V. Chauhan, T. Rungta, M. Rawat, K. Goyal, Y. Gupta, M.P. Singh, Excavating SARS-coronavirus 2 genome for epitope-based subunit vaccine synthesis using immunoinformatics approach, *J. Cell. Physiol.* 236 (2021), <https://doi.org/10.1002/jcp.29923>.
- [259] S. Banerjee, K. Majumder, G.J. Gutierrez, D. Gupta, B. Mittal, Immunoinformatics approach for multi-epitope vaccine designing against SARS-CoV-2, *BioRxiv.* (2020).
- [260] R. Anand, S. Biswal, R. Bhatt, B.N. Tiwary, Computational perspectives revealed prospective vaccine candidates from five structural proteins of novel SARS corona virus 2019 (SARS-CoV-2), *PeerJ.* 8 (2020), <https://doi.org/10.7717/peerj.9855>.
- [261] D. Forni, R. Cagliani, C. Pontremoli, A. Mozzi, U. Pozzoli, M. Clerici, M. Sironi, Antigenic variation of SARS-CoV-2 in response to immune pressure, *Mol. Ecol.* (2020), <https://doi.org/10.1111/mec.15730>.
- [262] W.R. Martin, F. Cheng, A rational design of a multi-epitope vaccine against SARS-CoV-2 which accounts for the glycan shield of the spike glycoprotein, *ChemRxiv.* (2020).
- [263] D. Requena, A. Médico, R.D. Chacón, M. Ramírez, O. Marín-Sánchez, Identification of Novel Candidate Epitopes on SARS-CoV-2 Proteins for South America: A Review of HLA Frequencies by Country, *Front. Immunol.* 11 (2020), <https://doi.org/10.3389/fimmu.2020.02008>.
- [264] G. Liu, B. Carter, T. Bricken, S. Jain, M. Viard, M. Carrington, D.K. Gifford, Computationally Optimized SARS-CoV-2 MHC Class I and II Vaccine Formulations Predicted to Target Human Haplotype Distributions, *Cell Syst.* 11 (2020) 131–144.e6, <https://doi.org/10.1016/j.cels.2020.06.009>.
- [265] A.S. Shomuradova, M.S. Vagida, S.A. Sheetikov, K.V. Zornikova, D. Kiryukhin, A. Titov, I.O. Peshkova, A. Khmelevskaya, D.V. Dianov, M. Malasheva, A. Shmelev, Y. Serdyuk, D.V. Bagaev, A. Pivnyuk, D.S. Shcherbinin, A.V. Maleeva, N.T. Shakirova, A. Pilunov, D.B. Malko, E.G. Khamaganova, B. Biderman, A. Ivanov, M. Shugay, G.A. Efimov, SARS-CoV-2 Epitopes Are Recognized by a Public and Diverse Repertoire of Human T Cell Receptors, *Immunity.* 53 (2020), <https://doi.org/10.1016/j.immuni.2020.11.004>.
- [266] P. Mani Mishra, V.N. Uversky, C.K. Nandi, Serum albumin-mediated strategy for the effective targeting of SARS-CoV-2, *Med. Hypotheses.* 140 (2020), <https://doi.org/10.1016/j.mehy.2020.109790> 109790.
- [267] B. Dearlove, E. Lewitus, H. Bai, Y. Li, D.B. Reeves, M.G. Joyce, P.T. Scott, M.F. Amare, S. Vasan, N.L. Michael, K. Modjarrad, M. Rolland, A SARS-CoV-2 vaccine candidate would likely match all currently circulating variants, *Proc. Natl. Acad. Sci. U. S. A.* 117 (2020) 23652–23662, <https://doi.org/10.1073/pnas.2008281117>.
- [268] C.H. Lee, H. Koohy, In silico identification of vaccine targets for 2019-nCoV, *F1000Research.* 9 (2020), <https://doi.org/10.12688/f1000research.22507.2>.
- [269] C.R.B. Bonin, G.C. Fernandes, R.W. dos Santos, M. Lobosco, A qualitatively validated mathematical-computational model of the immune response to the yellow fever vaccine, *BMC Immunol.* 19 (2018) 15, <https://doi.org/10.1186/s12865-018-0252-1>.
- [270] A.B. Pigozzo, G.C. Macedo, R.W. dos Santos, M. Lobosco, On the computational modeling of the innate immune system S7 S7 *BMC Bioinformatics.* 14 (2013), <https://doi.org/10.1186/1471-2105-14-S6-S7>.
- [271] X. Li, M. Bohner, C.-K. Wang, Impulsive differential equations: Periodic solutions and applications, *Automatica.* 52 (2015), <https://doi.org/10.1016/j.automatica.2014.11.009>.
- [272] A.N. Chatterjee, F. Al Basir, A Model for SARS-CoV-2 Infection with Treatment, *Comput. Math. Methods Med.* (2020 2020).), <https://doi.org/10.1155/2020/1352982>.
- [273] J. Mata, M. Cohn, Cellular automata-based modeling program: synthetic immune system, *Immunol. Rev.* 216 (2007) 198–212, <https://doi.org/10.1111/j.1600-065X.2007.00511.x>.
- [274] A. Bankhead, E. Mancini, A.C. Sims, R.S. Baric, S. McWeeney, P.M.A. Sloot, A simulation framework to investigate in vitro viral infection dynamics, *J. Comput. Sci.* 4 (2013), <https://doi.org/10.1016/j.jocs.2011.08.007>.
- [275] F. Fatehi, R.J. Bingham, E.C. Dykeman, P.G. Stockley, R. Twarock, Comparing antiviral strategies against COVID-19 via multi-scale within host modelling, *ArXiv.* (2020).