VHH Structural Modelling Approaches: A Critical Review

Poonam Vishwakarma, Akhila Melarkode Vattekatte, Nicolas Shinada, Julien Diharce, Carla Martins, Frédéric Cadet, Fabrice Gardebien, Catherine Etchebest, Aravindan Arun Nadaradjane, Alexandre de Brevern

▶ **To cite this version:**

Poonam Vishwakarma, Akhila Melarkode Vattekatte, Nicolas Shinada, Julien Diharce, Carla Martins, et al.. VHH Structural Modelling Approaches: A Critical Review. International Journal of Molecular Sciences, 2022, 23 (7), pp.3721. 10.3390/ijms23073721 . hal-03665842

HAL Id: hal-03665842

https://u-paris.hal.science/hal-03665842

Submitted on 12 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

International Journal of
Molecular Sciences

MDPI

Review

V_HH Structural Modelling Approaches: A Critical Review

Poonam Vishwakarma [1,2,†], Akhila Melarkode Vattekatte [1,2,†], Nicolas Shinada [3,‡], Julien Diharce [1,‡], Carla Martins [1,2], Frédéric Cadet [2,4], Fabrice Gardebien [2], Catherine Etchebest [1], Aravindan Arun Nadaradjane [1,2,§] and Alexandre G. de Brevern [1,2,*,§]

1 INSERM UMR_S 1134, BIGR, DSIMB Team, Université de Paris and Université de la Réunion, F-75015 Paris, France; poonam.vishwakarma@univ-reunion.fr (P.V.); akhila.melarkode-vattekatte@univ-reunion.fr (A.M.V.); julien.diharce@univ-paris-diderot.fr (J.D.); cmartins@insa-toulouse.fr (C.M.); catherine.etchebest@univ-paris-diderot.fr (C.E.); aravindan.nadaradjane@univ-reunion.fr (A.A.N.)
2 INSERM UMR_S 1134, BIGR, DSIMB Team, Université de Paris and Université de la Réunion, F-97715 Saint Denis Messag, France; frederic.cadet.run@gmail.com (F.C.); fabrice.gardebien@univ-reunion.fr (F.G.)
3 3 SBX Corp., Tokyo-to, Shinagawa-ku, Tokyo 141-0022, Japan; shinada@sbx-corp.com
4 PEACCEL, Artificial Intelligence Department, Square Albin Cachot, F-75013 Paris, France
* Correspondence: alexandre.debrevern@univ-paris-diderot.fr; Tel.: +33-1-44493000
† These authors contributed equally to this work.
‡ These authors contributed equally to this work.
§ These authors contributed equally to this work.

Abstract: V_HH, i.e., VH domains of camelid single-chain antibodies, are very promising therapeutic agents due to their significant physicochemical advantages compared to classical mammalian antibodies. The number of experimentally solved V_HH structures has significantly improved recently, which is of great help, because it offers the ability to directly work on 3D structures to humanise or improve them. Unfortunately, most V_HHs do not have 3D structures. Thus, it is essential to find alternative ways to get structural information. The methods of structure prediction from the primary amino acid sequence appear essential to bypass this limitation. This review presents the most extensive overview of structure prediction methods applied for the 3D modelling of a given V_HH sequence (a total of 21). Besides the historical overview, it aims at showing how model software programs have been shaping the structural predictions of V_HHs. A brief explanation of each methodology is supplied, and pertinent examples of their usage are provided. Finally, we present a structure prediction case study of a recently solved V_HH structure. According to some recent studies and the present analysis, AlphaFold 2 and NanoNet appear to be the best tools to predict a structural model of V_HH from its sequence.

Keywords: antibodies; frameworks; Complementarity Determining Regions; single-domain antibody; secondary structure; sequence–structure relationship; homology and comparative modelling; threading; deep learning; docking

check for updates

Citation: Vishwakarma, P.; Vattekatte, A.M.; Shinada, N.; Diharce, J.; Martins, C.; Cadet, F.; Gardebien, F.; Etchebest, C.; Nadaradjane, A.A.; de Brevern, A.G. V_HH Structural Modelling Approaches: A Critical Review. *Int. J. Mol. Sci.* **2022**, *23*, 3721. https://doi.org/10.3390/ijms23073721

Academic Editor: Honoo Satake

Received: 25 February 2022
Accepted: 23 March 2022
Published: 28 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Proteins carry out most of the functions in a cell. Among them, antibodies (Abs) or Immunoglobulins (Ig) play a major role in the immune response. The antibodies are found in mammals (see Figure 1a) and are classified into five main classes, namely IgA, IgG, IgD, IgE and IgM, IgG being the most abundant Immunoglobulin. IgM and IgA are multimeric Abs, while IgG, IgE and IgD are monomeric.

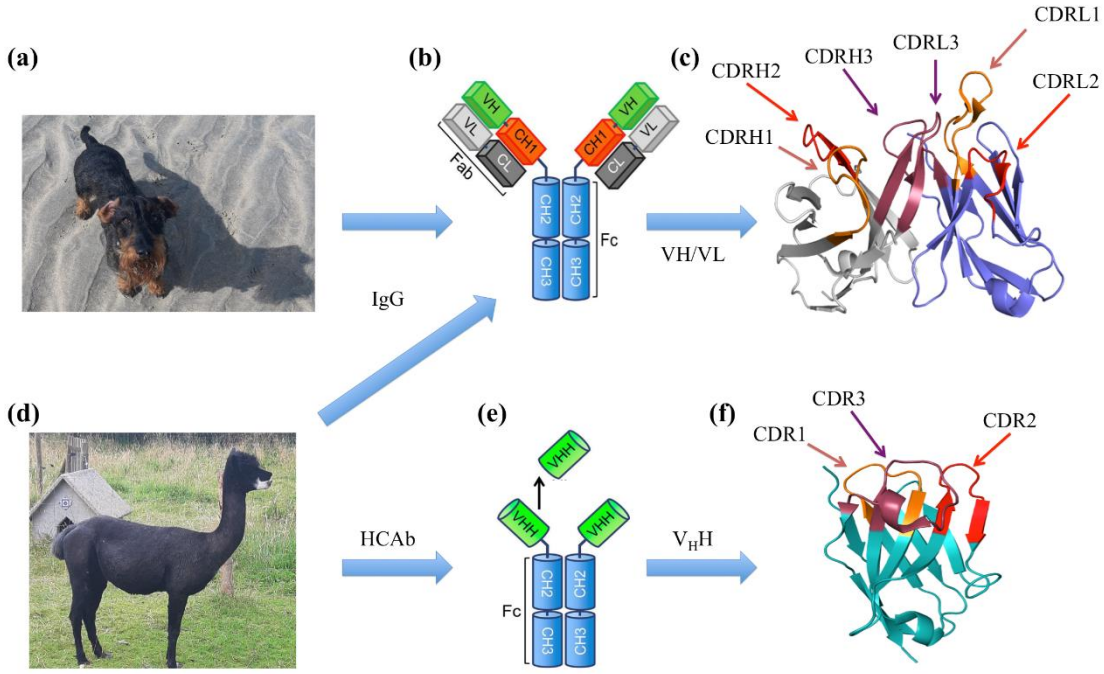

Figure 1. Antibody domains. Immunoglobulins are found in mammals such as the dachshund (*Canis familiaris*) named Snoopy (**a**) and a Vicuña (**d**); both pictures were taken in Normandy (France) in August 2021. The immunoglobulins are shown by domains in (**b**) Classical Immunoglobulin Gamma (IgG) and (**e**) HCAb (only found in the *Camelidae* family). (**c**,**f**) Antigen-binding domain of IgG and HCAb, namely the VH/VL domain and V_HH domain. The FRs in the VH domain are shown in grey, in the VL domain in dark blue and in the V_HH domain in cyan. The CDRs of the heavy chain are represented as CDRH1, CDRH2 and CDRH3, while CDRL1, CDRL2 and CDRL3 are CDRs of the light chain. The CDRs are represented in red shades: CDR1 (orange), CDR2 (red) and CDR3 (raspberry). Visualisation was performed using PyMOL [1–3]. Please notice that Snoopy was not harmed when taking these pictures and received an optimised dose of beef and turkey dog treats that he likes a lot.

They all contain four chains: two identical heavy chains and two identical light chains. The heavy chains comprise three constant domains (CH_1, CH_2 and CH_3), followed by one variable domain (VH). In contrast, each light chain has only one constant domain (CL_1) and one variable domain (VL, see Figure 1b). The major function of antibodies is antigen binding, i.e., the capacity of recognising and binding to a specific target determined by the sequence and structural characteristics. The anatomy of the variable region (VH and VL) of an antibody can be further detailed with (i) four Framework Regions (FRs) and (ii) three (hyper)variable loops named Complementarity Determining Regions (CDRs) (see Figure 1c). CDRs constitute the main part of the so-called paratope region and are directly implicated in the interaction with the epitope (part of the antigen specifically recognised by the antibody). In terms of affinity, interaction ranges are in the nanomolar to micromolar range. This high affinity between a molecule and its antibody has led to a number of applications for diagnosis, therapeutics and in research fields [4–6]. One of the most interesting developments is the emergence of bispecific antibodies (BsAbs). BsAbs neutralise two specific targets using two different epitope binding regions formed by two variable fragments from each chain. Three of them are available on the market [7]: (i) Catumaxomab is used against chemotherapy-refractory ovarian cancer and recurrent malignant ascites by targeting EpCAM antigen on tumour cells and the CD3 antigen on T cells [8]; (ii) Blinatumomab against Philadelphia chromosome-negative (Ph-)-B-cell acute lymphoblastic leukaemia, which acts on CD3 and CD19, and a complete remission can be achieved (which was prevented before the approval of this drug) [9] and (iii) Emicizumab is used for the treatment of patients with congenital factor VIII deficiency [10].

Even though promising therapeutic strategies have been developed and are available in the market, Abs remain quite challenging to tackle in terms of manufacturing and purification [11,12]. One of the important limitations lies in the process cost to obtain a stable and functional molecule. Other factors like stability, bioavailability and the expected immune response are also difficult to optimise [13]. Hence, one relevant strategy to optimise the cost of production consists of using in silico approaches beforehand. Since the precise 3D structure plays a major role in the specificity of the recognition, a key step requires the elucidation of this 3D structure. Besides experimental procedures, theoretical methodologies constitute an alternative and promising way for predicting and studying their structures. Moreover, they are extremely valuable tools for understanding more deeply their properties and also offer the possibility to design new antibodies [14–16]. In this context, the analysis and prediction of IgG structures and complexes has been a major research field for 30 years. Two consecutive competitions, namely AMA-I and –II (Antibody Modelling Assessment), were initiated in the 2010s in order to assess the antibody structure prediction methodologies [17,18]. Their evaluations of state-of-the-art software like Accelrys [19], PIGS [20] and the Rosetta Antibody modelling suite [21] underlined the difficulty in properly predicting the conformations of the six CDRs (three from heavy chains and three from light chains) [17,18].

Interestingly, some animals have developed immunoglobulins with slightly different architectures that often encompass only one chain. The first one comes from cartilaginous fishes such as sharks. They have a special antibody called the "New Antigen Receptor antibody" (IgNAR). The IgNAR is longer than human IgG heavy chains, with dissimilar sequences [22]. The variable domain of the IgNAR is stable and small in size, which is very valuable for drug discovery [23]. Nearly thirty years ago, in addition to classical antibodies, a new class of immunoglobulin was discovered in camelids (this family includes Bactrian camel, dromedary camel, guanaco, llama, alpaca and vicuña; see Figure 1d) [24,25], where the light chain and, also, the CH_1 domain of the heavy chain were absent. Consequently, they were named Heavy Chain-only Antibodies (HCAb, see Figure 1e). Interestingly, their VH domain, named V_HH for Variable domain of the Heavy chain of HCAbs or commercially nanobody, is composed of, as any VH or VL domain, four frameworks and three CDRs (see Figure 1e). They are less than 130 residues long but often have a longer CDR3 than their conventional equivalents. Their smaller size makes them easier to manufacture (and at a low cost) than Abs. Moreover, V_HH can be used independently from HCAb, while VH and VL must be combined for classical Abs [26,27].

V_HHs are thermostable, robust and can be tailored depending on the goal of their utility [28–31]. Furthermore, they bring advantages similar to classical Abs: they can be used for in vivo and in vitro diagnosis, therapy and research [31], e.g., in medicine [32], CAR-T [33], cancers [34–40], the detection of pathogens [41] or as biosensors [42]. For in vivo diagnosis, V_HHs have the capacity to pass through the blood vessels and diffuse rapidly to the tissues, which is helpful for in vivo imaging techniques like SPECT (single-photon emission computerised tomography) and PET/CT (positron emission tomography/computerised tomography). In terms of diagnosis, V_HHs have been efficient in diagnosing and monitoring the evolution [42] of HER2-positive breast carcinoma, atherosclerotic plaques and arthritis. V_HHs are also adapted for enzyme-linked immunosorbent assays (ELISA), a technique that is used for diagnosis and research purposes [43–45] for quantifying the amount of molecules in a biological sample. Another example is the Double-Antibody Sandwich ELISA (DAS-ELISA), where V_HHs are used to capture and detect the presence of a molecule of interest, e.g., to detect the alpha serum protein in foetal blood [46] or Staphylococcal enterotoxin C in dairy products [47].

V_HHs are also highly promising therapeutic molecules [48]. Yet, a humanisation step is first required to make V_HHs an effective therapeutic molecule for humans [49]. As an example, Caplacizumab was the first therapeutic V_HH authorised by FDA against acquired thrombotic thrombocytopenic purpura (aTTP) [50]. Ozoralizumab and vobarilizumab are two other V_HHs currently utilised in clinical trials for rheumatoid arthritis [51,52]. Very

recently, in the context of the COVID-19 pandemic, an impressive number of V_HHs have also been developed against the SARS-CoV-2 spike protein [53–65]. Therefore, a large number of pharmaceutical companies are now developing V_HHs. For instance, AbLynx (now part of Sanofi) has a maximum of six V_HHs in different phases, and three out of six are bispecific V_HHs [31].

The production of V_HHs is mainly obtained from immunising the response of camelids: the molecule of interest is injected to the animal, and the produced clones of V_HHs are then recovered. Specific sequences of V_HHs (e.g., by using next-generation sequencing) targeting a given antigen have to be determined, and the best proper clone(s) need to be expressed. In this aim, different methods have been developed [66] that use, for instance, phage display libraries or even plants [67–69]. For one particular antigen, a large number of experiments can be required to identify a lead V_HH with a pertinent affinity.

As for classical Abs, the understanding of the interaction between V_HH and its partner of interest is better achieved when the 3D structure of the complex is available [70]. Indeed, it would allow designing new V_HH sequences with improved affinity or specificity or, in the context of biotechnology, grafting other partners (including V_HHs) onto it. Despite the increasing number of V_HH structures available in the Protein Data Bank (PDB) [71] in recent years, the structures of a large number of V_HH sequences are far from being fully solved. It is therefore necessary to use computational approaches to access this 3D information of interest [72]. However, as was shown by the first modelling of a V_HH targeting the human DARC (Duffy Antigen/Receptor for Chemokine, now called Atypical Chemokine Receptor 1) protein in 2010, getting a relevant V_HH 3D model [73] is quite difficult and remains extremely challenging, as exemplified by a very recent study [74].

While AMA-I and –II competitions have analysed IgG molecular modelling success, none have been done for V_HHs [17,18]. Similarly, the modelling of IgG structures benefits from a large number of dedicated tools such as PyIgClassify to classify the structures of their CDRs [75], but nothing equivalent exists specifically for V_HHs. For example, no dedicated approach to specifically model V_HHs had been developed until last year. Accordingly, for the first time, the present review aims at giving an overview of the specificities of the V_HH 3D model proposals and their improvements over the last 12 years. Different tools available will be presented, and examples from the literature will be discussed. The advantages and performances of the modelling tools can be more easily compared and discussed, since they are applied to the same family of interest.

2. V_HH Modelling

2.1. General Principle, a Short History

The evolution of sequencing techniques over the last 50 years implies that entire genomes are now accessible in a few hours, as is their proteome [76,77]. The gap between the number of protein sequences and experimental structures continues to grow despite the increased number of structures deposited in the PDB [71,78,79]. Additionally, for more than 40 years, methods for predicting 3D structural models from a sequence using computational approaches have made impressive progress [80,81]. The first approaches were based on simplified physicochemical principles, such as lattice folding in the early 1970s [82]. They were quickly supplanted by the use of analogies between proteins based on evolution. This paradigm states that mutations within a sequence could happen and accumulate themselves, leading to a minor modification in the 3D structure while keeping the same function. This conclusion gives rise to the following principle: "structure is more conserved than sequence". These homology or comparative modelling approaches have been developed from the 1990s onwards [83] with the famous software called "Modeller" [84]. Nonetheless, the main pitfall of this approach lies in the threshold of the sequence identity between the query and the template sequences. When it is too low, it precludes the use of this kind of method. A solution can be found by looking for the compatibility between known folds and very distant sequences, namely the threading approach. These techniques had some nice results in the late 1990s, e.g., GenThreader [85], or, more recently, ORION [86,87], despite a

higher computational cost. In parallel, approaches that combine small fragments from very distant proteins and optimise tens of thousands of 3D models are being developed and are reaching maturity with Rosetta [88,89] and I-Tasser [90]. Finally, very recently, Deep Learning approaches have allowed reaching a new level, sometimes leading to models of almost atomic quality with AlphaFold 2 [91] and trRosetta [92,93]. These approaches, composed of multiple complex neural networks, combine very long-distance evolutionary searches and advanced local compositional proposals. These advances have been achieved thanks to the rising computational power of GPUs in the last few years and better mathematical representations [94].

An important point to consider is the need for accurate metrics to quantify the distances between 3D structures (or structural models) to estimate the reliability of these predictions. The most classical, but also the strictest, is the Root Mean Square Deviation (RMSD) [95], as it corresponds to a classical Euclidean distance between two structures. Its evaluation therefore depends on the length of the protein of interest. To take into account the inherent flexibility of proteins (a long and poorly structured loop can lead to a high RMSD, despite a good overall prediction), many metrics weighting these atomic distances have been developed, the best known being the GDT_TS [96] and the TM-Score [97,98]. They take into account the quality of the folding compared to the reference structure and are thus very much used in the evaluation of Critical Assessment for Structure Prediction (CASP) competitions [99].

2.2. Abs and V_HHs Specificities

As presented before, Abs and V_HHs have specific global and local topologies, i.e., FRs and CDRs [100–102], or a VH/VL interface [103,104]. Hence, some specific Abs modelling tools have been developed. Of course, classical protein modelling tools (see Figure 2a) can be used for Abs and V_HHs. However, because of the various implications of Abs in biotechnology and biomedical domains and their diversity (Figure 2), the need for dedicated Abs tools has emerged (see Figure 2b). Abs software can model classical antibodies, but some of them offer options to model heavy chains or light chains only. V_HH sequences are often compared to heavy chains of antibodies with lengthy CDR3 sequences. These tools can accept a V_HH sequence and predict an associated structural model. In this review, modelling software programs are classified into simple and more advanced techniques. The number of V_HH structural models remains limited in the literature, but their variety is quite rare and remarkable.

2.3. Modeller and ModWeb

Modeller [84] is probably the most important software of comparative/homology modelling. Developed in 1993, it is still maintained (https://salilab.org/modeller/, last accessed date: 7 February 2022), free for academics and included in commercial packages. Modeller automatically produces a model encompassing all nonhydrogen atoms based on (multiple) sequence alignment for a sequence (provided by the user) to be modelled with known related structures [106]. It takes as the input an alignment file of the sequence query with one or several sequences of structural template(s). Different refinement options exist. After multiple cycles of building and model evaluations, the chosen number of structural models is produced. The user must choose the best one associated with the lowest DOPE score [107]. Moreover, developers have proposed an online web server named ModWeb that uses PSI-BLAST [108] and IMPALA [109] to find compatible structural templates and Modeller to build the structural model [84].

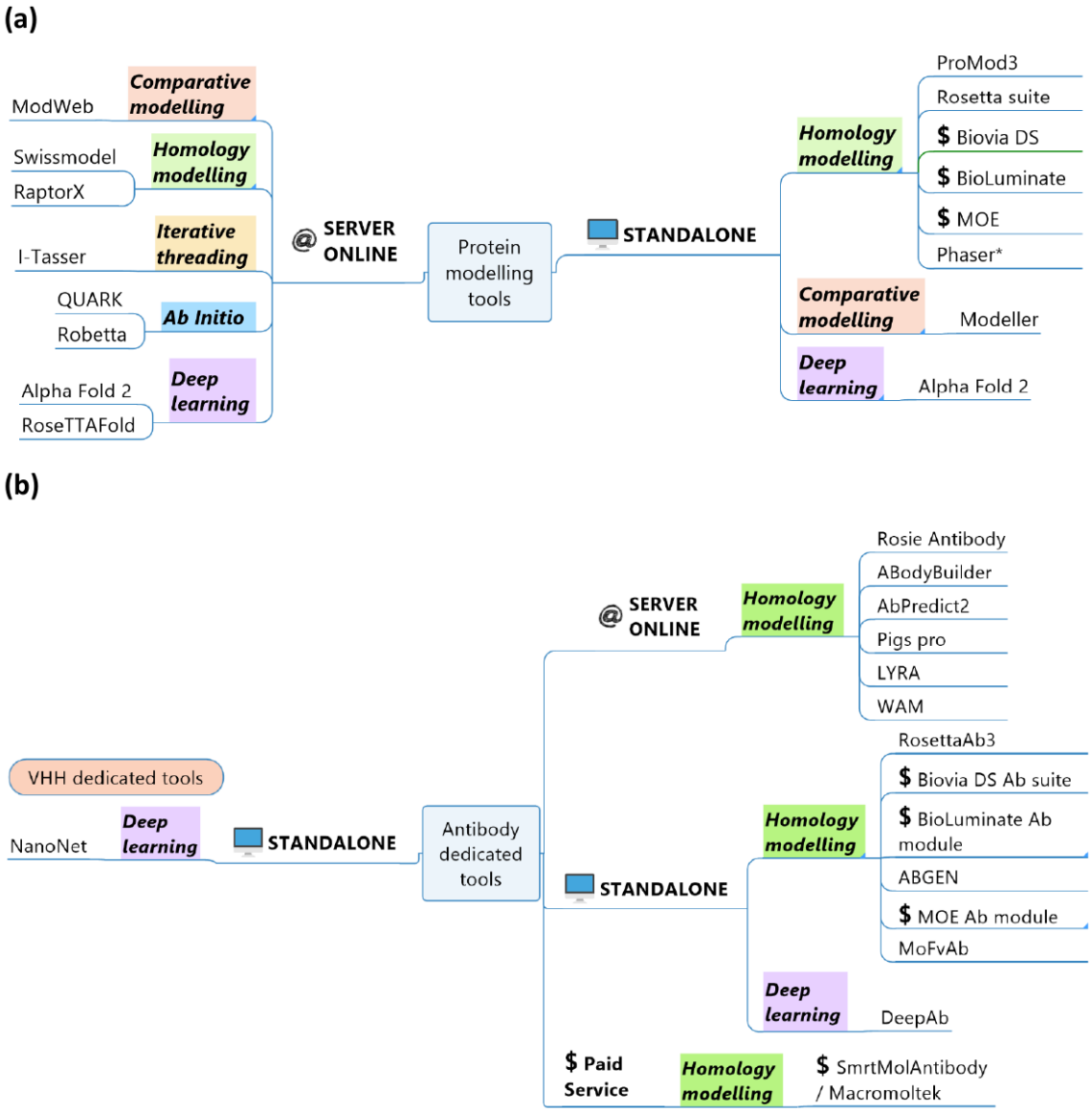

Figure 2. Protein structural prediction tools. The above representation is a summary of (**a**) general protein modelling tools and (**b**) Antibody and V_HH modelling tools. All the tools are classified according to their availability (standalone and/or online server, academic or commercial) and the methodologies used (comparative/homology modelling, threading, ab initio and deep learning). * The phaser is dedicated to molecular replacement [105].

The V_HH fold looks well-conserved in the frameworks [70] but quite variable in CDRs. Comparative modelling using Modeller is a legitimate solution to predict the unknown structure of an antibody [110] and of V_HH from its sequence. Nonetheless, the high sequence variability at the CDRs makes it a difficult challenge. Nowadays, Modeller is the most used software to model V_HH structural models in theoretical studies.

In 2010, our team modelled the first V_HH [111]. This V_HH was optimised against human DARC/ACKR1, which is the entry point of a malarial parasite, *Plasmodium vivax*, into red blood cells. The structure of this anti-DARC V_HH (named CA52) was built using two templates (PDB IDs: 1OP9 and 1JTO), selected using PSI-BLAST against PDB. The sequence query is 124 amino acids long and has an extra disulphide bond between CDR1 and CDR3 and was constrained as found in the second template. Two hundred models were generated using Modeller. Comparisons of the models were done by superimposition;

the model with the median RMSD was selected. This un-classical approach was chosen, as the DOPE score was non-discriminative enough due to the extremely high structural similarity of the models. Indeed, this high similarity is explained by two reasons: the extra disulphide bridge constraint and an entirely equivalent fold of the frameworks [73]. This model was compared to other (and distinct) V_HH clones obtained from the same camel and helped to search for a peptide able to bind DARC; an analysis of the electrostatic surface determinants was particularly important [112].

Another example focused on the cytokine TNF, a well-known drug target for several inflammatory diseases, such as Crohn's disease. Two anti-human TNFR1 V_HHs were experimentally generated and linked with an anti-albumin V_HH to create "TNF Receptor-One Silencer" (TROS) [113]. The two V_HHs were built with Modeller using multiple templates from four different structures (PDB IDs: 4FZE, 4JVP, 2KH2 and 3P0G). RAMPAGE software, a tool based on the local protein geometry, was used to validate the generated models [114]. The best models ranked were used for docking using ClusPro [115] with its human target TRNF1. This last step revealed the different binding sites of the two V_HHs. The authors used it to design bispecific V_HHs and tested them experimentally [113].

Other examples were done using the single structural templates for the design of V_HH focusing on several targets: vascular endothelial growth factor receptor 2 (VEGFR2) for antiangiogenic strategies in cancer therapy [116], bone morphogenic factor 4 (BMP4, implicated in carcinogenesis and tumour progression [117]), CD47 molecule, a promising cancer biomarker [118] and Vascular Endothelial Growth Factor 165 implicated in tumour angiogenesis and metastasis [119]. One can also cite their use for an antivenom therapy inhibiting two myotoxic phospholipases from *Bothrops jararacussu* venom: Bothropstoxin I and II [120]. Furthermore, multiple templates were employed to model a V_HH against matrix metalloproteinase 8 (MMP8) linked to several pathological conditions, e.g., lethal hepatitis and the systemic inflammatory response syndrome [121]. All these studies modelled a V_HH structure from its sequence and used it to perform docking [122]. Studying the interaction and affinity between V_HH and its target of interest was the major goal of these papers. Additional experimental information allowed refining the models and/or docking, e.g., V_HH anti-MMP8 second poses were considered better in regards to ELISA data [121] and the development of a structure-based engineering approach to abrogate pre-existing antibody binding in Biotherapeutics [123]. Regarding the actual sanitary context, Modeller was used for the design of a V_HH that neutralised SARS-CoV-2. The docking made with the conceived V_HH was in excellent agreement with the experiments and explained the difference of binding between V_HHs [61].

Modeller was also used for the refinement of small-angle X-ray scattering analysis results in the context of the interaction of a V_HH interacting with the Disrupted-in-Schizophrenia 1 (DISC1) protein, involved in neurodevelopment and chronic mental illness. Modeller, a more supervised approach, was shown to be more efficient than a more complex method, QUARK (detailed later) [124].

2.4. ABGEN

ABGEN software was made available in 1996. ABGEN is a homology-based algorithm that models an antibody and has some strong similarities with Modeller. It consists of two modules called ABalign and ABbuild [125]. ABalign finds templates for the heavy and light chains and provides an identity score for each template identified. Based on the best template selected, ABbuild constructs a rigid model for all chains; then, the loops and side chains are optimised. Finally, ABGEN includes dedicated mechanistic approaches with XPLOR [126] and GROMOS [127] to refine the obtained models. ABGEN is no longer accessible and has not been used on V_HHs but was the first to specifically optimise the CDR loops.

2.5. Web Antibody Modelling

Web antibody modelling (WAM) was a very interesting hierarchical approach published in 2000 [128] and dedicated to antibodies. It improved previous works [129,130] and can be described in five consecutive steps:

1. A similarity search is performed to build the framework (backbone and side chains) and canonical loop backbones with closer structures (in terms of sequence similarity).
2. Using CONGEN [131], the canonical loop side chains are constructed using an iterative placement technique searching for the global minimum energy conformation.
3. Depending on the loop length, alternative conformations are produced by using directly the PDB or CONGEN again.
4. A specific solvent-modified version of the Valence force field is tested to assess the different conformations.
5. Finally, the conformation is selected from the five lowest energy conformations. The final model is the conformation with the set of torsion angles closest to the canonical one, as defined in the chosen clustering approach.

Theoretically, this tool can be used to model a V_HH sequence, but this has not been done, and its website has been unavailable for some years. Still, it represents the first real attempt to optimise part-by-part an antibody.

2.6. SWISS-MODEL

SWISS-MODEL is a famous online protein homology modelling server that was firstly published in 2003 [132]. It has been frequently updated with both interface ergonomic and methodological improvements [133,134], but the basis is rooted in the principles of the Modeller algorithm (see above). The research of compatible sequences is made by PSI-BLAST [108], combined with delicate HHblits [135]. Then, the user can select which template(s) to use. An interesting point is the different measures provided to assess the qualities of the several models built, i.e., the Global Model Quality Estimate (GMQE) [133]. Building of the structural model is done with locally developed OpenStructure, i.e., an integrated software framework for computational structural biology [136]. In a similar way, they also developed measures to assess the quality of the models, i.e., QMEAN scoring function [137]. While it is a generic approach, it has been used for the modelling of different classical antibodies [138–141].

Their in-house software ProMod3 was recently provided freely to download as a standalone package [142]. It can be used to identify structural templates after sequence alignments for a given V_HH sequence. Then, it produces multiple homology models in batch mode.

SWISS-MODEL was used to generate structural models of humanised V_HHs. Murakami and co-workers used SWISS-MODEL directly to generate their humanised V_HHs and later relaxed the 3D structural models with advanced molecular dynamics (MD) simulations. They previously classified CDR3 conformations into four clusters carried out on MD to see their stability. Indeed, their clusters seemed highly biologically relevant [143]. SWISS-MODEL was also used to generate a structural model of a V_HH that binds to the famous spike protein of SARS-COV-2. This particular V_HH was used to understand the structural basis for the induction of a spike trimer [144]. Noorden and co-workers built a complex of a V_HH and SAN2 protein and provided further comprehension of a nuclear pore complex structure [145]. Thanongsaksrikul and co-workers developed libraries of VHs and V_HHs against botulism. They succeeded in the identification of Abs against the light chain of type A *Clostridium botulinum* (BoTxA/LC) neurotoxin. Using SWISS-MODEL, they obtained VH and V_HH structural models. The interface binding between BoTxA/LC and the selected VH/V_HH was determined by using ZDOCK and RDOCK modules on Discovery Studio 2.1 (Accelrys Inc., now the Biovia-Dassault System). As they noted, "BoTxA/LC neutralization by the V_HH should be due to the V_HH insertion into the enzymatic cleft of the toxin, which is usually inaccessible to a conventional antibody molecule" [146]. Higashida and Matsunaga modelled all their V_HHs with SWISS-MODEL before doing

advanced molecular dynamics, especially to look at the important parameters to properly gain the inner flexibility of V_HH CDR3 [147]. Orlov and co-workers also modelled a V_HH with SWISS-MODEL to gain the structural basis of V_HH recognition of grapevine fanleaf virus implicated in the virus resistance loss. The proposition of this structural model allows a better understanding of the epitope and helps to design experiments to confirm it [148]. Mahajan and co-workers used SWISS-MODEL V_HH structural models (refined by molecular dynamics) in structure-based computational methods to optimise the binding affinity of the non-amyloid component of human α-synuclein, a natively disordered protein implicated in the pathogenesis of Parkinson's disease (with a large number of experiments) [149].

2.7. MoFvAb

MoFvAb was published in 2015 [150] and focuses on "Modelling the Fv region of Antibodies". In the first step, it annotates FRs and CDRs using the WolfGuy numbering system [151]. Based on the segments, templates are identified for heavy and light chains. For all chains, rigid models are produced, except for CDR-H3. The latter undergoes de novo building. In the next step, the CDRs are relaxed, and a side chain refinement is performed based on the neighbourhood algorithm. The final steps involve VH/VL orientation prediction and refinement, rotamer optimisation and CHARMM minimisation [150]. The methodology was tested on the AMA-I and -II datasets with good results but was never applied on V_HH. It might be adapted for V_HH, but the approach is not currently available.

2.8. Prediction of Immunoglobulin Structure

These web servers were developed by Anna Tramontano's group. Anna Tramontano worked previously on antibody structures, sequence–structure relationship and modelling using classical approaches [152–155]. The first version named "Prediction of ImmunoGlobulin Structure" (PIGS) was published in 2008 [20], as the authors found that WAM had numerous limitations and was not very flexible. On the basis of the results from previous research, CDRH3 was modelled differently according to its length [156]. PIGS proposed structural models of the complex VH/VL for the user-provided sequences. The evaluation of the approaches was tested in the original paper but also in AMA-I and -II competitions, with good results [17,18]. The second version, named PIGS PRO, was proposed in 2017 [157]. As WAM, it can be divided into five steps:

1. Frameworks are prepared: Structural templates for the frameworks are selected using sequence identity with protein structures from PDB.
2. Five of the six CDRs are built: CDRs L1–L3 and H1 and H2 are modelled by getting conformations from antibodies with the same canonical structures discarding the identity sequence.
3. To complete, a CDRH3 loop is proposed using the structural template with the highest sequence identity with the query sequence.
4. All is merged: The complex VH–VL is modelled.
5. At last, SCWRL is used to optimise the side chain conformations of the predicted model.

Interestingly, the user had the possibility of providing a V_HH sequence instead of a VH sequence by putting in a "fake" canonical VL sequence to initiate the prediction job. This V_HH/VL structural model could have provided an interesting approach. The web server is not currently online.

2.9. AbodyBuilder

ABodyBuilder is a dedicated tool developed in Deane's lab for antibody structure prediction. The algorithm first annotates and finds templates for VH and VL alone and for the complex VH–VL [158]. A new update has been deployed, and it is now also usable for V_HH alone [159]. The approach is hierarchical, as often seen. The FREAD algorithm tries to identify the templates for CDRs. If no templates are found, then the Sphinx loop is

triggered ab initio. The final step is to use PEARS to predict the side chain conformations using the IMGT position-dependent rotamer library. An interesting feature is the estimation of the expected RMSD with a confidence value of the resulting antibody model [158]. As a valuable example, this method has been used recently for the analysis of SARS-CoV-1-specific V_HH to apprehend the conformational diversity of the CDR region, its affinity and stability [160].

2.10. LYmphocyte Receptor Automated Modelling

LYmphocyte Receptor Automated modelling (LYRA) was an improved approach to predict B- and T-cell receptor structural models [161], presented as an online web server (https://services.healthtech.dtu.dk/service.php?LYRA-1.0, last accessed date: 8 February 2022). The authors have done important developments regarding curation both in terms of sequences and structures. The sequences have been annotated and numbered with the Kabat–Chothia approach adopted by Abhinandan and Martin [162] and reduced in terms of redundancy. Research of the compatible optimal sequences was done using a Hidden Markov Model [163]. The best-scoring profile is used to deduce the receptor (TCR or BCR) and chain type (heavy or light). At this stage, the alignments are recertified to correctly identify the heavy and light chains. An important point is the generation of different variations of the scenario by looking at (most of) 20 different structural templates to ensure there are no clashes. Similarly, to refine it, CDRs are searched on a defined library based on a specific CDR clustering [164]. Finally, to create the final model, the templates from both chains are assembled, and the side chains are repacked [161]. This job process is fast and takes less than a minute on average. The user can replace the heavy sequence with the V_HH sequence and get a model for this V_HH sequence in a similar way to PIGS and PIGS PRO.

2.11. Phyre2

Phyre [165] and its successor Phyre2 [166] are generic approaches very similar to techniques such as SWISS-MODEL that combine evolutionary information with a dedicated analysis of the protein structure dataset. It is an available online web server (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index, last accessed date: 9 February 2022). Phyre2 has two modes: a normal and an intensive mode.

The normal mode consists of three steps:

1. Related sequences (with or without structures) are detected by HHblits [135] and HHsearch [167] to produce multiple-sequence alignment and to search for adequate templates for the input sequence. Once the templates are identified, the models are constructed but roughly only with the backbone.
2. A specific step is done to optimise the loop modelling where the indels (insertions and deletions) in these models are found.
3. The side chains are grafted to build the final Phyre2 model.

The intensive mode has two steps. The first step produces many alternative models of the same query with the basic mode. Different metrics are then tested related to the confidence interval and coverage. When a region is considered not covered by "homologs", then the Poing algorithm enters the stage to predict the structure using ab initio components. This algorithm mimics a virtual ribosome by performing a folding process only with Cα atoms. It focuses at first on repetitive structures, then the loops [168]. At the end, the backbone is completed with the Pulchra algorithm [169], and the side chains are finally added with R3 [170], such as for the normal mode.

Different V_HH structural models were generated with Phyre2. A first example implies Urease C (UreC) of *Helicobacter pylori* has an important role to play in bacterial colonisation of the gastric mucosa. Loss of its activity has been speculated to arrest *H. pylori* colonisation, i.e., a pertinent target for therapeutic intervention. Different clones of V_HH were produced against UreC. A high-affinity V_HH, called HMR23, was optimised and selected for further analysis. To understand the difference of binding properties between parent anti-UreC

V_HH and HMR23, structural models were generated using Phyre2. The modelling of parent and mutant V_HHs was essential to understand the structural conservation in terms of fold in the latter [171].

A second example focuses on *Acinetobacter baumannii*, a multidrug-resistant bacterial species responsible for many hospital-derived infections. Its ability to form biofilms helps its survival in hospital conditions. Inhibiting the formation of biofilms can contribute to the reduction of infection. In a previous study, researchers detected an efficient V_HH clone against the Biofilm-associated protein (Bap) of this bacterium [172]; the purpose of this study was to go further using 3D structural information. An interesting point to notice is the use of multiple methodologies to propose the structural models. They were produced with (i) Phyre2 [166], (ii) the Protein Structure Prediction Server (PS^2V^2) [173] that builds the models with Modeller [84] and (iii) LOMETS [174]; the last two approaches are methodologically equivalent. A total of ten models for each template were generated with ten distinct structural templates. Model evaluations were done using ProSA [175,176] and then further refined using Modrefiner [177] to select the best ones. It was one of the few studies where multiple approaches were really tested. Then, they performed a docking analysis with ZDOCK [178,179]. Hence, the authors defined important V_HH residues and their interactions for the ligand-binding site. They proposed binding modes between Bap and the V_HH. Finally, the functional residues in the largest cleft implied in ligand binding were identified [180]. An identical approach was used for a V_HH against chronic inflammatory disease caused by *Porphyromonas gingivalis* [181].

2.12. RaptorX

RaptorX is an automated server that performs template-based tertiary structure modelling [182,183]. It derives from RAPTOR, a classical threading approach [184,185]. The improvement of RaptorX versus RAPTOR consists in the use of a nonlinear alignment scoring function with a conditional random field (CRF), providing local entropy. In the first step, the sequence is cut into domains, and the predictions of different features, such as disorder prediction, are performed. Then, the templates are searched and ranked through a threading process. The alignment quality is assessed by Artificial Neural Networks. At the end, RaptorX provides multiple final models, including models that are produced based on multiple top-ranked templates. The latest version of RaptorX has now evolved and includes a deep convolutional neural field (DeepCNF) to predict the secondary structure [186]. DeepCNF has two components: a deep convolutional neural network (DCNN) and a conditional random field (CRF) for the input and label (output) layers that are CRFs. It has recently been enriched by the use of deep learning to predict protein contacts [187]. RaptorX [182,183,188] was used in exciting research on *Bordetella pertussis*. This aerobic, non-spore-forming, Gram-negative coccobacillus was implicated in the renewed outbreak of whooping cough in the elderly. Its adenylate cyclase-haemolysin toxin (CyaA) plays an important role during the early phase of infection. A specific subdomain named CyaA-RTX is involved in toxins binding to target cells. Through the screening of a VH/V_HH phage display library, two VHs and two V_HHs clones were identified after several optimisation rounds [146,189–191]. Three-dimensional structural models of these proteins were built using RaptorX [192]. The loop conformations were obtained using the dedicated FALC loop modelling web server [193]. The best models were further validated using various algorithms available on the NIH SAVE server. Then, the best resulting models were well-refined and optimised by energy minimisation performed with Gromacs software [194]. Afterwards, in order to predict their mode of interaction, the V_HH models were docked on the target CyaA-RTX domains with the ClusPro web server [115]. Interestingly, all the nanobodies were found to interact with the target, especially in the linker region between the two domains CyaA and RTX, i.e., in excellent accordance with the experimental data.

2.13. Rosetta, Robetta, Rosetta Antibody and V_HH Modelling Application

Developed under the guidance of David Baker, the Rosetta story began in the last century. Originally, it mainly consisted in an ab initio approach [195], but progressively, it evolved towards a de novo strategy that combines small structural fragments obtained from a sequence search in a dataset. The addition of evolution information and local protein constraints improved the quality of the results [89,196,197]. A last round of various improvements led to erasing non-native contacts. Finally, an intensive production of alternative structural models and the clustering of models coming from homologous sequences contribute to make Rosetta a great successful tool in the proposition of structural models [198,199]. Rosetta development was impressive and can be analysed through different prims. RosettaCM, dedicated to optimise comparative modelling steps, competed with the best similar approaches [199]. A web server named Robetta is freely available (https://robetta.bakerlab.org/, last accessed date: 10 February 2022) [200,201].

This powerful tool is also freely downloadable for academics and can be installed locally. Rosetta software has evolved a lot since its beginning, being highly scriptable and customisable [93]. Besides 3D predictions, it encompasses extremely valuable services, e.g., protein design [202] or protein docking [203,204], which are of particular interest in the context of V_HH exploration. Rosetta was applied for the study of an anti-glycoprotein (cAbAn33), anti-lysozyme (cAbLys3) and anti-prostate-specific antigen (cAbPSA-N7). The contribution to the stability of these proteins of an extra-disulphide bond (between CDR1 and CDR3) present in camels and llamas was more precisely explored. Fortunately, cAbAn33 and cAbLys3 have available protein structures that were used for creating models of mutants in which the extra cysteines CDR1 and CDR3 were replaced. ESyPred3D [200], i.e., a methodology based on Modeller and Robetta [200], was used to generate the mutant models. The authors mentioned that the models from ESyPred correspond to those generated by Robetta (data were not shown). Hence, molecular modelling in this study has been useful in understanding the structural conservation of mutants in terms of fold. The models of the mutants predicted had very few differences compared to their parent V_HH structures [205]. From the generalist Rosetta software, a specific antibody development was done, namely RosettaAntibody [206]. This tool is available online and can be used through Rosie (https://rosie.rosettacommons.org/antibody/, last accessed date: 10 February 2022). The protocol consists of three steps for a given antibody sequence [21,207,208]:

1. In the initial step, the CDRs are identified using the Kabat CDR definition, and the residues are renumbered using the Chothia scheme. The template selection is then carried out for all the frameworks, and five of the six CDRs (CDRH1 and CDRH2 and CDRL1–CDR3),
2. From the selected templates, a preliminary model is created using homology modelling, as it was shown to be more accurate than a completely de novo approach.
3. CDRH3 de novo loop modelling completes the model prediction, along with the optimisation of the VH–VL interface.

As the Rosetta modelling suite also incorporates a docking approach, this last was optimised for antibodies [209]. However, RosettaAntibody appeared less efficient with single-chain antibodies and, so, with V_HHs. Accordingly, a specific adaptation in the loop definition was done [210] to better take account of the specificity of the V_HH CDRs, especially CDR3 [211]. Until 2021, it was the only specific development and assessment dedicated to V_HHs. This study underlines the difficulty in V_HH modelling and, specifically, the fact that an approach for IgG is not optimal for V_HH. For a while, the specific script for V_HH was not available in the new versions of RosettaAntibody, but it is now usable again. Even if the tool is not used directly, it has been frequently cited as an analysis of CDRs that is quite precise, i.e., for optimising synthetic V_HHs [212].

2.14. AbPredict2

AbPredict is implemented in the Rosetta modelling suite through two versions: ABPredict 1 [213] and ABPredict 2 [214]. It can be downloaded (http://abpredict.weizmann

.ac.il/bin/steps, last accessed date: 14 February 2022) and is also free for academia. In short, AbPredict exploits known antibody 3D structures and optimally combines backbone conformations with a low-energy Monte Carlo search. A few details are given in the next section.

AbPredict methodology is based on five independent databases comprising backbone torsion angles of the segments of VH, VL, CDRH3, CDRL3 and rigid body orientation of heavy and light chains. The initialisation starts by combining random segments from the five databases. A biased simulated annealing Monte Carlo (SA-MC) sampling is performed, with thousands of independent trajectories. The lowest energy conformation is extracted from each of these trajectories. From the latter list of conformations, the final predicted antibody model is the one that has the lowest energy.

AbPredict has been benchmarked using the AMA-II antibody set and compared to the methods presented therein. It performed in the upper third of all the compared methods [208]. It was recently compared to RosettaAntibody and RosettaCM, a generic version, in the case of long CDRH3. Surprisingly for these very hard cases, RosettaCM was slightly better than the two other dedicated methodologies [215]. Thus, despite some success in Ab modelling, it seems that ABPredict is not used for V_HH modelling nowadays.

2.15. Biovia Discovery Studio/Antibody Modelling

Discovery studio (DS) is a commercial toolbox coming historically from the Accelrys company located at San Diego, CA, USA (a merging of Molecular Simulations Inc., Synopsys Scientific Systems, Oxford Molecular, the Genetics Computer Group (GCG), Synomics Ltd. located at Witney, Oxfordshire, United Kingdom and some others) and is now owned by Dassault Systèmes Company. It was renamed BioVia. BioVia Discovery Studio uses Modeller at the backend to model protein structures from a user-provided sequence [84]. DS proposes different strategies based on homology modelling but with dedicated tools to model the antibodies named Biotherapeutics and Antibody modelling. They were positively tested in AMI-I and -II and validated [216].

The first and simplest strategy (named single template) consists of finding the best templates (based on the sequence identity) for the heavy and the light chains of a given antibody sequence. Both of these templates should belong to the same antibody structure. Spatial orientation or the interface between the heavy and the light chains will also be used as a template from the same antibody.

In contrast, in the second strategy (named as chimeric), the best templates for the heavy and light chains and the interface between the chains can come from different antibody structures. After identifying the top five best templates, five models are built using Modeller [84]. If the sequence identity between the template and the target is too low (inferior to 10%), the template is discarded. This initial step is followed by the CDR refinement of the top five models. The CDR regions are identified using the IMGT, Chothia, Honegger or Kabat numbering schemes [101]. The CDR templates are searched and ranked by sequence similarity according to the BLOSUM62 matrix and, finally, rated based on the crystallographic resolution. The CDR loops are built with Modeller using the best-ranked templates while maintaining the frameworks in these models.

Two V_HH modelling applications have been done using DS. Jullien and colleagues worked on histones, proteins implicated in the nucleosome complex. Some posttranslational modifications of histones are indicators of gene expression. This feature can be exploited by binding exogenous proteins to these histones, which helps to understand the chromatin dynamics. In this study, V_HH rose against the H2A–H2B complex in the chromatin, i.e., named chromatibody, and were experimentally tested and automatically modelled using Modeller from Discovery Studio. The best models were selected based on the Modeller scoring function (molpdf) and DOPE scores [107] and then refined. This molecular modelling has been helpful in suggesting the appropriate placement of the chromatin-binding motif in the β-hairpin region of the CDR3 into the H2A–H2B protein cavity (Jullien et al. 2016).

The second case dealt with a V_HH that targets Phospholipase A2 (PLA2), an enzyme present in snake venom that digests cell membrane lipids. The number of snakebite cases, especially due to monocled cobras (*Naja kaouthia*) in Thailand, is on the rise. Hence, an antibody against PLA2 has practical interest, especially a camelid V_HH, due to its thermostable properties. A humanised dromedary V_HH phage library was previously obtained [146]. Two V_HH clones from PLA2 bio-panning (P3-1 and P3-3) were selected for further studies of modelling and docking. Both V_HH clones are classical V_HHs, along with P3-3, having an extra CDR1–CDR3 disulphide bond. The templates were selected using BLAST against PDB, and the structural models were built with DS. The steric hindrance of the models was assessed and, also, geometrical properties with the Ramachandran plot. The best models were selected for docking with PLA2 using ZDOCK. The docking experiments revealed that the V_HH models bind to the Ca^{2+}-binding site, the active site and the phospholipid-binding site of PLA2 [191]. Despite their interest, these findings have not yet been experimentally validated.

2.16. MOE/Antibody Pipeline

MOE belongs to the Chemical Computing Groups Company, located at Montreal, QC, Canada, and has a pipeline for antibody and biologics designs [217]. The tool works for classical antibodies and V_HH. For a V_HH sequence, MOE searches for templates in two databases: in-house-built and PDB databases. In the in-house-built database searching process, the template will be kept if the sequence similarity with the target is more than 50%. From the in-house-built and PDB databases, the top 10 and top five hits will be retained, respectively. The templates are then constructed by grafting the frameworks and the CDRs. Multiple models are built, and then, an energy minimisation is performed. From this group of models, a consensus model is established, and different optimised final models are proposed.

With the final model, CDR3 conformations are further explored and, lastly, clustered. A unique CDR3 conformation is finally grafted in the V_HH model. Once again, an energetical evaluation is done. The top three models with the lowest binding energies will be the final output of this program [217].

Different studies have been made with classical antibodies. A recent study combined IgG and V_HH as bispecific antibodies. MOE structural models were tested by molecular dynamics and analysed in light of the experimental data. They could compare different V_HH behaviours and especially observed problematic interactions that were too strong between one IgG chain and one V_HH [218].

2.17. Schrödinger BioLuminate and Antibody Pipeline

BioLuminate is a suite developed by Schrödinger Company, located at New-York, NY, USA. It is a pipeline for general homology modelling but includes a specific design for antibodies. The modelling protocol is:

1. Frameworks are detected in a curated antibody database, providing structural templates (selected using a sequence identity).
2. Next, a set of CDRs are grafted after scanning another custom database containing only CDRs and a selection based on structural clustering, sequence similarity and stem residue geometry matching.
3. In the final step, the in-house Prime software repacks the side chains and minimises the antibody model.
4. A CDR3 loop is built using the ab initio method.

Schrödinger does provide detailed information about BioLuminate's Antibody modelling tool and tutorials: (https://www.schrodinger.com/training/building-homology-models-multiple-sequence-viewereditor214, https://www.schrodinger.com/training/antibody-visualization-and-modeling-bioluminate-workshop-tutorial214, last accessed date: 15 February 2022).

Their antibody pipeline was used in a study on *Clostridium difficile* Toxin (CDT). This last is a potent toxin responsible for antibiotic-associated diarrhoea. These toxins belong to the C2 class of toxins, which means the presence of two subunits (CDTa and CDTb). The CDTa subunit is responsible for the ADP-ribosylation of actin, and CDTb is responsible for forming a toxin–pore complex through which the CDTa subunit is internalised. This study attempted to inhibit the function of the CDTa subunit using V_HH from lama. Three anti-CDT V_HH clones were modelled using the BioLuminate package. Two of the clones had an extra disulphide bond that was preserved in the structural models. These models were used in a docking experiment to understand their precise mode of interaction with their target. The models were docked onto CDTa using the PIPER module in Discovery Studio. All three V_HHs were found to bind to the NAD^+-binding cavity in CDTa. Two compete with each other, whereas the last has a different binding site. This study also reported interactions in FRs of the V_HH used to lead to a better understanding of the experimental data.

2.18. Macromoltek's SmrtMolAntibody

The Macromoltek Company, located in Austin, TX, USA, created SmrtMolAntibody. The first step of this algorithm is to search an antibody database to detect templates for the VH and VL query sequences with the BLAST approach [219]. Then, a similar search is done on a loop database to find more adequate templates for all six CDRs. Based on the identified templates, assembling the frameworks and grafting the CDRs allow building an initial antibody model. A specific modelling step is done for the CDR(H)3 loop. The top 50 models with the best (energy) CDRH3 loop conformations are retained. For each of these 50 models, the side chains are repacked, and their torsions are minimised. The above-mentioned final steps are subject to multiple evaluations based on an energy score with modified and softened Lennard-Jones associated with side-chain rotamer frequencies [219].

2.19. I-TASSER and C-I-TASSER

During multiple rounds of CASP, with David Baker's Rosetta, I-TASSER [90] was the best available method to propose pertinent structural models [220]. I-TASSER is based on Local Meta-Threading Server (LOMETS) [221]. LOMETS exploits deep learning-based threading methods and profile-based programs to identify the best templates and to annotate structure-based protein functions. Once the templates are identified, I-TASSER uses a Monte Carlo-based simulation to produce and to refine the protein structural model. The latest development includes deep learning-based contact predictions in the I-TASSER methodology, which is now named C-I-TASSER [222].

The I-TASSER web server was used to model different V_HHs. A first study looked at the hepatitis C virus (HCV). This virus has six non-structural proteins (NS). Fusion proteins NS3 and NS4 form a serine protease, which cleaves the linker between NS5A and NS5B. This last is directly implicated in the virus life cycle and is a good target. Three V_HH clones were generated against the fusion HCV protease NS3/4A. The I-TASSER web server was used to build the three structural models. They were refined using ModRefiner and fragment-guided molecular dynamics [177]. Computational docking analyses revealed that CDR2 and CDR3 of the three V_HH would bind to the NS3/4A catalytic triad residues. Two V_HHs residues from FRs were also implicated in interactions with NS3/4A [189]. This analysis brought up a plausible explanation for better understanding the experimental data previously obtained.

The second example looks at Staphylococcal Protein A (PrA). It binds to the FC and sometimes Fab regions of antibodies and makes it an attractive tool for the in vitro isolation of antibodies. This study attempts to understand the binding of PrA to V_HHs. Fridy and collaborators designed V_HHs with minimal CDRs to remove their contribution in binding to PrA. The LaP-1 (llama antibody against Protein A) V_HH was optimised, and a few mutants were selected [223,224]. The I-TASSER web server was used to model the structures of these V_HHs. The best models were chosen based on the c-score of I-TASSER.

The models provided insights on the PrA binding to antibodies and, also, into the structural stability of the mutants and the parent V_HH [225].

2.20. QUARK and C-QUARK

QUARK is a reference ab initio structural prediction strategy [226]. It is based on a force field composed of 11 energetic terms, classically representing three resolution levels (atomic, residue and topology). The strategy can be described in three stages.

Based on the input sequence, the secondary structures, torsion angles and solvent accessibility and generated fragments are predicted. Then, Multiple Replica Exchange Monte Carlo (REMC) simulations, using a semi-reduced protein model, are performed to build multiple alternative conformations. The final step uses SPICKER for clustering [227] and performs refinement on the full-atom structure.

The latest version of this software is named C-QUARK [228,229]. When compared to QUARK, two steps are added at the start: (1) DeepMSA-generated multiple sequence alignment [230] and (2) contact map prediction based on deep learning methods. Even if QUARK and C-QUARK performed well in CASP competitions [231,232], they were not used for antibody and V_HH modelling.

2.21. AlphaFold 2

AlphaFold was developed by the industrial laboratory DeepMind [233]. It joined for the first time the CASP competition in 2018 (CASP13) [234] and won the Free Modelling category [235], i.e., to predict novel protein folds. The Template-Based Modelling category, i.e., protein folds already in the PDB, was won by Zhang's group [90]. Two years later, they created a shockwave by greatly improving the quality of their methodology [91] and clearly won this CASP competition. Its inclusion of advanced deep learning approaches coupled with the GPU power of DeepMind allowed a fantastic success [236,237]. It reduced the percentage of the dark human proteome [238] and opened opportunities for drug design [239].

AlphaFold 2 uses multiple sequence alignment (MSA), residue pairing information and structural templates for a given sequence that should be modelled [91]. A transformer, called Evoformer, processes all of these inputs. An important innovation point about this latter is that it allows information exchange between the MSA and residue pairing blocks. The authors emphasised that this Evoformer block helps in establishing a rational between the spatial and the evolutionary information of residues. After passing through Evoformer, the resultant information is processed by structural blocks that output the coordinates directly. One particularity is that the intermediate MSA, residue pairing information and the predicted structure are reinjected into the Evoformer blocks. By default, these intermediate results are reintroduced in the network three times, which improves the predictions. AlphaFold 2 can be downloaded from a GitHub repository (https://github.com/deepmind/alphafold, last accessed date: 9 February 2022). However, users need good computers with performing GPU cards and a good amount of memory. Interestingly, an online Jupyter notebook can also be used (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/batch/AlphaFold2_batch.ipynb, last accessed date: 9 February 2022). It produces five models by default [240]. For each residue in the model, a confidence score is provided, and it is used to rank the models among them.

AlphaFold 2 was a hot topic for 2020 and 2021, leading to a revolution in protein structural model building [241]. However, AlphaFold 2 does not always give 100% correct/meaningful predictions, i.e., some globular proteins are still not properly modelled [242], and transmembrane proteins are not close to a near-native structure, as they are hard targets. Antibodies were not specifically tested, but different examples showed, as seen for Rosetta [207], that classical comparative modelling still performs better. It is so logical that no V_HH modelling has been performed at this time.

2.22. RoseTTAFold and DeepAb

Deep learning has made a paradigm change in structural bioinformatics [243–248]. Therefore, the famous Rosetta has been upgraded and now leverages these new advances [201]. trRosetta is based on deep learning approaches with an all-atom energy function [249] combined with inter-residue distance constraints and orientation distributions [92,250].

Based on the properties and successes of AlphaFold 2 [91], Baker's team extensively studied the different neural network architectures and proposed a new deep learning-based methodology called RoseTTAfold [251]. It is based on a three-track (1D, 2D and 3D) neural network, which can simultaneously process multiple sequence alignments (MSA), inter-residue contacts and refinement of the predicted structure. The connections in this network allow efficient learning about the relationships between the protein sequences, distances and the coordinates. The interlinked 1D and 2D neural networks take and process cropped MSA and templates as the input. The 3D track first provides a backbone-only model, and then, SE(3)-transformer performs an iterative refinement by using the inference of a relationship between the sequences and the templates. By the end of this stage, the user will have a full-atom model. The Robetta server with RoseTTAfold methodology provides predicted models with a confidence index (Å error estimate per position), which is used to rank the five final models [251]. As for AlphaFold 2, since the approach is extremely recent, no use in predicting the structure of V_HHs has been observed.

While preparing the submission of this review, DeepAb was published [252]. It can be considered as the natural evolution of RosettaAntibody with deep learning RoseTTAfold. The algorithm can be divided into two main steps. The first part has a deep residual convolutional network, which consists of 1D ResNet, a Bi-Long short-term memory (Bi-LSTM) repertoire encoder and 2D ResNet. The 1D ResNet and Bi-LSTM encoder processes structural and evolutionary features. The 2D ResNet will output inter-residue distances and orientations. In the second step, Rosetta minimisation is used with distance and angle constraints. Concerning V_HH modelling, Ruffolo et al. observed that CDR3 resembles that of CDRH3 [252]. Nevertheless, they suggest that the methodology should be refined with a database dedicated to V_HHs. In the present stage, DeepAb seems to be more specific for classical antibodies than V_HHs.

2.23. NanoNet

NanoNet was proposed to the scientific community in August 2021 [253]. It is truly the first optimised V_HH approach. This deep learning approach was trained with classical antibodies and V_HHs, as a large amount of data is needed to train the neural network and obtain relevant results. Its architecture is made by a convolutional neural network (CNN) with two 1D residual neural networks (ResNet). The first ResNet learns the frameworks and CDR hypervariable loops, while the second will apprehend inter-residue interactions. After a dropout layer to avoid overfitting, the last layer outputs C-alpha coordinates for a given input V_HH sequence. Full backbone and side chains can then be constructed using Pulchra [169]. This program is available online (https://github.com/dina-lab3D/NanoNet, last accessed date: 13 February 2022). In our experience, it took less than 15 s to predict a full atomic structure of a V_HH using a computer with 8 GB RAM, 4 CPUs and an Intel i5 processor. No particular library is required for the use of this program.

The assessment of NanoNet was done against AlphaFold 2 with 16 V_HH deposited in the PDB in 2021, i.e., absent from AlphaFold 2 training. NanoNet performed better with a mean RMSD of 2.69 (±1.49) Å vs. 3.23 (±2.49) Å for AlphaFold 2 (and 1.57 (±0.41) Å vs. 2.04 (±2.09) Å on CDR3, reps.). Similar outcomes were obtained with 37 V_HHs when confronted with the Rosetta Antibody modelling suite. NanoNet performed better with a mean RMSD of 1.68 (±0.57) Å vs. 2.71 (±1.13) Å for Rosetta Antibody (and 2.99 (±1.48) Å vs. 5.73 (±2.33) Å on CDR3, resp.).

Hence, NanoNet appears to be a very interesting and promising new tool dedicated to V_HH with excellent results. It has been, in particular, used to optimise CDR3 loop predictions (associated with the experimental data) in optimising anti-SARS-CoV-2 V_HHs [254].

3. Discussion

In the previous sections, we listed the different methodologies, provided their specificities and gave examples of V_HH structural modelling. They are all very different and mainly used for docking purposes. However, it is highly difficult to precisely know the quality of each of these approaches/results. Indeed, they are all well-established approaches, but no real benchmark was performed until now, and the evaluations are limited to specific and extremely rare cases. A few examples are detailed hereafter.

(i) The dedicated development of V_HH scripts for the RosettaAntibody modelling suite is as expected associated with a specific evaluation [211] but mainly focuses on classical RosettaAntibody modelling than another tool [207].

(ii) More recently, as Modeller [84] remains the most used comparative modelling approach, we evaluated the Modeller quality in V_HH modelling [74]. Using 100 different V_HHs, different template selection strategies for comparative modelling using Modeller have been extensively assessed. This study analysed the conformational changes in both the FRs and CDRs using an original strategy of conformational discretisation based on a structural alphabet [255,256]. It showed that, often, multi-template is the best method to obtain a correct V_HH model and that the DOPE score is a relevant measure to select this model [107]. Nonetheless, it remains difficult to propose satisfactory models for some V_HHs. Even sometimes, to use the closest V_HH in terms of the RMSD is not always the best choice to obtain a good model, underlying the possibility for future improvement.

(iii) Finally, NanoNet has demonstrated its superiority over RosettaAntibody and AlphaFold 2 but with a limited number of examples (see the previous subsection) [253].

To better understand this complexity, we decided to evaluate and show a rather classical example of V_HH modelling. We selected a structure of V_HH that binds with lanthanides (PDB ID: 6XYF) [257] deposited in the PDB recent enough to never have been used in any of the previously cited approaches. 6XYF was obtained by X-ray crystallography with a very good resolution of 1.11 Å. Please note that this case is a "classical" case of structural prediction, i.e., representative of V_HH modelling, as others were also tested. Eight different methods were carried out with six different software and online tools: Modeller with single-template and multi-template [84], ModWeb [258,259], SwissModel with the best sequence identity template and a second time with the best GMQE score [134], AlphaFold 2 [91], RoseTTAfold [251] and NanoNet [253].

At first, Modeller was used with the simplest approach, i.e., a single template. A template was selected with the highest sequence identity, i.e., a V_HH domain (PDB ID: 5LMW chain A) used as a crystallisation chaperone for different constructs associated with the type IX secretion system from *Porphyromonas gingivalis* [260]. PDB 5LMW had a sequence identity of 89.9% with the target sequence. In addition, Modeller was used in a multiple-template approach with three templates. We decided to use this method because of its remarkable efficiency, as demonstrated in Reference [74]. Two new V_HH structures were selected: anti-HIV V_HH (PDB ID: 3R0M chain B) [261] and anti-HIV-1 gp120 V_HH (PDB ID: 2XA3 chain A) [262], in addition to the 5LMW template. They also shared high sequence identity with 6XYF V_HH: 89.3% and 88.9%, respectively. Each time, 100 models were generated, and the best model was selected using the DOPE score [107].

The ModWeb server was also used to model the 6XYF sequence. This online job provided one unique model based on a single template (PDB ID: 6XYM [257]) that had an extremely high sequence identity percentage (98%) with the target sequence. Indeed, this V_HH came from the same lab as our test V_HH [257].

These approaches can be easily compared to SwissModel. The first step with this method is to look at the best sequence identity. A V_HH anti-SARS-CoV-1 (PDB ID: 6WAQ chain A) [54] was identified that shares 82.3% sequence identity with our target. An alternative model with the best GMQE score was built with a V_HH anti-Vsig4 (PDB ID: 5IML chain B) with a sequence identity of 80.0%.

Regarding more advanced techniques, deep learning-based methodologies (RoseTTAfold, AlphaFold 2 and NanoNet) were used to model 6XYF. RoseTTAfold predicted five models. Their mean estimated RMS errors (in Å) per position were 1.10 ± 0.87, 1.13 ± 0.82, 1.12 ± 0.8, 1.12 ± 0.79 and 1.13 ± 0.84, respectively (displayed according to their ranks). AlphaFold 2 gave five models. These models had mean confidences (representing the accuracy in terms of the RMS) that were the following: 91.45 ± 10.38, 90.71 ± 11.05, 90.44 ± 11.54, 90.42 ± 11.83 and 89.99 ± 12.16. NanoNet produced only a single model for each query sequence.

Figure 3 shows every best structural model superimposed with the corresponding X-ray structure. The model with the best DOPE score obtained with a single-template Modeller approach showed a RMSD of 2.28 Å (see Figure 3a). This value was slightly improved with the addition of two close V_HHs, dropping to 2.18 Å (see Figure 3b). The template used by the ModWeb server resulted in a model with a RMSD value of 1.97 Å (Figure 3c), but we chose to discard this option, as this template was made by the same research group as the query [257]. Please note that, in our experience, ModWeb sometimes provides incomplete models. Actually, it depends on the pertinence of the identified template and the structural coverage. Concerning SwissModel, the best sequence identity model had the best RMSD of all the comparative approaches, with a value of 1.85 Å (see Figure 3d; the GMQE score and QMEANDisCo Global score [263] were 0.78 and 0.77 ± 0.07, respectively). For the SwissModel solution, considering the best GMQE, its RMSD was 2.05 Å (see Figure 3e; GMQE score = 0.81 and QMEANDisCo Global score = 0.79 ± 0.07).

The results for the deep learning-based techniques were also very relevant. The RoseTTAfold model ranked first had a RMSD of 1.87 Å (see Figure 3g) but was surpassed by the models from NanoNet and AlphaFold 2 with RMSDs of 1.55 Å (see Figure 3h) and 1.62 Å (see Figure 3f), respectively.

According to our case study based on 6XYF, it was demonstrated that NanoNet and AlphaFold 2 predicted the closest models to the experimental structure. This tendency was also found when CDR3 was specifically analysed. NanoNet and AlphaFold 2 ranked first and second (1.81 and 1.98 Å). SwissModel (with the best sequence identity) also performed well here, with a similar accuracy to AlphaFold 2, while the other methods were less efficient, with RMSDs ranging between 2.05 and 3.80 Å.

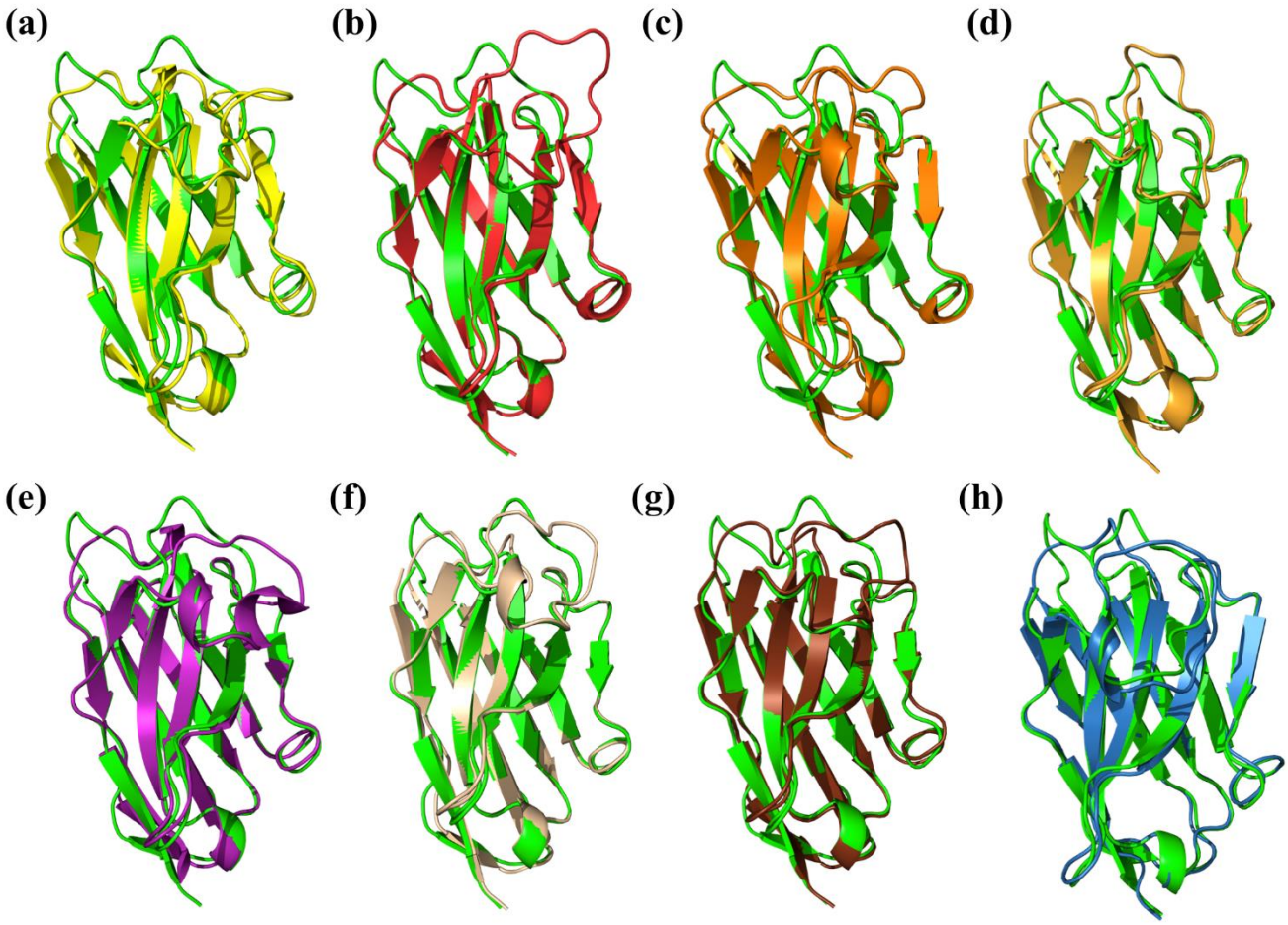

Figure 3. Superimposition of the V_HH structural models with the X-ray structure of PDB ID: 6XYF. The V_HH structure (in green) is superimposed with a structural model from the (**a**) Modeller best model for single-template (in yellow, RMSD = 2.28 Å, RMSD CDR3 = 3.80 Å), (**b**) Modeller best model for multi-template (in red, RMSD = 2.18 Å, RMSD CDR3 = 3.32 Å), (**c**) ModWeb (in orange, RMSD = 1.92 Å, RMSD CDR3 = 2.05 Å), (**d**) SwissModel best model with the best sequence identity (in yellow ochre, RMSD = 1.85 Å, RMSD CDR3 = 1.98 Å), (**e**) SwissModel best model with a GMQE score (in purple, RMSD = 2.05 Å, RMSD CDR3 = 2.75 Å), (**f**) AlphaFold 2 (in wheat, RMSD = 1.62 Å, RMSD CDR3 = 1.98 Å), (**g**) RoseTTAfold (in chocolate brown, RMSD = 1.87 Å, RMSD CDR3 = 2.56 Å) and (**h**) NanoNet (in blue, RMSD = 1.55 Å, RMSD CDR3 = 1.81 Å). Visualisation was performed using PyMOL [1–3].

4. Conclusions and Perspectives

V_HHs have been known for almost 30 years, but in recent years, they have received a phenomenal revival of interest from an experimental point of view. Regarding bioinformatics, it has taken much longer to gain insight into their structures and functions due to inherent limitations. For example, the number of precise analyses of the sequence–structure relationship analyses of V_HH domains, which is required to develop their in silico design and optimisation, is still limited [72,264,265].

There have also been some recent developments of web repositories dedicated to V_HHs. For instance, the Institute Collection and Analysis of Nanobodies (iCAN) has the first collected large number of V_HH sequences coming from scientific publications and patents (http://ican.ils.seu.edu.cn, not operative at the moment of reviewing this paper, last try: 19 February 2022) [266]. Llamanade is a dedicated open-source computational pipeline for robust V_HH humanisation [267]. More recently, an integrated nanobody database for immunoinformatics (INDI) lists V_HHs from all the major public outlets of biological

sequences and patents (http://research.naturalantibody.com/nanobodies, last accessed date: 11 February 2022) from GenBank, next-generation sequencing repositories and structure databases and publications [268]. Even the famous SAbDab has an extension for V_HH with SAbDab-nano (http://opig.stats.ox.ac.uk/webapps/newsabdab/nano/, last accessed date: 12 February 2022) [159]. All V_HH structures could also be mined from the PDB.

Concerning the structural features of V_HHs, specific developments are quite rare. We have shed light on the first optimised V_HH structural modelling approach, NanoNet [253]. Since V_HH docking is a key point in the development of those methods, a dedicated reranking approach of the V_HH docking poses has been proposed with interest, namely NbX [269], but unfortunately, no tool is currently available to test it. Similarly, they can be positively integrated into epitope prediction approaches, as developed in Reference [270]. For a few years, the number of V_HH structures deposited in the PDB greatly increased (see Figures 4 and A1), which confirms the increasing interest in these small protein domains.

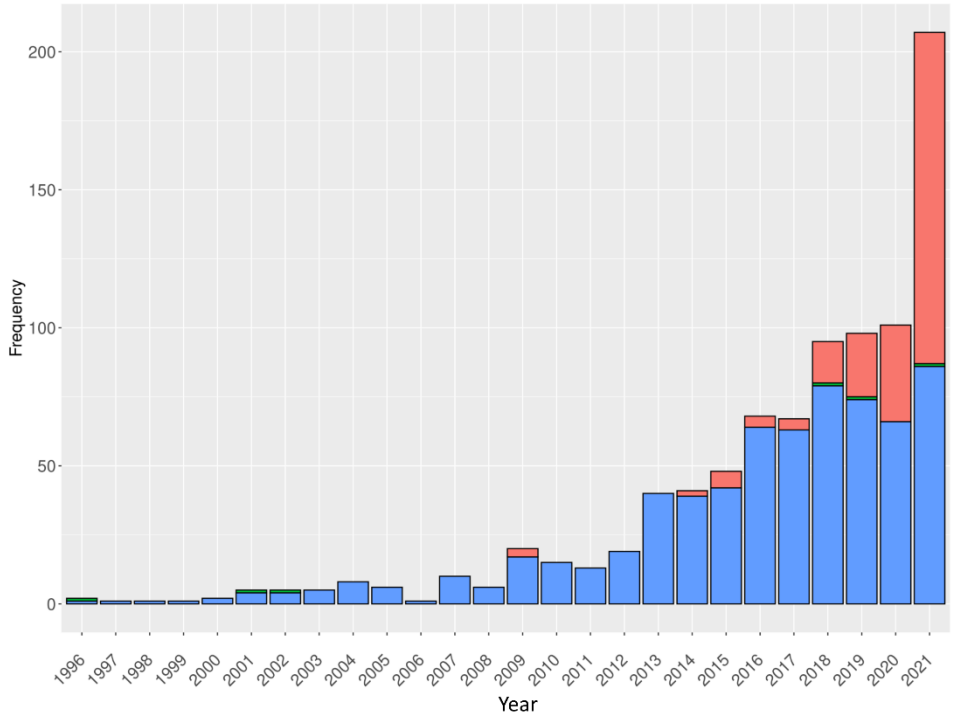

Figure 4. Number of V_HH structures deposited in the PDB per year. Structures solved using different experimental methods are colour-coded accordingly (electron microscopy, solution NMR and X-ray diffraction has been represented in red, green and blue respectively). Data extracted from Reference [159].

Besides 3D model production, there is also an interesting number of molecular dynamics simulations of V_HHs that are often used for their design optimisation [271–275]. Yet, there is no dedicated approach to globally apprehend their dynamics. An important point is that their CDRs are supposed to be variable and flexible. Moreover, molecular dynamics simulations are also interesting for going further with antibodies [276], as stated in his pioneer works by Nobel Prize Martin Karplus [277,278].

In this review, we have seen the limitations of our knowledge, i.e., the difficulty to apprehend the ability of each methodology for V_HH modelling. The number of scenarios used to build V_HH structural models is impressive and is often successful for apprehending atomistic-level epitopes–paratopes docking and experimental data. A more extensive modelling assessment would be good, especially with recent deep learning approaches such as AlphaFold 2 [91] and NanoNet [253], since multiple examples have shown how pertinent they might be.

Author Contributions: Conceptualisation, A.A.N. and A.G.d.B.; methodology, A.A.N. and A.G.d.B.; formal analysis, A.M.V., A.A.N. and A.G.d.B.; resources, P.V., N.S., J.D. and A.G.d.B.; data curation, A.M.V., C.M., F.C., A.A.N. and A.G.d.B.; writing—original draft preparation, A.A.N. and A.G.d.B.; writing—review and editing, P.V., J.D., C.E., A.A.N. and A.G.d.B.; visualisation, C.M., C.E., A.A.N. and A.G.d.B.; supervision, A.G.d.B.; project administration, F.G. and A.G.d.B. and funding acquisition, A.G.d.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the POE FEDER 2014-20 of the Conseil Régional de La Réunion (S3D VHH program, no. SYNERGIE RE0022962), EU-H2020 and Université de la Réunion. This work was supported by grants from the Ministry of Research (France), Université Paris Cité (formerly University Paris Diderot, Sorbonne, Paris Cité, France and formerly Université de Paris), Université de la Réunion, National Institute for Blood Transfusion (INTS, France), National Institute for Health and Medical Research (INSERM, France), IdEx ANR-18-IDEX-0001 and labex GR-Ex. The labex GR-Ex, reference ANR-11-LABX-0051, was funded by the program "Investissements d'avenir" of the French National Research Agency, reference ANR-11-IDEX-0005-02. A.G.d.B. acknowledges the Indo-French Centre for the Promotion of Advanced Research/CEFIPRA for a collaborative grant (number 5302-2). A.G.d.B. and J.D. acknowledge the French National Research Agency for grant ANR-19-CE17-0021 (BASIN).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank Anne-Elisabeth Molza and Madjid Zemmouche for the fruitful discussions and Lynda Saminadin, Juan-Manuel Mora-Rey and Damien Guimond for the administrative support. We would also like to take this opportunity to thank all the researchers who have worked on this class of macromolecules, without whom our work would never have been possible.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of the data; in the writing of the manuscript or in the decision to publish the results.

Appendix A

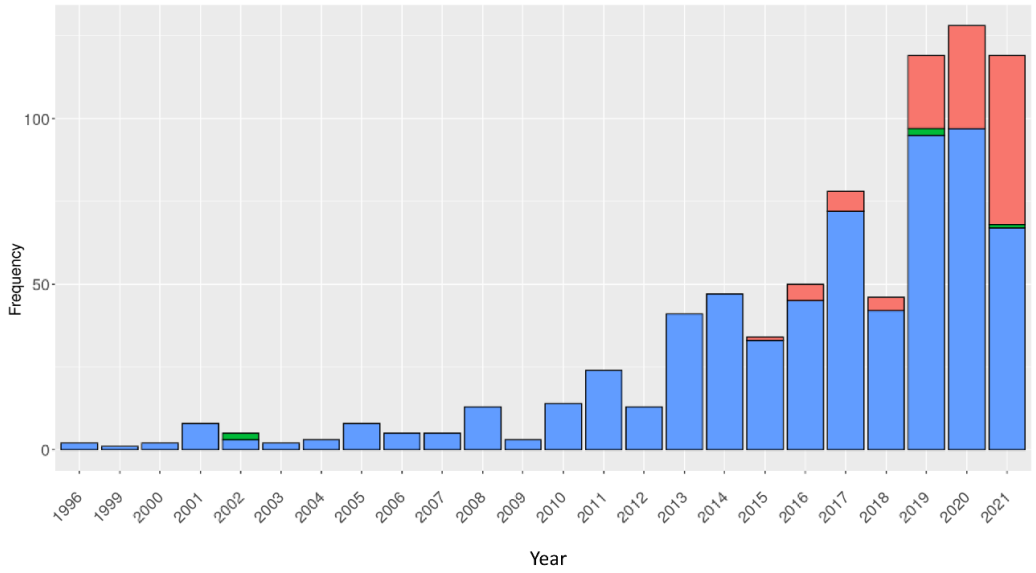

Figure A1. Number of V_HH structures deposited in the PDB per year. Structures solved using different experimental methods are colour-coded accordingly (our own analysis) (electron microscopy, solution NMR and X-ray diffraction has been represented in red, green and blue respectively). It differs from Figure 4, as only nonmodified camelid V_HHs were analysed.

References

1. DeLano, W.L. *Pymol*, Version 2.5; Delano Scientific: San Carlos, CA, USA, 2002.
2. Delano, W.L. The Pymol Molecular Graphics System on World Wide Web. Available online: http://www.pymol.org (accessed on 20 February 2022).
3. Schrodinger, L.L.C. *The Pymol Molecular Graphics System*, Version 1.7.2.2; Delano Scientific: San Carlos, CA, USA, 2015.
4. Berger, M.; Shankar, V.; Vafai, A. Therapeutic applications of monoclonal antibodies. *Am. J. Med. Sci.* **2002**, *324*, 14–30. [CrossRef]
5. Ecker, D.M.; Jones, S.D.; Levine, H.L. The therapeutic monoclonal antibody market. *mAbs* **2015**, *7*, 9–14. [CrossRef]
6. Gao, Y.; Huang, X.; Zhu, Y.; Lv, Z. A brief review of monoclonal antibody technology and its representative applications in immunoassays. *J. Immunoass. Immunochem.* **2018**, *39*, 351–364. [CrossRef]
7. Ma, J.; Mo, Y.; Tang, M.; Shen, J.; Qi, Y.; Zhao, W.; Huang, Y.; Xu, Y.; Qian, C. Bispecific antibodies: From research to clinical application. *Front. Immunol.* **2021**, *12*, 626616. [CrossRef] [PubMed]
8. Linke, R.; Klein, A.; Seimetz, D. Catumaxomab: Clinical development and future directions. *mAbs* **2010**, *2*, 129–136. [CrossRef] [PubMed]
9. Dees, S.; Ganesan, R.; Singh, S.; Grewal, I.S. Bispecific antibodies for triple negative breast cancer. *Trends Cancer* **2021**, *7*, 162–173. [CrossRef] [PubMed]
10. Lenting, P.J.; Denis, C.V.; Christophe, O.D. Emicizumab, a bispecific antibody recognizing coagulation factors ix and x: How does it actually compare to factor viii? *Blood* **2017**, *130*, 2463–2468. [CrossRef] [PubMed]
11. Samaranayake, H.; Wirth, T.; Schenkwein, D.; Räty, J.K.; Ylä-Herttuala, S. Challenges in monoclonal antibody-based therapies. *Ann. Med.* **2009**, *41*, 322–331. [CrossRef]
12. Sapra, P.; Shor, B. Monoclonal antibody-based therapies in cancer: Advances and challenges. *Pharmacol. Ther.* **2013**, *138*, 452–469. [CrossRef]
13. Sifniotis, V.; Cruz, E.; Eroglu, B.; Kayser, V. Current advancements in addressing key challenges of therapeutic antibody design, manufacture, and formulation. *Antibodies* **2019**, *8*, 36. [CrossRef] [PubMed]
14. Keyt, B.A.; Baliga, R.; Sinclair, A.M.; Carroll, S.F.; Peterson, M.S. Structure, function, and therapeutic use of igm antibodies. *Antibodies* **2020**, *9*, 53. [CrossRef] [PubMed]
15. Wu, N.C.; Wilson, I.A. Structural insights into the design of novel anti-influenza therapies. *Nat. Struct. Mol. Biol.* **2018**, *25*, 115–121. [CrossRef]
16. Zhao, J.; Nussinov, R.; Wu, W.J.; Ma, B. In silico methods in antibody design. *Antibodies* **2018**, *7*, 22. [CrossRef] [PubMed]
17. Almagro, J.C.; Beavers, M.P.; Hernandez-Guzman, F.; Maier, J.; Shaulsky, J.; Butenhof, K.; Labute, P.; Thorsteinson, N.; Kelly, K.; Teplyakov, A.; et al. Antibody modeling assessment. *Proteins* **2011**, *79*, 3050–3066. [CrossRef] [PubMed]
18. Almagro, J.C.; Teplyakov, A.; Luo, J.; Sweet, R.W.; Kodangattil, S.; Hernandez-Guzman, F.; Gilliland, G.L. Second antibody modeling assessment (ama-ii). *Proteins* **2014**, *82*, 1553–1562. [CrossRef] [PubMed]
19. Fasnacht, M.; Butenhof, K.; Goupil-Lamy, A.; Hernandez-Guzman, F.; Huang, H.; Yan, L. Automated antibody structure prediction using accelrys tools: Results and best practices. *Proteins* **2014**, *82*, 1583–1598. [CrossRef] [PubMed]
20. Marcatili, P.; Rosi, A.; Tramontano, A. Pigs: Automatic prediction of antibody structures. *Bioinformatics* **2008**, *24*, 1953–1954. [CrossRef] [PubMed]
21. Weitzner, B.D.; Jeliazkov, J.R.; Lyskov, S.; Marze, N.; Kuroda, D.; Frick, R.; Adolf-Bryfogle, J.; Biswas, N.; Dunbrack, R.L.; Gray, J.J. Modeling and docking of antibody structures with rosetta. *Nat. Protoc.* **2017**, *12*, 401–416. [CrossRef] [PubMed]
22. Marchalonis, J.J.; Schluter, S.F.; Bernstein, R.M.; Hohman, V.S. Antibodies of sharks: Revolution and evolution. *Immunol. Rev.* **1998**, *166*, 103–122. [CrossRef]
23. Juma, S.N.; Gong, X.; Hu, S.; Lv, Z.; Shao, J.; Liu, L.; Chen, G. Shark new antigen receptor (ignar): Structure, characteristics and potential biomedical applications. *Cells* **2021**, *10*, 1140. [CrossRef]
24. Arbabi Ghahroudi, M.; Desmyter, A.; Wyns, L.; Hamers, R.; Muyldermans, S. Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett.* **1997**, *414*, 521–526. [CrossRef]
25. Hamers-Casterman, C.; Atarhouch, T.; Muyldermans, S.; Robinson, G.; Hamers, C.; Songa, E.B.; Bendahman, N.; Hamers, R. Naturally occurring antibodies devoid of light chains. *Nature* **1993**, *363*, 446–448. [CrossRef]
26. Muyldermans, S.; Cambillau, C.; Wyns, L. Recognition of antigens by single-domain antibody fragments: The superfluous luxury of paired domains. *Trends Biochem. Sci.* **2001**, *26*, 230–235. [CrossRef]
27. De Genst, E.; Saerens, D.; Muyldermans, S.; Conrath, K. Antibody repertoire development in camelids. *Dev. Comp. Immunol.* **2006**, *30*, 187–198. [CrossRef]
28. Omidfar, K.; Rasaee, M.J.; Kashanian, S.; Paknejad, M.; Bathaie, Z. Studies of thermostability in camelus bactrianus (bactrian camel) single-domain antibody specific for the mutant epidermal-growth-factor receptor expressed by pichia. *Biotechnol. Appl. Biochem.* **2007**, *46*, 41–49.
29. Perruchini, C.; Pecorari, F.; Bourgeois, J.P.; Duyckaerts, C.; Rougeon, F.; Lafaye, P. Llama vhh antibody fragments against gfap: Better diffusion in fixed tissues than classical monoclonal antibodies. *Acta Neuropathol.* **2009**, *118*, 685–695. [CrossRef]
30. Tu, Z.; Huang, X.; Fu, J.; Hu, N.; Zheng, W.; Li, Y.; Zhang, Y. Landscape of variable domain of heavy-chain-only antibody repertoire from alpaca. *Immunology* **2020**, *161*, 53–65. [CrossRef]
31. Muyldermans, S. Applications of nanobodies. *Annu. Rev. Anim. Biosci.* **2021**, *9*, 401–421. [CrossRef] [PubMed]

32. Smolarek, D.; Bertrand, O.; Czerwinski, M. Variable fragments of heavy chain antibodies (vhhs): A new magic bullet molecule of medicine? *Postepy Hig. I Med. Dosw. (Online)* **2012**, *66*, 348–358. [CrossRef] [PubMed]
33. Bao, C.; Gao, Q.; Li, L.L.; Han, L.; Zhang, B.; Ding, Y.; Song, Z.; Zhang, R.; Zhang, J.; Wu, X.H. The application of nanobody in car-t therapy. *Biomolecules* **2021**, *11*, 238. [CrossRef] [PubMed]
34. Liu, M.; Li, L.; Jin, D.; Liu, Y. Nanobody-a versatile tool for cancer diagnosis and therapeutics. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnology* **2021**, *13*, e1697. [CrossRef]
35. Hosseindokht, M.; Bakherad, H.; Zare, H. Nanobodies: A tool to open new horizons in diagnosis and treatment of prostate cancer. *Cancer Cell Int.* **2021**, *21*, 580. [CrossRef]
36. Moradi, A.; Pourseif, M.M.; Jafari, B.; Parvizpour, S.; Omidi, Y. Nanobody-based therapeutics against colorectal cancer: Precision therapies based on the personal mutanome profile and tumor neoantigens. *Pharmacol. Res.* **2020**, *156*, 104790. [CrossRef]
37. Park, S.R.; Lee, J.H.; Kim, K.; Kim, T.M.; Lee, S.H.; Choo, Y.K.; Kim, K.S.; Ko, K. Expression and in vitro function of anti-breast cancer llama-based single domain antibody vhh expressed in tobacco plants. *Int. J. Mol. Sci.* **2020**, *21*, 1354. [CrossRef] [PubMed]
38. Verhaar, E.R.; Woodham, A.W.; Ploegh, H.L. Nanobodies in cancer. *Semin. Immunol.* **2021**, *52*, 101425. [CrossRef]
39. Zhai, T.; Wang, C.; Xu, Y.; Huang, W.; Yuan, Z.; Wang, T.; Dai, S.; Peng, S.; Pang, T.; Jiang, W.; et al. Generation of a safe and efficacious llama single-domain antibody fragment (vhh) targeting the membrane-proximal region of 4-1bb for engineering therapeutic bispecific antibodies for cancer. *J. Immunother. Cancer* **2021**, *9*, e002131. [CrossRef]
40. Iezzi, M.E.; Policastro, L.; Werbajh, S.; Podhajcer, O.; Canziani, G.A. Single-domain antibodies and the promise of modular targeting in cancer imaging and treatment. *Front. Immunol.* **2018**, *9*, 273. [CrossRef]
41. Wang, J.; Mukhtar, H.; Ma, L.; Pang, Q.; Wang, X. Vhh antibodies: Reagents for mycotoxin detection in food products. *Sensors* **2018**, *18*, 485. [CrossRef] [PubMed]
42. Bever, C.S.; Dong, J.X.; Vasylieva, N.; Barnych, B.; Cui, Y.; Xu, Z.L.; Hammock, B.D.; Gee, S.J. Vhh antibodies: Emerging reagents for the analysis of environmental chemicals. *Anal. Bioanal. Chem.* **2016**, *408*, 5985–6002. [CrossRef] [PubMed]
43. Gong, X.; Zhu, M.; Li, G.; Lu, X.; Wan, Y. Specific determination of influenza h7n2 virus based on biotinylated single-domain antibody from a phage-displayed library. *Anal. Biochem.* **2016**, *500*, 66–72. [CrossRef] [PubMed]
44. Dash, L.; Subramaniam, S.; Khulape, S.A.; Prusty, B.R.; Pargai, K.; Narnaware, S.D.; Patil, N.V.; Pattnaik, B. Development and utilization of vhh antibodies derived from camelus dromedarius against foot-and-mouth disease virus. *Anim. Biotechnol.* **2019**, *30*, 57–62. [CrossRef]
45. Li, H.; Dekker, A.; Sun, S.; Burman, A.; Kortekaas, J.; Harmsen, M.M. Novel capsid-specific single-domain antibodies with broad foot-and-mouth disease strain recognition reveal differences in antigenicity of virions, empty capsids, and virus-like particles. *Vaccines* **2021**, *9*, 620. [CrossRef] [PubMed]
46. Su, B.; Wang, Y.; Pei, H.; Sun, Z.; Cao, H.; Zhang, C.; Chen, Q.; Liu, X. Phage-mediated double-nanobody sandwich immunoassay for detecting alpha fetal protein in human serum. *Anal. Methods Adv. Methods Appl.* **2020**, *12*, 4742–4748. [CrossRef]
47. Ji, Y.; Chen, L.; Wang, Y.; Zhang, K.; Wu, H.; Liu, Y.; Wang, Y.; Wang, J. Development of a double nanobody-based sandwich immunoassay for the detecting staphylococcal enterotoxin c in dairy products. *Foods* **2021**, *10*, 2426. [CrossRef]
48. Steeland, S.; Vandenbroucke, R.E.; Libert, C. Nanobodies as therapeutics: Big opportunities for small antibodies. *Drug Discov. Today* **2016**, *21*, 1076–1113. [CrossRef]
49. Rossotti, M.A.; Bélanger, K.; Henry, K.A.; Tanha, J. Immunogenicity and humanization of single-domain antibodies. *FEBS J.* **2021**, *in press*. [CrossRef] [PubMed]
50. Scully, M.; Cataland, S.R.; Peyvandi, F.; Coppo, P.; Knöbl, P.; Kremer Hovinga, J.A.; Metjian, A.; de la Rubia, J.; Pavenski, K.; Callewaert, F.; et al. Caplacizumab treatment for acquired thrombotic thrombocytopenic purpura. *N. Engl. J. Med.* **2019**, *380*, 335–346. [CrossRef]
51. Jovčevska, I.; Muyldermans, S. The therapeutic potential of nanobodies. *BioDrugs Clin. Immunother. Biopharm. Gene Ther.* **2020**, *34*, 11–26. [CrossRef] [PubMed]
52. Senolt, L. Emerging therapies in rheumatoid arthritis: Focus on monoclonal antibodies. *F1000Research* **2019**, *8*, 1549. [CrossRef] [PubMed]
53. Huo, J.; Le Bas, A.; Ruza, R.R.; Duyvesteyn, H.M.E.; Mikolajek, H.; Malinauskas, T.; Tan, T.K.; Rijal, P.; Dumoux, M.; Ward, P.N.; et al. Neutralizing nanobodies bind sars-cov-2 spike rbd and block interaction with ace2. *Nat. Struct. Mol. Biol.* **2020**, *27*, 846–854. [CrossRef]
54. Wrapp, D.; De Vlieger, D.; Corbett, K.S.; Torres, G.M.; Wang, N.; Van Breedam, W.; Roose, K.; van Schie, L.; Hoffmann, M.; Pöhlmann, S.; et al. Structural basis for potent neutralization of betacoronaviruses by single-domain camelid antibodies. *Cell* **2020**, *181*, 1004–1015.e15. [CrossRef]
55. Chen, F.; Liu, Z.; Jiang, F. Prospects of neutralizing nanobodies against sars-cov-2. *Front. Immunol.* **2021**, *12*, 690742. [CrossRef] [PubMed]
56. Güttler, T.; Aksu, M.; Dickmanns, A.; Stegmann, K.M.; Gregor, K.; Rees, R.; Taxer, W.; Rymarenko, O.; Schünemann, J.; Dienemann, C.; et al. Neutralization of sars-cov-2 by highly potent, hyperthermostable, and mutation-tolerant nanobodies. *EMBO J.* **2021**, *40*, e107985. [CrossRef] [PubMed]
57. Hanke, L.; Vidakovics Perez, L.; Sheward, D.J.; Das, H.; Schulte, T.; Moliner-Morro, A.; Corcoran, M.; Achour, A.; Karlsson Hedestam, G.B.; Hällberg, B.M.; et al. An alpaca nanobody neutralizes sars-cov-2 by blocking receptor interaction. *Nat. Commun.* **2020**, *11*, 4420. [CrossRef]

58. Koenig, P.A.; Das, H.; Liu, H.; Kümmerer, B.M.; Gohr, F.N.; Jenster, L.M.; Schiffelers, L.D.J.; Tesfamariam, Y.M.; Uchima, M.; Wuerth, J.D.; et al. Structure-guided multivalent nanobodies block sars-cov-2 infection and suppress mutational escape. *Science* **2021**, *371*, eabe6230. [CrossRef]
59. Schoof, M.; Faust, B.; Saunders, R.A.; Sangwan, S.; Rezelj, V.; Hoppe, N.; Boone, M.; Billesbølle, C.B.; Puchades, C.; Azumaya, C.M.; et al. An ultrapotent synthetic nanobody neutralizes sars-cov-2 by stabilizing inactive spike. *Science* **2020**, *370*, 1473–1479. [CrossRef] [PubMed]
60. Wu, Y.; Li, C.; Xia, S.; Tian, X.; Kong, Y.; Wang, Z.; Gu, C.; Zhang, R.; Tu, C.; Xie, Y.; et al. Identification of human single-domain antibodies against sars-cov-2. *Cell Host Microbe* **2020**, *27*, 891–898.e5. [CrossRef]
61. Xiang, Y.; Nambulli, S.; Xiao, Z.; Liu, H.; Sang, Z.; Duprex, W.P.; Schneidman-Duhovny, D.; Zhang, C.; Shi, Y. Versatile and multivalent nanobodies efficiently neutralize sars-cov-2. *Science* **2020**, *370*, 1479–1484. [CrossRef]
62. Xu, J.; Xu, K.; Jung, S.; Conte, A.; Lieberman, J.; Muecksch, F.; Lorenzi, J.C.C.; Park, S.; Schmidt, F.; Wang, Z.; et al. Nanobodies from camelid mice and llamas neutralize sars-cov-2 variants. *Nature* **2021**, *595*, 278–282. [CrossRef]
63. Favorskaya, I.A.; Shcheblyakov, D.V.; Esmagambetov, I.B.; Dolzhikova, I.V.; Alekseeva, I.A.; Korobkova, A.I.; Voronina, D.V.; Ryabova, E.I.; Derkaev, A.A.; Kovyrshina, A.V.; et al. Single-domain antibodies efficiently neutralize sars-cov-2 variants of concern. *Front. Immunol.* **2022**, *13*, 822159. [CrossRef] [PubMed]
64. Weinstein, J.B.; Bates, T.A.; Leier, H.C.; McBride, S.K.; Barklis, E.; Tafesse, F.G. A potent alpaca-derived nanobody that neutralizes sars-cov-2 variants. *iScience* **2022**, *25*, 103960. [CrossRef] [PubMed]
65. Verkhivker, G. Allosteric determinants of the sars-cov-2 spike protein binding with nanobodies: Examining mechanisms of mutational escape and sensitivity of the omicron variant. *Int. J. Mol. Sci.* **2022**, *23*, 2172. [CrossRef]
66. Chabrol, E.; Stojko, J.; Nicolas, A.; Botzanowski, T.; Fould, B.; Antoine, M.; Cianférani, S.; Ferry, G.; Boutin, J.A. Vhh characterization.Recombinant vhhs: Production, characterization and affinity. *Anal. Biochem.* **2020**, *589*, 113491. [CrossRef]
67. De Buck, S.; Virdi, V.; De Meyer, T.; De Wilde, K.; Piron, R.; Nolf, J.; Van Lerberge, E.; De Paepe, A.; Depicker, A. Production of camel-like antibodies in plants. *Methods Mol. Biol.* **2012**, *911*, 305–324.
68. Meng, L.; Gao, X.; Liu, X.; Sun, M.; Yan, H.; Li, A.; Yang, Y.; Bai, Z. Enhancement of heterologous protein production in corynebacterium glutamicum via atmospheric and room temperature plasma mutagenesis and high-throughput screening. *J. Biotechnol.* **2021**, *339*, 22–31. [CrossRef] [PubMed]
69. Reader, R.H.; Workman, R.G.; Maddison, B.C.; Gough, K.C. Advances in the production and batch reformatting of phage antibody libraries. *Mol. Biotechnol.* **2019**, *61*, 801–815. [CrossRef] [PubMed]
70. Noël, F.; Malpertuy, A.; de Brevern, A.G. Global analysis of vhhs framework regions with a structural alphabet. *Biochimie* **2016**, *131*, 11–19. [CrossRef]
71. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The protein data bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [CrossRef]
72. Melarkode Vattekatte, A.; Shinada, N.K.; Narwani, T.J.; Noël, F.; Bertrand, O.; Meyniel, J.P.; Malpertuy, A.; Gelly, J.C.; Cadet, F.; de Brevern, A.G. Discrete analysis of camelid variable domains: Sequences, structures, and in-silico structure prediction. *PeerJ* **2020**, *8*, e8408. [CrossRef]
73. Smolarek, D.; Bertrand, O.; Czerwinski, M.; Colin, Y.; Etchebest, C.; de Brevern, A.G. Multiple interests in structural models of darc transmembrane protein. *Transfus. Clin. Biol. J. Soc. Fr. Transfus. Sang.* **2010**, *17*, 184–196. [CrossRef]
74. Melarkode Vattekatte, A.; Cadet, F.; Gelly, J.C.; de Brevern, A.G. Insights into comparative modeling of v(h)h domains. *Int. J. Mol. Sci.* **2021**, *22*, 9771. [CrossRef] [PubMed]
75. Adolf-Bryfogle, J.; Xu, Q.; North, B.; Lehmann, A.; Dunbrack, R.L. Pyigclassify: A database of antibody cdr structural classifications. *Nucleic Acids Res.* **2015**, *43*, D432–D438. [CrossRef] [PubMed]
76. Chen, X.; Kang, Y.; Luo, J.; Pang, K.; Xu, X.; Wu, J.; Li, X.; Jin, S. Next-generation sequencing reveals the progression of covid-19. *Front. Cell. Infect. Microbiol.* **2021**, *11*, 632490. [CrossRef]
77. Makałowski, W.; Shabardina, V. Bioinformatics of nanopore sequencing. *J. Hum. Genet.* **2020**, *65*, 61–67. [CrossRef]
78. Berman, H.M.; Battistuz, T.; Bhat, T.N.; Bluhm, W.F.; Bourne, P.E.; Burkhardt, K.; Feng, Z.; Gilliland, G.L.; Iype, L.; Jain, S.; et al. The protein data bank. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2002**, *58*, 899–907. [CrossRef] [PubMed]
79. Richardson, J.S.; Richardson, D.C.; Goodsell, D.S. Seeing the pdb. *J. Biol. Chem.* **2021**, *296*, 100742. [CrossRef] [PubMed]
80. Garnier, J. Protein structure prediction. *Biochimie* **1990**, *72*, 513–524. [CrossRef]
81. Baker, D.; Sali, A. Protein structure prediction and structural genomics. *Science* **2001**, *294*, 93–96. [CrossRef] [PubMed]
82. Gło, N. Statistical mechanics of protein folding, unfolding and fluctuation. *Adv. Biophys.* **1976**, 65–113.
83. Greer, J. Comparative modeling of homologous proteins. *Methods Enzymol.* **1991**, *202*, 239–252.
84. Sali, A.; Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815. [CrossRef] [PubMed]
85. Jones, D.T. Genthreader: An efficient and reliable protein fold recognition method for genomic sequences. *J. Mol. Biol.* **1999**, *287*, 797–815. [CrossRef]
86. Ghouzam, Y.; Postic, G.; de Brevern, A.G.; Gelly, J.C. Improving protein fold recognition with hybrid profiles combining sequence and structure evolution. *Bioinformatics* **2015**, *31*, 3782–3789. [CrossRef] [PubMed]
87. Ghouzam, Y.; Postic, G.; Guerin, P.E.; de Brevern, A.G.; Gelly, J.C. Orion: A web server for protein fold recognition and structure prediction using evolutionary hybrid profiles. *Sci. Rep.* **2016**, *6*, 28268. [CrossRef]

88. Bystroff, C.; Shao, Y. Fully automated ab initio protein structure prediction using i-sites, hmmstr and rosetta. *Bioinformatics* **2002**, *18*, 54–61. [CrossRef]
89. Rohl, C.A.; Strauss, C.E.; Misura, K.M.; Baker, D. Protein structure prediction using rosetta. *Methods Enzymol.* **2004**, *383*, 66–93. [PubMed]
90. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The i-tasser suite: Protein structure and function prediction. *Nat. Methods* **2015**, *12*, 7–8. [CrossRef]
91. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with alphafold. *Nature* **2021**, *596*, 583–589. [CrossRef]
92. Du, Z.; Su, H.; Wang, W.; Ye, L.; Wei, H.; Peng, Z.; Anishchenko, I.; Baker, D.; Yang, J. The trrosetta server for fast and accurate protein structure prediction. *Nat. Protoc.* **2021**, *16*, 5634–5651. [CrossRef]
93. Leman, J.K.; Weitzner, B.D.; Lewis, S.M.; Adolf-Bryfogle, J.; Alam, N.; Alford, R.F.; Aprahamian, M.; Baker, D.; Barlow, K.A.; Barth, P.; et al. Macromolecular modeling and design in rosetta: Recent methods and frameworks. *Nat. Methods* **2020**, *17*, 665–680. [CrossRef]
94. Nguyen, D.D.; Cang, Z.; Wei, G.W. A review of mathematical representations of biomolecular data. *Phys. Chem. Chem. Phys. PCCP* **2020**, *22*, 4343–4367. [CrossRef] [PubMed]
95. Meli, R.; Biggin, P.C. Spyrmsd: Symmetry-corrected rmsd calculations in python. *J. Cheminf.* **2020**, *12*, 49. [CrossRef]
96. Jauch, R.; Yeo, H.C.; Kolatkar, P.R.; Clarke, N.D. Assessment of casp7 structure predictions for template free targets. *Proteins* **2007**, *69*, 57–67. [CrossRef]
97. Zhang, Y.; Skolnick, J. Scoring function for automated assessment of protein structure template quality. *Proteins* **2004**, *57*, 702–710. [CrossRef]
98. Zhang, Y.; Skolnick, J. Tm-align: A protein structure alignment algorithm based on the tm-score. *Nucleic Acids Res.* **2005**, *33*, 2302–2309. [CrossRef] [PubMed]
99. Ghosh, S.; Gadiyaram, V.; Vishveshwara, S. Validation of protein structure models using network similarity score. *Proteins* **2017**, *85*, 1759–1776. [CrossRef] [PubMed]
100. Lefranc, M.P. Antibody informatics: Imgt, the international immunogenetics information system. *Microbiol. Spectr.* **2014**, *2*, 2.2.01. [CrossRef]
101. Lefranc, M.P.; Lefranc, G. Immunoglobulins or antibodies: Imgt(®) bridging genes, structures and functions. *Biomedicines* **2020**, *8*, 319. [CrossRef] [PubMed]
102. Pallarès, N.; Lefebvre, S.; Contet, V.; Matsuda, F.; Lefranc, M.P. The human immunoglobulin heavy variable genes. *Exp. Clin. Immunogenet.* **1999**, *16*, 36–60. [CrossRef] [PubMed]
103. Mariuzza, R.A.; Phillips, S.E.; Poljak, R.J. The structural basis of antigen-antibody recognition. *Annu. Rev. Biophys. Biophys. Chem.* **1987**, *16*, 139–159. [CrossRef] [PubMed]
104. Vargas-Madrazo, E.; Paz-García, E. An improved model of association for vh-vl immunoglobulin domains: Asymmetries between vh and vl in the packing of some interface residues. *J. Mol. Recognit. JMR* **2003**, *16*, 113–120. [CrossRef] [PubMed]
105. McCoy, A.J.; Grosse-Kunstleve, R.W.; Adams, P.D.; Winn, M.D.; Storoni, L.C.; Read, R.J. Phaser crystallographic software. *J. Appl. Cryst.* **2007**, *40*, 658–674. [CrossRef] [PubMed]
106. Webb, B.; Sali, A. Comparative protein structure modeling using modeller. *Curr. Protoc. Bioinform.* **2016**, *54*, 5.6.1–5.6.37. [CrossRef] [PubMed]
107. Shen, M.Y.; Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci. A Publ. Protein Soc.* **2006**, *15*, 2507–2524. [CrossRef]
108. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped blast and psi-blast: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [CrossRef] [PubMed]
109. Schäffer, A.A.; Wolf, Y.I.; Ponting, C.P.; Koonin, E.V.; Aravind, L.; Altschul, S.F. Impala: Matching a protein sequence against a collection of psi-blast-constructed position-specific score matrices. *Bioinformatics* **1999**, *15*, 1000–1011. [CrossRef]
110. Choong, Y.S.; Lee, Y.V.; Soong, J.X.; Law, C.T.; Lim, Y.Y. Computer-aided antibody design: An overview. *Adv. Exp. Med. Biol.* **2017**, *1053*, 221–243. [PubMed]
111. Smolarek, D.; Hattab, C.; Hassanzadeh-Ghassabeh, G.; Cochet, S.; Gutiérrez, C.; de Brevern, A.G.; Udomsangpetch, R.; Picot, J.; Grodecka, M.; Wasniowska, K.; et al. A recombinant dromedary antibody fragment (vhh or nanobody) directed against human duffy antigen receptor for chemokines. *Cell. Mol. Life Sci. CMLS* **2010**, *67*, 3371–3387. [CrossRef] [PubMed]
112. Smolarek, D.; Hattab, C.; Buczkowska, A.; Kaczmarek, R.; Jarząb, A.; Cochet, S.; de Brevern, A.G.; Lukasiewicz, J.; Jachymek, W.; Niedziela, T.; et al. Studies of a murine monoclonal antibody directed against darc: Reappraisal of its specificity. *PLoS ONE* **2015**, *10*, e0116472.
113. Steeland, S.; Puimège, L.; Vandenbroucke, R.E.; Van Hauwermeiren, F.; Haustraete, J.; Devoogdt, N.; Hulpiau, P.; Leroux-Roels, G.; Laukens, D.; Meuleman, P.; et al. Generation and characterization of small single domain antibodies inhibiting human tumor necrosis factor receptor 1. *J. Biol. Chem.* **2015**, *290*, 4022–4037. [CrossRef]
114. Lovell, S.C.; Davis, I.W.; Arendall III, W.B.; de Bakker, P.I.; Word, J.M.; Prisant, M.G.; Richardson, J.S.; Richardson, D.C. Structure validation by calpha geometry: Phi,psi and cbeta deviation. *Proteins* **2003**, *50*, 437–450. [CrossRef] [PubMed]
115. Comeau, S.R.; Gatchell, D.W.; Vajda, S.; Camacho, C.J. Cluspro: A fully automated algorithm for protein-protein docking. *Nucleic Acids Res.* **2004**, *32*, W96–W99. [CrossRef] [PubMed]

116. Shahangian, S.S.; Sajedi, R.H.; Hasannia, S.; Jalili, S.; Mohammadi, M.; Taghdir, M.; Shali, A.; Mansouri, K.; Sariri, R. A conformation-based phage-display panning to screen neutralizing anti-vegf vhhs with vegfr2 mimicry behavior. *Int. J. Biol. Macromol.* **2015**, *77*, 222–234. [CrossRef] [PubMed]
117. Calpe, S.; Wagner, K.; El Khattabi, M.; Rutten, L.; Zimberlin, C.; Dolk, E.; Verrips, C.T.; Medema, J.P.; Spits, H.; Krishnadath, K.K. Effective inhibition of bone morphogenetic protein function by highly specific llama-derived antibodies. *Mol. Cancer Ther.* **2015**, *14*, 2527–2540. [CrossRef] [PubMed]
118. Cheng, X.; Wang, J.; Kang, G.; Hu, M.; Yuan, B.; Zhang, Y.; Huang, H. Homology modeling-based in silico affinity maturation improves the affinity of a nanobody. *Int. J. Mol. Sci.* **2019**, *20*, 4187. [CrossRef] [PubMed]
119. Khodabakhsh, F.; Salimian, M.; Ziaee, P.; Kazemi-Lomedasht, F.; Behdani, M.; Ahangari Cohan, R. Designing and development of a tandem bivalent nanobody against vegf(165). *Avicenna J. Med. Biotechnol.* **2021**, *13*, 58–64. [PubMed]
120. Prado, N.D.; Pereira, S.S.; da Silva, M.P.; Morais, M.S.; Kayano, A.M.; Moreira-Dill, L.S.; Luiz, M.B.; Zanchi, F.B.; Fuly, A.L.; Huacca, M.E.; et al. Inhibition of the myotoxicity induced by bothrops jararacussu venom and isolated phospholipases a2 by specific camelid single-domain antibody fragments. *PLoS ONE* **2016**, *11*, e0151363. [CrossRef] [PubMed]
121. Demeestere, D.; Dejonckheere, E.; Steeland, S.; Hulpiau, P.; Haustraete, J.; Devoogdt, N.; Wichert, R.; Becker-Pauly, C.; Van Wonterghem, E.; Dewaele, S.; et al. Development and validation of a small single-domain antibody that effectively inhibits matrix metalloproteinase 8. *Mol. Ther. J. Am. Soc. Gene Ther.* **2016**, *24*, 890–902. [CrossRef] [PubMed]
122. Pang, Q.; Chen, Y.; Mukhtar, H.; Xiong, J.; Wang, X.; Xu, T.; Hammock, B.D.; Wang, J. Camelization of a murine single-domain antibody against aflatoxin b(1) and its antigen-binding analysis. *Mycotoxin Res.* **2022**, *38*, 51–60. [CrossRef] [PubMed]
123. Lin, J.; Lee, S.L.; Russell, A.M.; Huang, R.F.; Batt, M.A.; Chang, S.S.; Ferrante, A.; Verdino, P. A structure-based engineering approach to abrogate pre-existing antibody binding to biotherapeutics. *PLoS ONE* **2021**, *16*, e0254944. [CrossRef] [PubMed]
124. Yerabham, A.S.K.; Müller-Schiffmann, A.; Ziehm, T.; Stadler, A.; Köber, S.; Indurkhya, X.; Marreiros, R.; Trossbach, S.V.; Bradshaw, N.J.; Prikulis, I.; et al. Biophysical insights from a single chain camelid antibody directed against the disrupted-in-schizophrenia 1 protein. *PLoS ONE* **2018**, *13*, e0191162. [CrossRef] [PubMed]
125. Mandal, C.; Kingery, B.D.; Anchin, J.M.; Subramaniam, S.; Linthicum, D.S. Abgen: A knowledge-based automated approach for antibody structure modeling. *Nat. Biotechnol.* **1996**, *14*, 323–328. [CrossRef] [PubMed]
126. Schwieters, C.D.; Bermejo, G.A.; Clore, G.M. Xplor-nih for molecular structure determination from nmr and other data sources. *Protein Sci. A Publ. Protein Soc.* **2018**, *27*, 26–40. [CrossRef] [PubMed]
127. Schmid, N.; Eichenberger, A.P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A.E.; van Gunsteren, W.F. Definition and testing of the gromos force-field versions 54a7 and 54b7. *Eur. Biophys. J. EBJ* **2011**, *40*, 843–856. [CrossRef] [PubMed]
128. Whitelegg, N.R.; Rees, A.R. Wam: An improved algorithm for modelling antibodies on the web. *Protein Eng.* **2000**, *13*, 819–824. [CrossRef]
129. Martin, A.C.; Cheetham, J.C.; Rees, A.R. Modeling antibody hypervariable loops: A combined algorithm. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 9268–9272. [CrossRef]
130. Martin, A.C.; Cheetham, J.C.; Rees, A.R. Molecular modeling of antibody combining sites. *Methods Enzymol.* **1991**, *203*, 121–153.
131. Bruccoleri, R.E.; Karplus, M. Prediction of the folding of short polypeptide segments by uniform conformational sampling. *Biopolymers* **1987**, *26*, 137–168. [CrossRef] [PubMed]
132. Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M.C. Swiss-model: An automated protein homology-modeling server. *Nucleic Acids Res.* **2003**, *31*, 3381–3385. [CrossRef]
133. Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Gallo Cassarino, T.; Bertoni, M.; Bordoli, L.; et al. Swiss-model: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **2014**, *42*, W252–W258. [CrossRef] [PubMed]
134. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. Swiss-model: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303. [CrossRef] [PubMed]
135. Remmert, M.; Biegert, A.; Hauser, A.; Söding, J. Hhblits: Lightning-fast iterative protein sequence searching by hmm-hmm alignment. *Nat. Methods* **2011**, *9*, 173–175. [CrossRef]
136. Biasini, M.; Schmidt, T.; Bienert, S.; Mariani, V.; Studer, G.; Haas, J.; Johner, N.; Schenk, A.D.; Philippsen, A.; Schwede, T. Openstructure: An integrated software framework for computational structural biology. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 701–709. [CrossRef]
137. Benkert, P.; Biasini, M.; Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **2011**, *27*, 343–350. [CrossRef] [PubMed]
138. Li, Z.; Zhang, F.; Zhang, C.; Wang, C.; Lu, P.; Zhao, X.; Hao, L.; Ding, J. Immunoinformatics prediction of omp2b and bcsp31 for designing multi-epitope vaccine against brucella. *Mol. Immunol.* **2019**, *114*, 651–660. [CrossRef]
139. Rahman, N.; Ali, F.; Basharat, Z.; Shehroz, M.; Khan, M.K.; Jeandet, P.; Nepovimova, E.; Kuca, K.; Khan, H. Vaccine design from the ensemble of surface glycoprotein epitopes of sars-cov-2: An immunoinformatics approach. *Vaccines* **2020**, *8*, 423. [CrossRef] [PubMed]
140. Ranieri, D.I.; Corgliano, D.M.; Franco, E.J.; Hofstetter, H.; Hofstetter, O. Investigation of the stereoselectivity of an anti-amino acid antibody using molecular modeling and ligand docking. *Chirality* **2008**, *20*, 559–570. [CrossRef]

141. Shen, H.D.; Tam, M.F.; Huang, C.H.; Chou, H.; Tai, H.Y.; Chen, Y.S.; Sheu, S.Y.; Thomas, W.R. Homology modeling and monoclonal antibody binding of the der f 7 dust mite allergen. *Immunol. Cell Biol.* **2011**, *89*, 225–230. [CrossRef] [PubMed]
142. Studer, G.; Tauriello, G.; Bienert, S.; Biasini, M.; Johner, N.; Schwede, T. Promod3-a versatile homology modelling toolbox. *PLoS Comput. Biol.* **2021**, *17*, e1008667. [CrossRef] [PubMed]
143. Murakami, T.; Kumachi, S.; Matsunaga, Y.; Sato, M.; Wakabayashi-Nakao, K.; Masaki, H.; Yonehara, R.; Motohashi, M.; Nemoto, N.; Tsuchiya, M. Construction of a humanized artificial vhh library reproducing structural features of camelid vhhs for therapeutics. *Antibodies* **2022**, *11*, 10. [CrossRef]
144. Hanke, L.; Das, H.; Sheward, D.J.; Perez Vidakovics, L.; Urgard, E.; Moliner-Morro, A.; Kim, C.; Karl, V.; Pankow, A.; Smith, N.L.; et al. A bispecific monomeric nanobody induces spike trimer dimers and neutralizes sars-cov-2 in vivo. *Nat. Commun.* **2022**, *13*, 155. [CrossRef]
145. Nordeen, S.A.; Andersen, K.R.; Knockenhauer, K.E.; Ingram, J.R.; Ploegh, H.L.; Schwartz, T.U. A nanobody suite for yeast scaffold nucleoporins provides details of the nuclear pore complex structure. *Nat. Commun.* **2020**, *11*, 6179. [CrossRef] [PubMed]
146. Thanongsaksrikul, J.; Srimanote, P.; Maneewatch, S.; Choowongkomon, K.; Tapchaisri, P.; Makino, S.I.; Kurazono, H.; Chaicumpa, W. A v h h that neutralizes the zinc metalloproteinase activity of botulinum neurotoxin type a. *J. Biol. Chem.* **2010**, *285*, 9657–9666. [CrossRef] [PubMed]
147. Higashida, R.; Matsunaga, Y. Enhanced conformational sampling of nanobody cdr h3 loop by generalized replica-exchange with solute tempering. *Life* **2021**, *11*, 1428. [CrossRef] [PubMed]
148. Orlov, I.; Hemmer, C.; Ackerer, L.; Lorber, B.; Ghannam, A.; Poignavent, V.; Hleibieh, K.; Sauter, C.; Schmitt-Keichinger, C.; Belval, L.; et al. Structural basis of nanobody recognition of grapevine fanleaf virus and of virus resistance loss. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 10848–10855. [CrossRef] [PubMed]
149. Mahajan, S.P.; Meksiriporn, B.; Waraho-Zhmayev, D.; Weyant, K.B.; Kocer, I.; Butler, D.C.; Messer, A.; Escobedo, F.A.; DeLisa, M.P. Computational affinity maturation of camelid single-domain intrabodies against the nonamyloid component of alpha-synuclein. *Sci. Rep.* **2018**, *8*, 17611. [CrossRef] [PubMed]
150. Bujotzek, A.; Fuchs, A.; Qu, C.; Benz, J.; Klostermann, S.; Antes, I.; Georges, G. Mofvab: Modeling the fv region of antibodies. *mAbs* **2015**, *7*, 838–852. [CrossRef] [PubMed]
151. Bujotzek, A.; Dunbar, J.; Lipsmeier, F.; Schäfer, W.; Antes, I.; Deane, C.M.; Georges, G. Prediction of vh-vl domain orientation for antibody variable domain modeling. *Proteins* **2015**, *83*, 681–695. [CrossRef] [PubMed]
152. Helmer-Citterich, M.; Rovida, E.; Luzzago, A.; Tramontano, A. Modelling antibody-antigen interactions: Ferritin as a case study. *Mol. Immunol.* **1995**, *32*, 1001–1010. [CrossRef]
153. Morea, V.; Lesk, A.M.; Tramontano, A. Antibody modeling: Implications for engineering and design. *Methods* **2000**, *20*, 267–279. [CrossRef] [PubMed]
154. Morea, V.; Tramontano, A.; Rustici, M.; Chothia, C.; Lesk, A.M. Antibody structure, prediction and redesign. *Biophys. Chem.* **1997**, *68*, 9–16. [CrossRef]
155. Tramontano, A.; Janda, K.; Napper, A.D.; Benkovic, S.J.; Lerner, R.A. Catalytic antibodies. *Cold Spring Harb. Symp. Quant. Biol.* **1987**, *52*, 91–96. [CrossRef] [PubMed]
156. Morea, V.; Tramontano, A.; Rustici, M.; Chothia, C.; Lesk, A.M. Conformations of the third hypervariable region in the vh domain of immunoglobulins. *J. Mol. Biol.* **1998**, *275*, 269–294. [CrossRef] [PubMed]
157. Lepore, R.; Olimpieri, P.P.; Messih, M.A.; Tramontano, A. Pigspro: Prediction of immunoglobulin structures v2. *Nucleic Acids Res.* **2017**, *45*, W17–W23. [CrossRef] [PubMed]
158. Leem, J.; Deane, C.M. High-throughput antibody structure modeling and design using abodybuilder. *Methods Mol. Biol.* **2019**, *1851*, 367–380. [PubMed]
159. Schneider, C.; Raybould, M.I.J.; Deane, C.M. Sabdab in the age of biotherapeutics: Updates including sabdab-nano, the nanobody structure tracker. *Nucleic Acids Res.* **2022**, *50*, D1368–D1372. [CrossRef] [PubMed]
160. George, S. Conformational diversity of cdr region during affinity maturation determines the affinity and stability of sars-cov-1 vhh-72 nanobody. *bioRxiv* **2020**. [CrossRef]
161. Klausen, M.S.; Anderson, M.V.; Jespersen, M.C.; Nielsen, M.; Marcatili, P. Lyra, a webserver for lymphocyte receptor structural modeling. *Nucleic Acids Res.* **2015**, *43*, W349–W355. [CrossRef]
162. Abhinandan, K.R.; Martin, A.C. Analysis and improvements to kabat and structurally correct numbering of antibody variable domains. *Mol. Immunol.* **2008**, *45*, 3832–3839. [CrossRef] [PubMed]
163. Marcatili, P.; Olimpieri, P.P.; Chailyan, A.; Tramontano, A. Antibody modeling using the prediction of immunoglobulin structure (pigs) web server [corrected]. *Nat. Protoc.* **2014**, *9*, 2771–2783. [CrossRef]
164. North, B.; Lehmann, A.; Dunbrack, R.L. A new clustering of antibody cdr loop conformations. *J. Mol. Biol.* **2011**, *406*, 228–256. [CrossRef] [PubMed]
165. Kelley, L.A.; Sternberg, M.J. Protein structure prediction on the web: A case study using the phyre server. *Nat. Protoc.* **2009**, *4*, 363–371. [CrossRef] [PubMed]
166. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J. The phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [CrossRef] [PubMed]
167. Söding, J. Protein homology detection by hmm-hmm comparison. *Bioinformatics* **2005**, *21*, 951–960. [CrossRef]

168. Jefferys, B.R.; Kelley, L.A.; Sternberg, M.J. Protein folding requires crowd control in a simulated cell. *J. Mol. Biol.* **2010**, *397*, 1329–1338. [CrossRef]
169. Rotkiewicz, P.; Skolnick, J. Fast procedure for reconstruction of full-atom protein models from reduced representations. *J. Comput. Chem.* **2008**, *29*, 1460–1465. [CrossRef] [PubMed]
170. Xie, W.; Sahinidis, N.V. Residue-rotamer-reduction algorithm for the protein side-chain conformation problem. *Bioinformatics* **2006**, *22*, 188–194. [CrossRef] [PubMed]
171. Hoseinpoor, R.; Mousavi Gargari, S.L.; Rasooli, I.; Rajabibazl, M.; Shahi, B. Functional mutations in and characterization of vhh against helicobacter pylori urease. *Appl. Biochem. Biotechnol.* **2014**, *172*, 3079–3091. [CrossRef]
172. Payandeh, Z.; Rasooli, I.; Mousavi Gargari, S.L.; Rajabi Bazl, M.; Ebrahimizadeh, W. Immunoreaction of a recombinant nanobody from camelid single domain antibody fragment with acinetobacter baumannii. *Trans. R. Soc. Trop. Med. Hyg.* **2014**, *108*, 92–98. [CrossRef]
173. Chen, C.C.; Hwang, J.K.; Yang, J.M. (ps)2-v2: Template-based protein structure prediction server. *BMC Bioinform.* **2009**, *10*, 366. [CrossRef]
174. Wu, S.; Zhang, Y. Lomets: A local meta-threading-server for protein structure prediction. *Nucleic Acids Res.* **2007**, *35*, 3375–3382. [CrossRef]
175. Sippl, M.J. Recognition of errors in three-dimensional structures of proteins. *Proteins* **1993**, *17*, 355–362. [CrossRef] [PubMed]
176. Wiederstein, M.; Sippl, M.J. Prosa-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **2007**, *35*, W407–W410. [CrossRef] [PubMed]
177. Xu, D.; Zhang, Y. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys. J.* **2011**, *101*, 2525–2534. [CrossRef] [PubMed]
178. Pierce, B.G.; Hourai, Y.; Weng, Z. Accelerating protein docking in zdock using an advanced 3d convolution library. *PLoS ONE* **2011**, *6*, e24657. [CrossRef]
179. Pierce, B.G.; Wiehe, K.; Hwang, H.; Kim, B.H.; Vreven, T.; Weng, Z. Zdock server: Interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics* **2014**, *30*, 1771–1773. [CrossRef] [PubMed]
180. Sefid, F.; Rasooli, I.; Payandeh, Z. Homology modeling of a camelid antibody fragment against a conserved region of acinetobacter baumannii biofilm associated protein (bap). *J. Theor. Biol.* **2016**, *397*, 43–51. [CrossRef]
181. Skottrup, P.D. Structural insights into a high affinity nanobody: Antigen complex by homology modelling. *J. Mol. Graph. Model.* **2017**, *76*, 305–312. [CrossRef] [PubMed]
182. Källberg, M.; Margaryan, G.; Wang, S.; Ma, J.; Xu, J. Raptorx server: A resource for template-based protein structure modeling. *Methods Mol. Biol.* **2014**, *1137*, 17–27.
183. Källberg, M.; Wang, H.; Wang, S.; Peng, J.; Wang, Z.; Lu, H.; Xu, J. Template-based protein structure modeling using the raptorx web server. *Nat. Protoc.* **2012**, *7*, 1511–1522. [CrossRef]
184. Xu, J.; Li, M. Assessment of raptor's linear programming approach in cafasp3. *Proteins* **2003**, *53*, 579–584. [CrossRef]
185. Xu, J.; Li, M.; Kim, D.; Xu, Y. Raptor: Optimal protein threading by linear programming. *J. Bioinform. Comput. Biol.* **2003**, *1*, 95–117. [CrossRef]
186. Wang, S.; Li, W.; Liu, S.; Xu, J. Raptorx-property: A web server for protein structure property prediction. *Nucleic Acids Res.* **2016**, *44*, W430–W435. [CrossRef] [PubMed]
187. Wang, S.; Sun, S.; Li, Z.; Zhang, R.; Xu, J. Accurate de novo prediction of protein contact map by ultra-deep learning model. *PLoS Comput. Biol.* **2017**, *13*, e1005324. [CrossRef] [PubMed]
188. Peng, J.; Xu, J. Raptorx: Exploiting structure information for protein alignment by statistical inference. *Proteins* **2011**, *79*, 161–171. [CrossRef]
189. Jittavisutthikul, S.; Thanongsaksrikul, J.; Thueng-In, K.; Chulanetra, M.; Srimanote, P.; Seesuay, W.; Malik, A.A.; Chaicumpa, W. Humanized-vhh transbodies that inhibit hcv protease and replication. *Viruses* **2015**, *7*, 2030–2056. [CrossRef]
190. Thueng-in, K.; Thanongsaksrikul, J.; Srimanote, P.; Bangphoomi, K.; Poungpair, O.; Maneewatch, S.; Choowongkomon, K.; Chaicumpa, W. Cell penetrable humanized-vh/v(h)h that inhibit rna dependent rna polymerase (ns5b) of hcv. *PLoS ONE* **2012**, *7*, e49254. [CrossRef]
191. Chavanayarn, C.; Thanongsaksrikul, J.; Thueng-In, K.; Bangphoomi, K.; Sookrung, N.; Chaicumpa, W. Humanized-single domain antibodies (vh/vhh) that bound specifically to naja kaouthia phospholipase a2 and neutralized the enzymatic activity. *Toxins* **2012**, *4*, 554–567. [CrossRef] [PubMed]
192. Malik, A.A.; Imtong, C.; Sookrung, N.; Katzenmeier, G.; Chaicumpa, W.; Angsuthanasombat, C. Structural characterization of humanized nanobodies with neutralizing activity against the bordetella pertussis cyaa-hemolysin: Implications for a potential epitope of toxin-protective antigen. *Toxins* **2016**, *8*, 99. [CrossRef] [PubMed]
193. Ko, J.; Lee, D.; Park, H.; Coutsias, E.A.; Lee, J.; Seok, C. The falc-loop web server for protein loop modeling. *Nucleic Acids Res.* **2011**, *39*, W210–W214. [CrossRef]
194. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J. Gromacs: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718. [CrossRef] [PubMed]
195. Simons, K.T.; Bonneau, R.; Ruczinski, I.; Baker, D. Ab initio protein structure prediction of casp iii targets using rosetta. *Proteins* **1999**, *3*, 171–176. [CrossRef]

196. Bystroff, C.; Thorsson, V.; Baker, D. Hmmstr: A hidden markov model for local sequence-structure correlations in proteins. *J. Mol. Biol.* **2000**, *301*, 173–190. [CrossRef]
197. Bradley, P.; Malmström, L.; Qian, B.; Schonbrun, J.; Chivian, D.; Kim, D.E.; Meiler, J.; Misura, K.M.; Baker, D. Free modeling with rosetta in casp6. *Proteins* **2005**, *61*, 128–134. [CrossRef] [PubMed]
198. Raman, S.; Vernon, R.; Thompson, J.; Tyka, M.; Sadreyev, R.; Pei, J.; Kim, D.; Kellogg, E.; DiMaio, F.; Lange, O.; et al. Structure prediction for casp8 with all-atom refinement using rosetta. *Proteins* **2009**, *77*, 89–99. [CrossRef] [PubMed]
199. Song, Y.; DiMaio, F.; Wang, R.Y.; Kim, D.; Miles, C.; Brunette, T.; Thompson, J.; Baker, D. High-resolution comparative modeling with rosettacm. *Structure* **2013**, *21*, 1735–1742. [CrossRef]
200. Kim, D.E.; Chivian, D.; Baker, D. Protein structure prediction and analysis using the robetta server. *Nucleic Acids Res.* **2004**, *32*, W526–W531. [CrossRef] [PubMed]
201. Park, H.; Kim, D.E.; Ovchinnikov, S.; Baker, D.; DiMaio, F. Automatic structure prediction of oligomeric assemblies using robetta in casp12. *Proteins* **2018**, *86*, 283–291. [CrossRef] [PubMed]
202. Schmitz, S.; Ertelt, M.; Merkl, R.; Meiler, J. Rosetta design with co-evolutionary information retains protein function. *PLoS Comput. Biol.* **2021**, *17*, e1008568. [CrossRef]
203. Chaudhury, S.; Berrondo, M.; Weitzner, B.D.; Muthu, P.; Bergman, H.; Gray, J.J. Benchmarking and analysis of protein docking performance in rosetta v3.2. *PLoS ONE* **2011**, *6*, e22477. [CrossRef]
204. Tivon, B.; Gabizon, R.; Somsen, B.A.; Cossar, P.J.; Ottmann, C.; London, N. Covalent flexible peptide docking in rosetta. *Chem. Sci.* **2021**, *12*, 10836–10847. [CrossRef]
205. Govaert, J.; Pellis, M.; Deschacht, N.; Vincke, C.; Conrath, K.; Muyldermans, S.; Saerens, D. Dual beneficial effect of interloop disulfide bond for single domain antibody fragments. *J. Biol. Chem.* **2012**, *287*, 1970–1979. [CrossRef] [PubMed]
206. Sivasubramanian, A.; Sircar, A.; Chaudhury, S.; Gray, J.J. Toward high-resolution homology modeling of antibody fv regions and application to antibody-antigen docking. *Proteins* **2009**, *74*, 497–514. [CrossRef]
207. Sircar, A.; Kim, E.T.; Gray, J.J. Rosettaantibody: Antibody variable region homology modeling server. *Nucleic Acids Res.* **2009**, *37*, W474–W479. [CrossRef] [PubMed]
208. Schoeder, C.T.; Schmitz, S.; Adolf-Bryfogle, J.; Sevy, A.M.; Finn, J.A.; Sauer, M.F.; Bozhanova, N.G.; Mueller, B.K.; Sangha, A.K.; Bonet, J.; et al. Modeling immunity with rosetta: Methods for antibody and antigen design. *Biochemistry* **2021**, *60*, 825–846. [CrossRef]
209. Jeliazkov, J.R.; Frick, R.; Zhou, J.; Gray, J.J. Robustification of rosettaantibody and rosetta snugdock. *PLoS ONE* **2021**, *16*, e0234282. [CrossRef] [PubMed]
210. Sircar, A. Methods for the homology modeling of antibody variable regions. *Methods Mol. Biol. (Clifton, N.J.)* **2012**, *857*, 301–311.
211. Sircar, A.; Sanni, K.A.; Shi, J.; Gray, J.J. Analysis and modeling of the variable region of camelid single-domain antibodies. *J. Immunol.* **2011**, *186*, 6357–6367. [CrossRef]
212. Zimmermann, I.; Egloff, P.; Hutter, C.A.; Arnold, F.M.; Stohler, P.; Bocquet, N.; Hug, M.N.; Huber, S.; Siegrist, M.; Hetemann, L.; et al. Synthetic single domain antibodies for the conformational trapping of membrane proteins. *eLife* **2018**, *7*, e34317. [CrossRef]
213. Norn, C.H.; Lapidoth, G.; Fleishman, S.J. High-accuracy modeling of antibody structures by a search for minimum-energy recombination of backbone fragments. *Proteins* **2017**, *85*, 30–38. [CrossRef] [PubMed]
214. Lapidoth, G.; Parker, J.; Prilusky, J.; Fleishman, S.J. Abpredict 2: A server for accurate and unstrained structure prediction of antibody variable domains. *Bioinformatics* **2019**, *35*, 1591–1593. [CrossRef] [PubMed]
215. Kodali, P.; Schoeder, C.T.; Schmitz, S.; Crowe, J.E.; Meiler, J. Rosettacm for antibodies with very long hcdr3s and low template availability. *Proteins* **2021**, *89*, 1458–1472. [CrossRef]
216. Kemmish, H.; Fasnacht, M.; Yan, L. Fully automated antibody structure prediction using biovia tools: Validation study. *PLoS ONE* **2017**, *12*, e0177923. [CrossRef] [PubMed]
217. Maier, J.K.; Labute, P. Assessment of fully automated antibody homology modeling protocols in molecular operating environment. *Proteins* **2014**, *82*, 1599–1610. [CrossRef] [PubMed]
218. Yanakieva, D.; Pekar, L.; Evers, A.; Fleischer, M.; Keller, S.; Mueller-Pompalla, D.; Toleikis, L.; Kolmar, H.; Zielonka, S.; Krah, S. Beyond bispecificity: Controlled fab arm exchange for the generation of antibodies with multiple specificities. *mAbs* **2022**, *14*, 2018960. [CrossRef]
219. Berrondo, M.; Kaufmann, S.; Berrondo, M. Automated aufbau of antibody structures from given sequences using macromoltek's smrtmolantibody. *Proteins* **2014**, *82*, 1636–1645. [CrossRef]
220. Zhang, Y. Interplay of i-tasser and quark for template-based and ab initio protein structure prediction in casp10. *Proteins* **2014**, *82*, 175–187. [CrossRef] [PubMed]
221. Roy, A.; Kucukural, A.; Zhang, Y. I-tasser: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* **2010**, *5*, 725–738. [CrossRef] [PubMed]
222. Zheng, W.; Zhang, C.; Li, Y.; Pearce, R.; Bell, E.W.; Zhang, Y. Folding non-homologous proteins by coupling deep-learning contact maps with i-tasser assembly simulations. *Cell Rep. Methods* **2021**, *1*, 100014. [CrossRef] [PubMed]
223. Frenken, L.G.; van der Linden, R.H.; Hermans, P.W.; Bos, J.W.; Ruuls, R.C.; de Geus, B.; Verrips, C.T. Isolation of antigen specific llama vhh antibody fragments and their high level secretion by saccharomyces cerevisiae. *J. Biotechnol.* **2000**, *78*, 11–21. [CrossRef]
224. Fridy, P.C.; Li, Y.; Keegan, S.; Thompson, M.K.; Nudelman, I.; Scheid, J.F.; Oeffinger, M.; Nussenzweig, M.C.; Fenyö, D.; Chait, B.T.; et al. A robust pipeline for rapid production of versatile nanobody repertoires. *Nat. Methods* **2014**, *11*, 1253–1260. [CrossRef]

225. Fridy, P.C.; Thompson, M.K.; Ketaren, N.E.; Rout, M.P. Engineered high-affinity nanobodies recognizing staphylococcal protein a and suitable for native isolation of protein complexes. *Anal. Biochem.* **2015**, *477*, 92–94. [CrossRef] [PubMed]
226. Xu, D.; Zhang, Y. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins* **2012**, *80*, 1715–1735. [CrossRef]
227. Zhang, Y.; Skolnick, J. Spicker: A clustering approach to identify near-native protein folds. *J. Comput. Chem.* **2004**, *25*, 865–871. [CrossRef]
228. Mortuza, S.M.; Zheng, W.; Zhang, C.; Li, Y.; Pearce, R.; Zhang, Y. Improving fragment-based ab initio protein structure assembly using low-accuracy contact-map predictions. *Nat. Commun.* **2021**, *12*, 5011. [CrossRef] [PubMed]
229. Zheng, W.; Li, Y.; Zhang, C.; Pearce, R.; Mortuza, S.M.; Zhang, Y. Deep-learning contact-map guided protein structure prediction in casp13. *Proteins* **2019**, *87*, 1149–1164. [CrossRef]
230. Zhang, C.; Zheng, W.; Mortuza, S.M.; Li, Y.; Zhang, Y. Deepmsa: Constructing deep multiple sequence alignment to improve contact prediction and fold-recognition for distant-homology proteins. *Bioinformatics* **2020**, *36*, 2105–2112. [CrossRef]
231. Kinch, L.N.; Li, W.; Monastyrskyy, B.; Kryshtafovych, A.; Grishin, N.V. Evaluation of free modeling targets in casp11 and roll. *Proteins* **2016**, *84*, 51–66. [CrossRef] [PubMed]
232. Tai, C.H.; Bai, H.; Taylor, T.J.; Lee, B. Assessment of template-free modeling in casp10 and roll. *Proteins* **2014**, *82*, 57–83. [CrossRef] [PubMed]
233. Senior, A.W.; Evans, R.; Jumper, J.; Kirkpatrick, J.; Sifre, L.; Green, T.; Qin, C.; Žídek, A.; Nelson, A.W.R.; Bridgland, A.; et al. Improved protein structure prediction using potentials from deep learning. *Nature* **2020**, *577*, 706–710. [CrossRef]
234. AlQuraishi, M. Alphafold at casp13. *Bioinformatics* **2019**, *35*, 4862–4865. [CrossRef]
235. Senior, A.W.; Evans, R.; Jumper, J.; Kirkpatrick, J.; Sifre, L.; Green, T.; Qin, C.; Žídek, A.; Nelson, A.W.R.; Bridgland, A.; et al. Protein structure prediction using multiple deep neural networks in the 13th critical assessment of protein structure prediction (casp13). *Proteins* **2019**, *87*, 1141–1148. [CrossRef]
236. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Applying and improving alphafold at casp14. *Proteins* **2021**, *89*, 1711–1721. [CrossRef] [PubMed]
237. Fersht, A.R. Alphafold-a personal perspective on the impact of machine learning. *J. Mol. Biol.* **2021**, *433*, 167088. [CrossRef]
238. Porta-Pardo, E.; Ruiz-Serra, V.; Valentini, S.; Valencia, A. The structural coverage of the human proteome before and after alphafold. *PLoS Comput. Biol.* **2022**, *18*, e1009818. [CrossRef]
239. Tong, A.B.; Burch, J.D.; McKay, D.; Bustamante, C.; Crackower, M.A.; Wu, H. Could alphafold revolutionize chemical therapeutics? *Nat. Struct. Mol. Biol.* **2021**, *28*, 771–772. [CrossRef] [PubMed]
240. Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. Colabfold-making protein folding accessible to all. *bioRxiv* **2022**. [CrossRef]
241. Callaway, E. It will change everything: Deepmind's ai makes gigantic leap in solving protein structures. *Nature* **2020**, *588*, 203–204. [CrossRef] [PubMed]
242. Thornton, J.M.; Laskowski, R.A.; Borkakoti, N. Alphafold heralds a data-driven revolution in biology and medicine. *Nat. Med.* **2021**, *27*, 1666–1669. [CrossRef]
243. Cao, C.; Liu, F.; Tan, H.; Song, D.; Shu, W.; Li, W.; Zhou, Y.; Bo, X.; Xie, Z. Deep learning and its applications in biomedicine. *Genom. Proteom. Bioinform.* **2018**, *16*, 17–32. [CrossRef]
244. Rifaioglu, A.S.; Atas, H.; Martin, M.J.; Cetin-Atalay, R.; Atalay, V.; Doğan, T. Recent applications of deep learning and machine intelligence on in silico drug discovery: Methods, tools and databases. *Brief. Bioinform.* **2019**, *20*, 1878–1912. [CrossRef]
245. Shi, Q.; Chen, W.; Huang, S.; Wang, Y.; Xue, Z. Deep learning for mining protein data. *Brief. Bioinform.* **2021**, *22*, 194–218. [CrossRef]
246. Torrisi, M.; Pollastri, G.; Le, Q. Deep learning methods in protein structure prediction. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 1301–1310. [CrossRef] [PubMed]
247. Cretin, G.; Galochkina, T.; de Brevern, A.G.; Gelly, J.C. Pythia: Deep learning approach for local protein conformation prediction. *Int. J. Mol. Sci.* **2021**, *22*, 8831. [CrossRef] [PubMed]
248. Vander Meersche, Y.; Cretin, G.; de Brevern, A.G.; Gelly, J.C.; Galochkina, T. Medusa: Prediction of protein flexibility from sequence. *J. Mol. Biol.* **2021**, *433*, 166882. [CrossRef] [PubMed]
249. Alford, R.F.; Leaver-Fay, A.; Jeliazkov, J.R.; O'Meara, M.J.; DiMaio, F.P.; Park, H.; Shapovalov, M.V.; Renfrew, P.D.; Mulligan, V.K.; Kappel, K.; et al. The rosetta all-atom energy function for macromolecular modeling and design. *J. Chem. Theory Comput.* **2017**, *13*, 3031–3048. [CrossRef] [PubMed]
250. Yang, J.; Anishchenko, I.; Park, H.; Peng, Z.; Ovchinnikov, S.; Baker, D. Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 1496–1503. [CrossRef]
251. Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G.R.; Wang, J.; Cong, Q.; Kinch, L.N.; Schaeffer, R.D.; et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **2021**, *373*, 871–876. [CrossRef] [PubMed]
252. Ruffolo, J.; Sulam, J.; Gray, J.J. Antibody structure prediction using interpretable deep learning. *Patterns* **2022**, *3*, 100406. [CrossRef] [PubMed]
253. Cohen, T.; Halfon, M.; Schneidman-Duhovny, D. Nanonet: Rapid end-to-end nanobody modeling by deep learning at sub angstrom resolution. *bioRxiv* **2021**. [CrossRef]

254. Sun, D.; Sang, Z.; Kim, Y.J.; Xiang, Y.; Cohen, T.; Belford, A.K.; Huet, A.; Conway, J.F.; Sun, J.; Taylor, D.J.; et al. Potent neutralizing nanobodies resist convergent circulating variants of sars-cov-2 by targeting diverse and conserved epitopes. *Nat. Commun.* **2021**, *12*, 4676. [CrossRef]
255. de Brevern, A.G.; Etchebest, C.; Hazout, S. Bayesian probabilistic approach for predicting backbone structures in terms of protein blocks. *Proteins* **2000**, *41*, 271–287. [CrossRef]
256. Barnoud, J.; Santuz, H.; Craveur, P.; Joseph, A.P.; Jallu, V.; de Brevern, A.G.; Poulain, P. Pbxplore: A tool to analyze local protein structure and deformability with protein blocks. *PeerJ* **2017**, *5*, e4013. [CrossRef]
257. Pompidor, G.; Zimmermann, S.; Loew, C.; Schneider, T. Engineered Nanobodies with a Lanthanide Binding Motif for Crystallographic Phasing. 2021. Available online: https://www.rcsb.org/structure/6XYF (accessed on 20 February 2022).
258. Pieper, U.; Eswar, N.; Stuart, A.C.; Ilyin, V.A.; Sali, A. Modbase, a database of annotated comparative protein structure models. *Nucleic Acids Res.* **2002**, *30*, 255–259. [CrossRef] [PubMed]
259. Pieper, U.; Webb, B.M.; Dong, G.Q.; Schneidman-Duhovny, D.; Fan, H.; Kim, S.J.; Khuri, N.; Spill, Y.G.; Weinkam, P.; Hammel, M.; et al. Modbase, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res.* **2014**, *42*, D336–D346. [CrossRef] [PubMed]
260. Duhoo, Y.; Roche, J.; Trinh, T.T.N.; Desmyter, A.; Gaubert, A.; Kellenberger, C.; Cambillau, C.; Roussel, A.; Leone, P. Camelid nanobodies used as crystallization chaperones for different constructs of porm, a component of the type ix secretion system from porphyromonas gingivalis. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2017**, *73*, 286–293. [CrossRef] [PubMed]
261. Chen, L.; McLellan, J.; Kwon, Y.; Schmidt, S.; Wu, X.; Zhou, T.; Yang, Y.; Zhang, B.; Forsman, A.; Weiss, R.; et al. Single-Headed Immunoglobulins Efficiently Penetrate Cd4-Binding Site and Effectively Neutralize Hiv-1. 2012. Available online: https://www.rcsb.org/structure/3r0m (accessed on 20 February 2022).
262. Hinz, A.; Lutje Hulsik, D.; Forsman, A.; Koh, W.W.; Belrhali, H.; Gorlani, A.; de Haard, H.; Weiss, R.A.; Verrips, T.; Weissenhorn, W. Crystal structure of the neutralizing llama v(hh) d7 and its mode of hiv-1 gp120 interaction. *PLoS ONE* **2010**, *5*, e10482. [CrossRef]
263. Studer, G.; Rempfer, C.; Waterhouse, A.M.; Gumienny, R.; Haas, J.; Schwede, T. Qmeandisco-distance constraints applied on model quality estimation. *Bioinformatics* **2020**, *36*, 1765–1771. [CrossRef]
264. Mitchell, L.S.; Colwell, L.J. Comparative analysis of nanobody sequence and structure data. *Proteins* **2018**, *86*, 697–706. [CrossRef] [PubMed]
265. Ahmed, A.M.; Brooks, C.L. X-ray crystal structure analysis of vhh-protein antigen complexes. *Methods Mol. Biol.* **2022**, *2446*, 513–530. [PubMed]
266. Zuo, J.; Li, J.; Zhang, R.; Xu, L.; Chen, H.; Jia, X.; Su, Z.; Zhao, L.; Huang, X.; Xie, W. Institute collection and analysis of nanobodies (ican): A comprehensive database and analysis platform for nanobodies. *BMC Genom.* **2017**, *18*, 797. [CrossRef]
267. Sang, Z.; Xiang, Y.; Bahar, I.; Shi, Y. Llamanade: An open-source computational pipeline for robust nanobody humanization. *Structure* **2022**, *30*, 418–429.e3. [CrossRef] [PubMed]
268. Deszyński, P.; Młokosiewicz, J.; Volanakis, A.; Jaszczyszyn, I.; Castellana, N.; Bonissone, S.; Ganesan, R.; Krawczyk, K. Indi-integrated nanobody database for immunoinformatics. *Nucleic Acids Res.* **2022**, *50*, D1273–D1281. [CrossRef] [PubMed]
269. Tam, C.; Kumar, A.; Zhang, K.Y.J. Nbx: Machine learning-guided re-ranking of nanobody-antigen binding poses. *Pharmaceuticals* **2021**, *14*, 968. [CrossRef] [PubMed]
270. Tahir, S.; Bourquard, T.; Musnier, A.; Jullian, Y.; Corde, Y.; Omahdi, Z.; Mathias, L.; Reiter, E.; Crépieux, P.; Bruneau, G.; et al. Accurate determination of epitope for antibodies with unknown 3d structures. *mAbs* **2021**, *13*, 1961349. [CrossRef]
271. Bekker, G.J.; Ma, B.; Kamiya, N. Thermal stability of single-domain antibodies estimated by molecular dynamics simulations. *Protein Sci. A Publ. Protein Soc.* **2019**, *28*, 429–438. [CrossRef]
272. Che, T.; Majumdar, S.; Zaidi, S.A.; Ondachi, P.; McCorvy, J.D.; Wang, S.; Mosier, P.D.; Uprety, R.; Vardy, E.; Krumm, B.E.; et al. Structure of the nanobody-stabilized active state of the kappa opioid receptor. *Cell* **2018**, *172*, 55–67.e15. [CrossRef]
273. Mohseni, A.; Molakarimi, M.; Taghdir, M.; Sajedi, R.H.; Hasannia, S. Exploring single-domain antibody thermostability by molecular dynamics simulation. *J. Biomol. Struct. Dyn.* **2019**, *37*, 3686–3696. [CrossRef]
274. Soler, M.A.; Fortuna, S.; de Marco, A.; Laio, A. Binding affinity prediction of nanobody-protein complexes by scoring of molecular dynamics trajectories. *Phys. Chem. Chem. Phys. PCCP* **2018**, *20*, 3438–3444. [CrossRef]
275. Zabetakis, D.; Shriver-Lake, L.C.; Olson, M.A.; Goldman, E.R.; Anderson, G.P. Experimental evaluation of single-domain antibodies predicted by molecular dynamics simulations to have elevated thermal stability. *Protein Sci. A Publ. Protein Soc.* **2019**, *28*, 1909–1912. [CrossRef] [PubMed]
276. Narwani, T.J.; Craveur, P.; Shinada, N.K.; Floch, A.; Santuz, H.; Vattekatte, A.M.; Srinivasan, N.; Rebehmed, J.; Gelly, J.C.; Etchebest, C.; et al. Discrete analyses of protein dynamics. *J. Biomol. Struct. Dyn.* **2020**, *38*, 2988–3002. [CrossRef]
277. McCammon, J.A.; Karplus, M. Internal motions of antibody molecules. *Nature* **1977**, *268*, 765–766. [CrossRef] [PubMed]
278. Novotný, J.; Bruccoleri, R.; Newell, J.; Murphy, D.; Haber, E.; Karplus, M. Molecular anatomy of the antibody binding site. *J. Biol. Chem.* **1983**, *258*, 14433–14437. [CrossRef]