



## OPEN ACCESS

## EDITED BY

Kathryn H. Ching,  
NextVivo, Inc., United States

## REVIEWED BY

Carol F. Webb,  
University of Oklahoma Health  
Sciences Center, United States  
Helen Dooley,  
University of Maryland, Baltimore,  
United States

## \*CORRESPONDENCE

Klaus R. Liedl  
Klaus.Liedl@uibk.ac.at

## SPECIALTY SECTION

This article was submitted to  
B Cell Biology,  
a section of the journal  
Frontiers in Immunology

RECEIVED 26 May 2022

ACCEPTED 15 August 2022

PUBLISHED 02 September 2022

## CITATION

Fernández-Quintero ML, Fischer A-LM,  
Kokot J, Waibl F, Seidler CA and  
Liedl KR (2022) The influence of  
antibody humanization on shark  
variable domain (VNAR)  
binding site ensembles.  
*Front. Immunol.* 13:953917.  
doi: 10.3389/fimmu.2022.953917

## COPYRIGHT

© 2022 Fernández-Quintero, Fischer,  
Kokot, Waibl, Seidler and Liedl. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use,  
distribution or reproduction is  
permitted which does not comply with  
these terms.

# The influence of antibody humanization on shark variable domain (VNAR) binding site ensembles

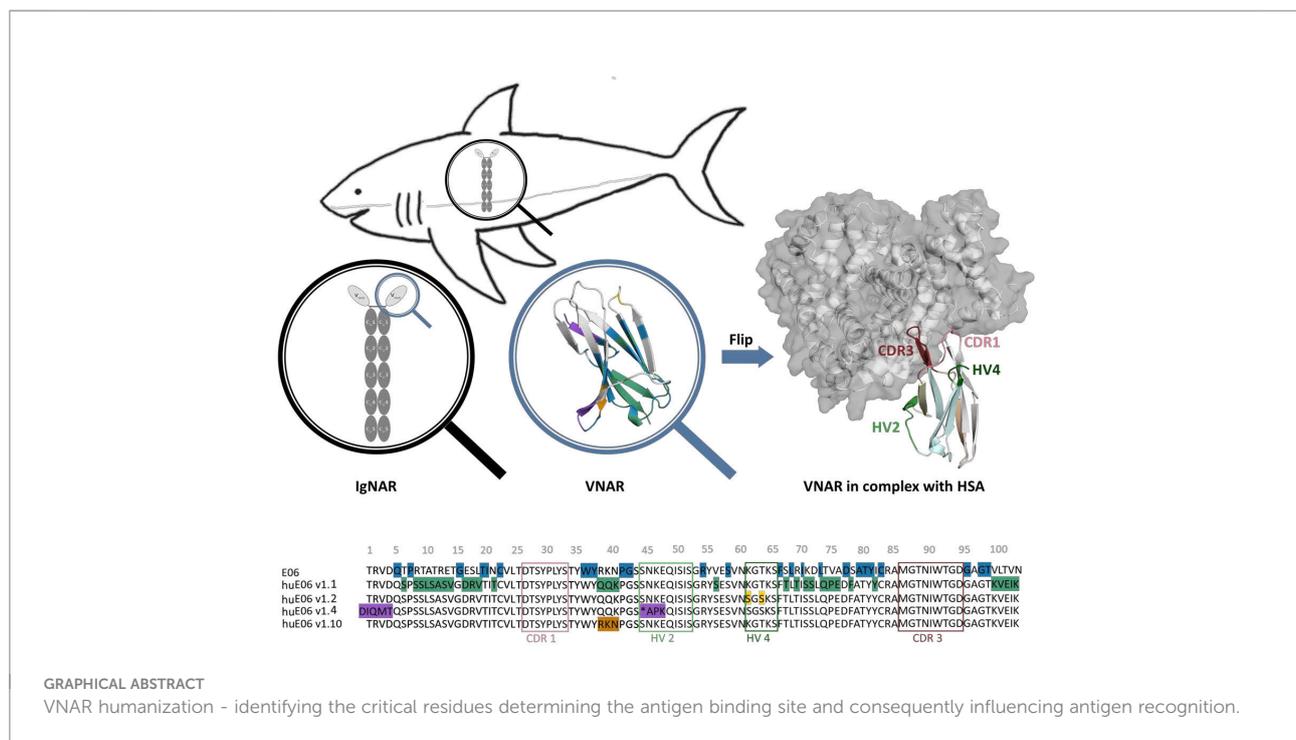
Monica L. Fernández-Quintero, Anna-Lena M. Fischer,  
Janik Kokot, Franz Waibl, Clarissa A. Seidler  
and Klaus R. Liedl\*

Department of General, Inorganic and Theoretical Chemistry, and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innsbruck, Austria

Sharks and other cartilaginous fish produce new antigen receptor (IgNAR) antibodies, as key part of their humoral immune response and are the phylogenetically oldest living organisms that possess an immunoglobulin (Ig)-based adaptive immune system. IgNAR antibodies are naturally occurring heavy-chain-only antibodies, that recognize antigens with their single domain variable regions (VNARs). In this study, we structurally and biophysically elucidate the effect of antibody humanization of a previously published spiny dogfish VNAR (parent E06), which binds with high affinity to the human serum albumin (HSA). We analyze different humanization variants together with the parental E06 VNAR and the human  $V\kappa 1$  light chain germline DPK9 antibody to characterize the influence of point mutations in the framework and the antigen binding site on the specificity of VNARs as reported by Kovalenko et al. We find substantially higher flexibility in the humanized variants, reflected in a broader conformational space and a higher conformational entropy, as well as population shifts of the dominant binding site ensembles in solution. A further variant, in which some mutations are reverted, largely restores the conformational stability and the dominant binding minimum of the parent E06. We also identify differences in surface hydrophobicity between the human  $V\kappa 1$  light chain germline DPK9 antibody, the parent VNAR E06 and the humanized variants. Additional simulations of VNAR-HSA complexes of the parent E06 VNAR and a humanized variant reveal that the parent VNAR features a substantially stronger network of stabilizing interactions. Thus, we conclude that a structural and dynamic understanding of the VNAR binding site upon humanization is a key aspect in antibody humanization.

## KEYWORDS

shark, VNAR, novel biotherapeutic formats, humanization, molecular dynamics simulations, hydrophobicity

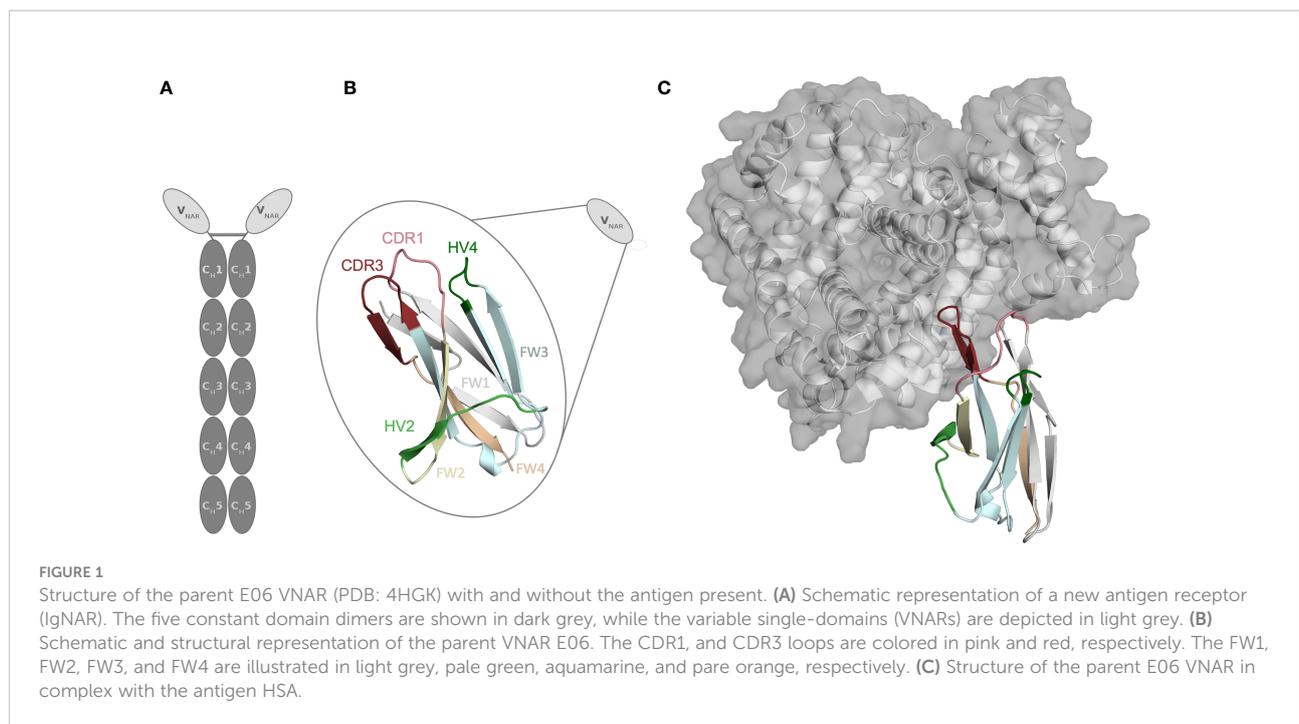


## Introduction

Approximately 500 million years ago, cartilaginous fish diverged from a common ancestor with other jawed vertebrates and can be divided into two extant subclasses, the holocephalans (chimeras and ratfish) and the elasmobranchs (sharks, rays and skates) (1–3). Cartilaginous fish are the phylogenetically oldest group of animals having an adaptive immune system, based on immunoglobulins (Ig), T-cell receptors and major histocompatibility complexes (MHC) (4–7). However, cartilaginous fish developed unique structural and immunological features and provide valuable insights in the evolution of the immune system (4, 8). Shark antibodies have evolved under challenging conditions, i.e., the high concentration of the protein denaturant urea in the blood serum and are therefore believed to be very stable (4, 9). The humoral immune response of sharks and their relatives comprises three Ig isotypes, namely IgM, IgW and so-called Ig new antigen receptor (IgNAR) (10). While the IgM and IgW are conventional heavy-light chain isotypes, the third atypical isotype IgNAR is a heavy-chain-only disulfide-bonded homodimer that does not pair with a light chain. The IgNAR consists of two identical heavy chains, which can be subdivided into two unpaired variable (VNAR) domains and five constant domain dimers ( $C_{H1}$ - $C_{H5}$ ) (Figure 1) (8, 11, 12). A schematic IgNAR structure, as well as an example of a VNAR crystal structure, is shown in Figures 1A, B (PDB accession code:

4HGK) (13). VNARs are small in size, soluble, exist as independent stable single-domains in solution and are known for their ability to bind and recognize a variety of antigens, including buried epitopes that are not accessible by conventional antibody variable domains (14–16). Due to their unique biophysical properties and characteristics, which are only shared by camelid single-chain antibodies, VNARs have received increasing attention as highly versatile proteins, which contribute to the success to date of alternative scaffolds and makes them attractive novel biotherapeutic proteins (14, 15, 17, 18). VNARs are structurally more similar to variable light chain and variable T-cell receptor domains than to variable heavy chain domains (19, 20). However, they contain only two complementarity determining region (CDR) loops, lacking the CDR2 loop and two  $\beta$ -strands. Instead, they contain other CDR like regions, namely the hypervariable region 2 (HV2), and the hypervariable region 4 (HV4), which reveal an elevated rate of somatic hypermutations. To compensate for the reduced size, the binding site features a long and extended CDR3 loop, which has the highest diversity in length, sequence, and structure (18, 21).

Another critical determinant of the structural diversity of VNARs is that they have further disulfide bonds in the CDRs and the framework region in addition to the canonical cysteine residues (Cys23-Cys88 – Kabat nomenclature) in the

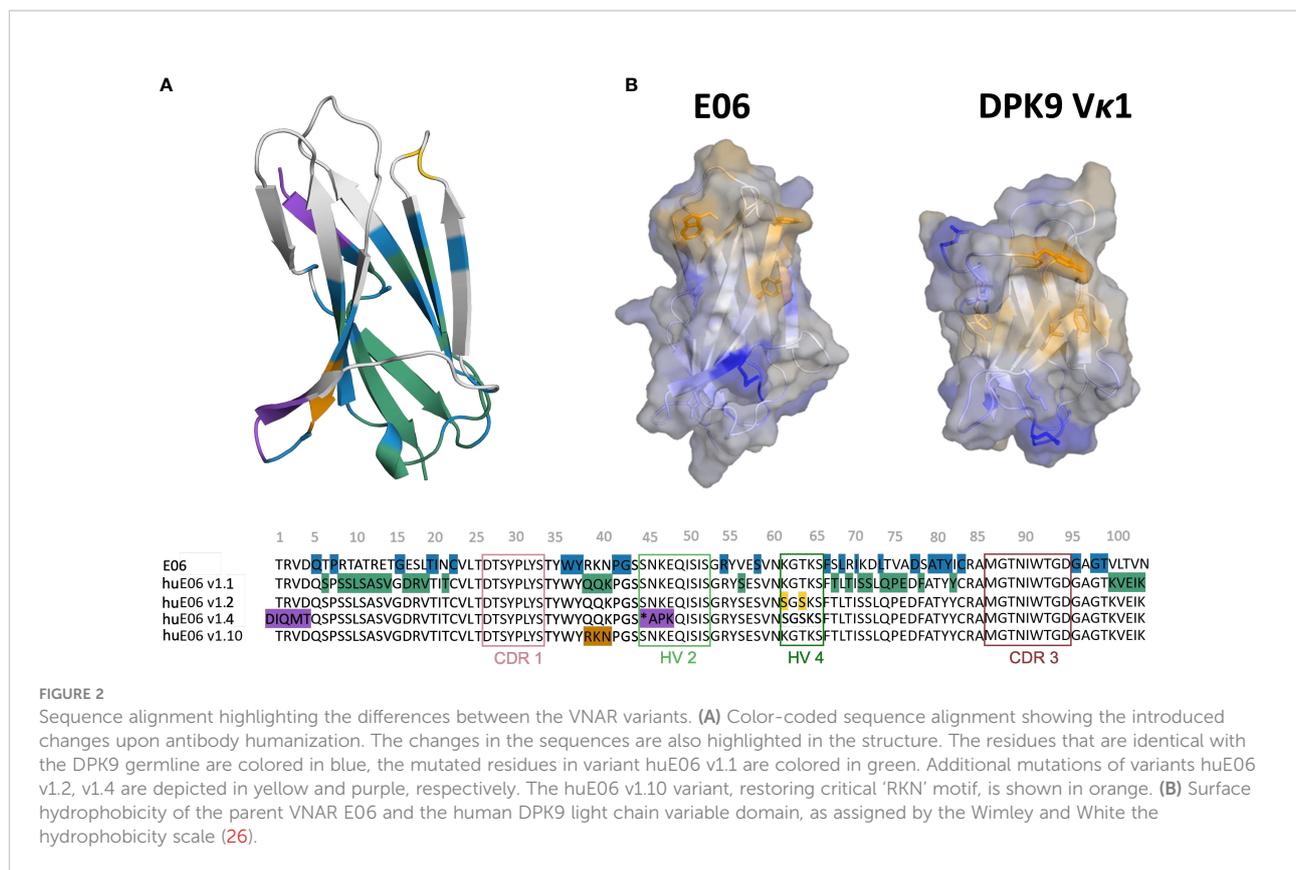


framework. Based on the number and position of additional cysteine residues, four types of naturally occurring IgNAR variable domains have been classified (4, 9, 12, 22, 23). Type I VNARs, which can be found in nurse sharks, contain two cysteines in CDR3 and two more in the framework regions 2 and 4. In contrast, type II and III VNARs have an additional cysteine pair to link CDR1 and CDR3 loops. Type III VNARs are predominantly present in neonatal shark development but have subsequently been found in adult spiny dogfish sharks and bamboo sharks. Structurally, these antibodies are characterized by a conserved tryptophan residue in the CDR1 loop and limited CDR3 loop diversity, however, the functional role of these antibodies remains still elusive. Type IV VNARs, which are primarily found in dogfish sharks, wobbegong, small-spotted catsharks and bamboo sharks contain only the two canonical cysteine residues (9, 13, 24, 25). In this study, we investigate the consequences and effects of antibody humanization of the spiny dogfish shark VNAR E06 (Figure 2). We thermodynamically and kinetically characterize the conformational diversity and the respective ensembles in solution of the E06 VNAR and compare them with different humanization variants published by Kovalenko et al., and the germline light chain V $\kappa$ 1 antibody, DPK9 (13). The investigated variants have been humanized by converting over 60% of non-CDR residues to those of the human germline and the resulting antibodies have mostly retained the specificity and affinity of the parent E06. For the first humanization variant (huE06 v1.1), the E06 VNAR has been humanized by replacing 63.5% of the framework residues (FW) of the VNAR E06, i.e., FW1 (residues 6-21), FW2 (residues 38-40), FW3 (residues 66-82) and FW4 (residues 99-103), with

residues from the human DPK9, while keeping the original shark residues for the first four N-terminal residues, CDR1, CDR3, HV2 and HV4 loop residues (Figure 2A). The DPK9 (V $\kappa$ 1) light chain was chosen for humanization, as it is one of the most stable and well-expressed human frameworks, that shares significant structural homology with E06 (27). The huE06 v1.2 variant was generated by introducing additional mutations in the HV4 loop. Additionally, mutations in the HV2 loop and in the N-terminus towards DPK9 were made to obtain huE06 v1.4. The variant E06 v1.10 was obtained by restoring critical contact residues (namely residues 38-40) with the antigen in FW2 (Figure 2A), which resulted in improved binding properties compared to E06 v1.1. Additionally, two crystal structures of the parental E06 VNAR and the huE06 v1.1 variant in complex with the HSA antigen were available (13). The two complex crystal structures show that not only CDR loops are responsible for interacting with the antigen, but that also several framework residues are in contact to HSA.

## Results

Where possible, we started our simulations from crystal structures. For, the other variants including the DPK9 (V $\kappa$ 1) light chain variable domain, we generated models by using “AlphaFold2” (Table 1) (13, 28). In Figure 2B, we present the surface hydrophobicity of the parent VNAR E06 and the human DPK9 light chain variable domain, as assigned using the hydrophobicity scale by Wimley and White (26). It shows that the DPK9 variable domain displays a hydrophobic patch at the



center of the interface, which corresponds to the contact area with the respective paired heavy chain. This hydrophobic patch in the DPK9 interface is replaced by a more hydrophilic surface in the E06 VNAR.

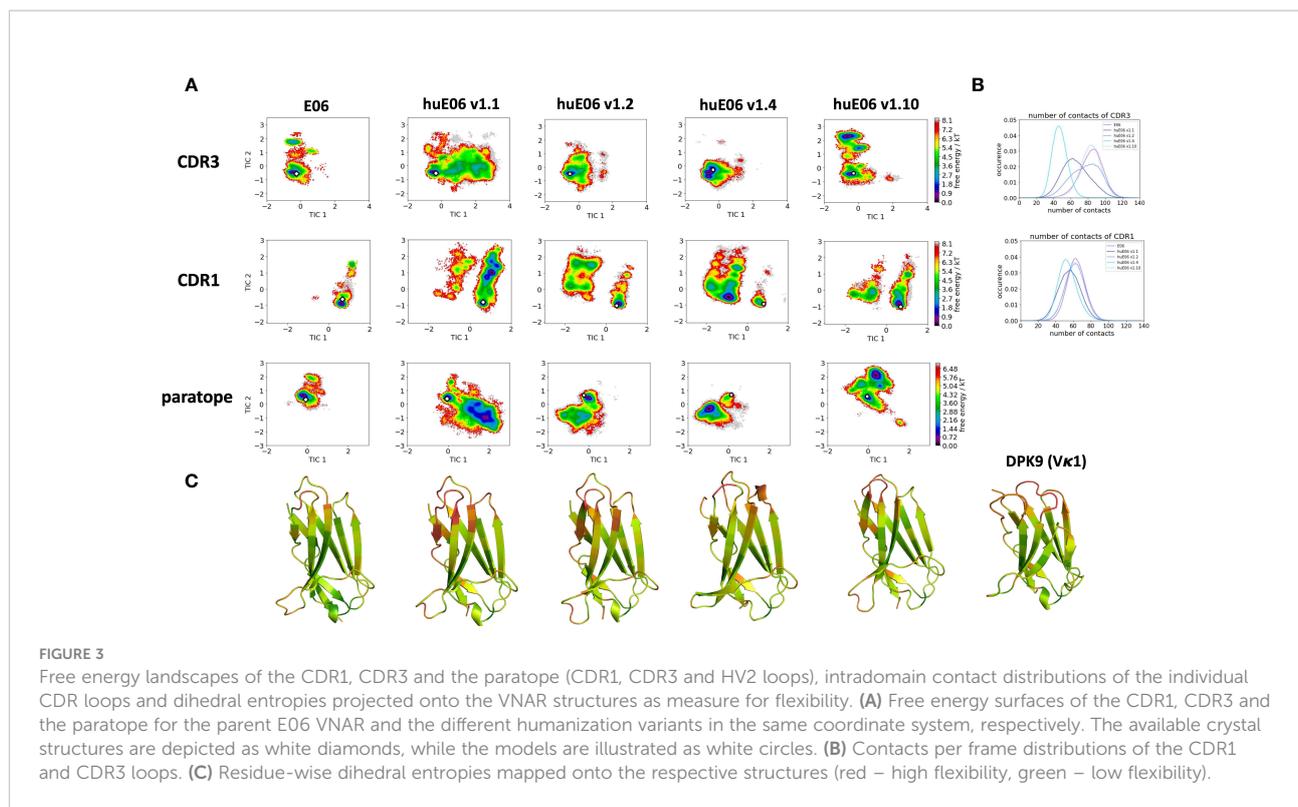
To characterize the conformational diversity of the E06 VNAR and the humanized variants in solution, we applied a protocol using the enhanced sampling technique metadynamics in combination with classical molecular dynamics (MD) simulations, to overcome the limitations of conformational sampling imposed by high energy barriers. The aggregated simulation time for each variant is summarized in [Table 1](#).

Moreover, we wanted to characterize and quantify the conformational diversity in the different humanized variants.

**TABLE 1** Overview of the investigated variants with the aggregated simulation times.

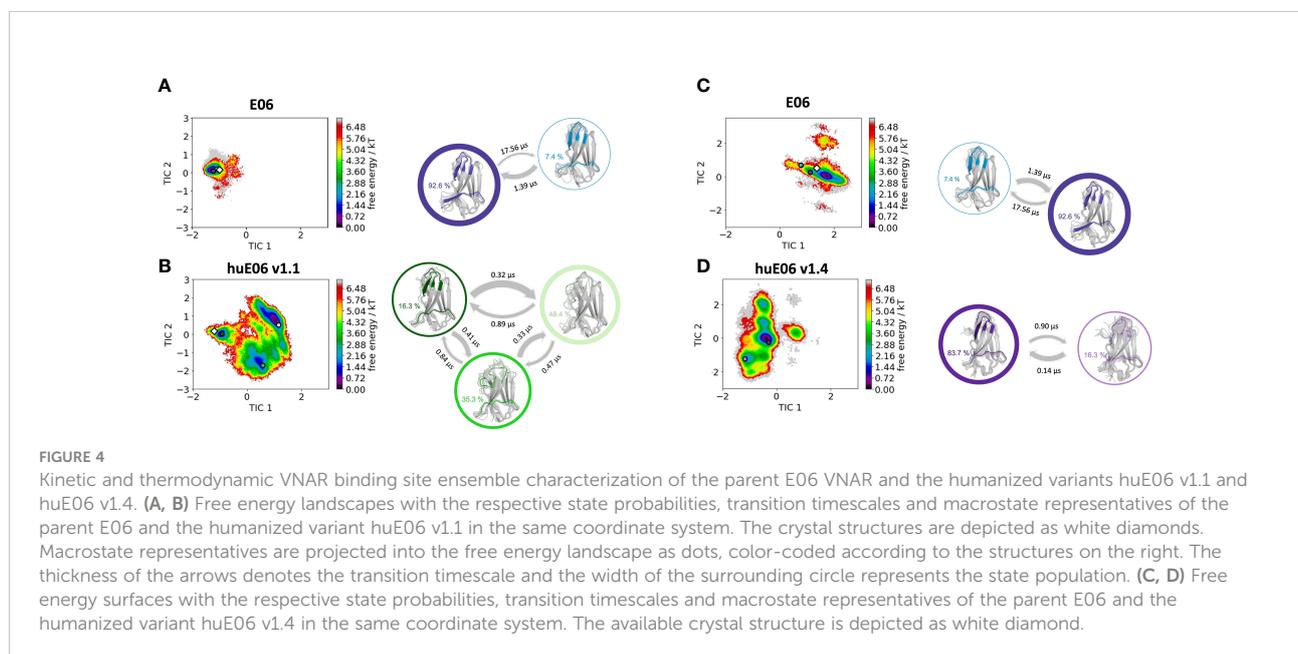
Variants	Starting structures	Simulation time/ $\mu$ s
E06	PDB: 4HGK	11.9
huE06 v1.1	PDB: 4HGM	28.2
huE06 v1.2	AlphaFold2	17.4
huE06 v1.4	AlphaFold2	43.1
huE06 v1.10	AlphaFold2	20.0
DPK9 (Vκ1)	AlphaFold2	33.8

[Figure 3A](#) shows the time-lagged independent component analysis (tICA) plots of the CDR3, CDR1 and of the whole paratope (comprising CDR1, CDR3 and the HV2 loop). tICA is a dimensionality-reduction technique, which detects the slowest-relaxing degrees of freedom. The free energy landscapes of the paratope clearly show that the parent E06 VNAR is less flexible than the humanized variants. This difference in conformational diversity is also reflected in the free energy surfaces of the individual CDR1 and CDR3 loops, which reveal different conformational states in solution. [Figure 3B](#) shows histograms of the number of intradomain contacts of the CDR3 and the CDR1, respectively. There is a clear reduction of the number of contacts in those variants which also show higher flexibility in the free energy surfaces. In agreement with experimental affinity and specificity measures determined by Kovalenko et al. (13), in which huE06 v1.10 is the closest to the parent E06 VNAR, we find that the parent E06 and the huE06 v1.10 display the highest number of intradomain contacts per frame. [Figure 3C](#) visualizes the residue-wise dihedral entropies projected onto the respective E06 VNAR and the humanized variant structures. We find clear differences in the dihedral entropy of the CDR loops and the HV2 and HV4 loops between the different variants and the parent VNAR. The huE06 v1.10 variant and the parent VNAR show the most limited flexibility, while the variant huE06 v1.4



reveals together with the germline DPK9 variable domain and the huE06 v1.1 variant the highest conformational diversity. A direct comparison of the CDR3 loop conformational spaces of the parent, the humanized variants, and the germline DPK9 variable domain is shown in SI Figure S1. The germline CDR3 loop reveals a high flexibility, reflected in a broader

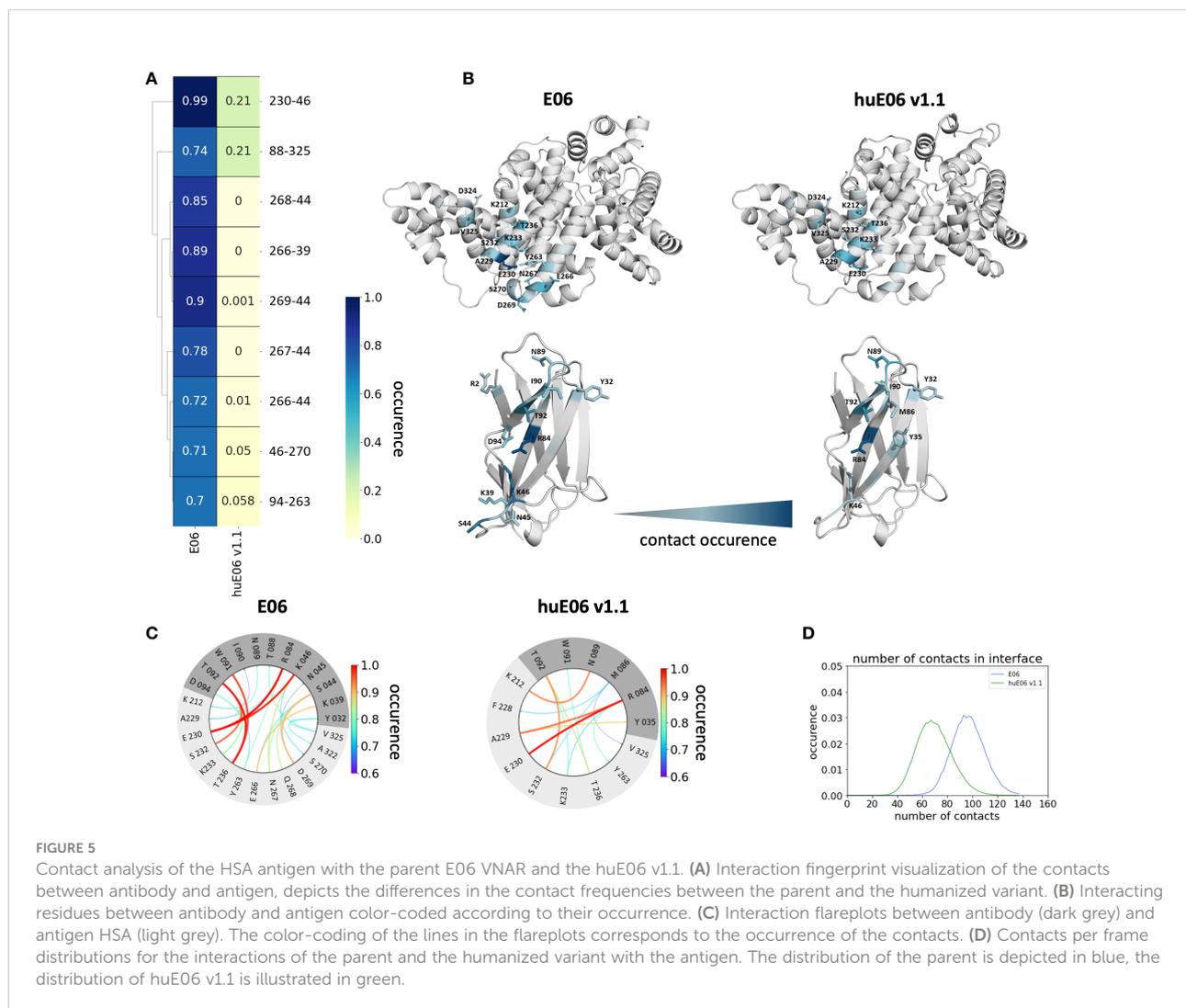
conformational space and a different dominant minimum in solution. To reconstruct thermodynamics and kinetics of different loop rearrangements, we built Markov-state models of the paratope based on the backbone torsions of the CDR1, CDR3, and HV2 (Figure 4). Figure 4 compares the parent E06 with the variant huE06 v1.1 (Figures 4A, B) and with the variant



huE06 v1.4 (Figures 4C, D) in the same coordinate system, respectively. We find overlaps of the conformational spaces of the parent and the humanized variants; however, the variants and the parent differ in flexibility and in their state probabilities. While the available complex X-ray structure (depicted as white diamond) lies in the dominant minimum in solution (92% probability) of the parent E06 VNAR, the probability of this state is substantially reduced to 16% in the huE06 v1.1 variant. The difference is more drastic for the comparison of the parent E06 VNAR with the huE06 v1.4.

In line with the results obtained for huE06 v1.1, we observe a bigger conformational landscape and find a substantially reduced probability for the binding competent state of the huE06 v1.4 (16%). As described in the methods section, we also performed simulations in complex with the antigen, as both available crystal structures were in complex with the HSA antigen. We do not only find a substantial increase of

flexibility in the antigen-binding site for the huE06 v1.1, compared to the parent E06 VNAR, but we also find substantial differences in the interaction networks. Figure 5 shows the interaction patterns of the parent E06-HSA complex compared with the huE06 v1.1-HSA complex. We observe a shift in the contact per frame distributions towards a lower number of contacts for the huE06 v1.1 variant (Figure 5D). This substantial reduction in the interaction network is also visualized in the antibody-antigen interaction fingerprints and flareplots (Figures 5A, C). The main differences in antibody-antigen interactions are also visualized in the interaction fingerprint plots, which provide a clustering of interactions based on their contact frequencies. These findings agree with the reduced binding affinity and specificity of the huE06 v1.1 to the HSA compared to the parent E06 VNAR and indicate that they result from missing interactions with the HV2 loop and the framework.



## Discussion

The rise of antibodies as biotherapeutic proteins has motivated numerous studies to characterize and understand the antibody binding interface as a pre-requisite for rational antibody design and engineering (29–35). Compared to conventional antibodies, small antibodies, such as nanobodies and VNARs, offer various advantages and reveal features that are desirable for drug discovery, i.e., higher stability and solubility. Additionally, they can work inside cells, recognize cryptic and buried epitopes and due to their small size wend into tissues (21). Furthermore, it has also been reported that VNARs can be used as blood-brain barrier (BBB) shuttles or transporter molecules (36).

Thus, to design and engineer these outstanding proteins, it is crucial to characterize the peculiar antibody binding site of VNARs structurally and dynamically and to elucidate the antibody-antigen recognition mechanism. In this study, we thermodynamically and kinetically characterize the consequences of humanization on VNAR binding site ensembles in solution and provide a description of the fundamental factors that contribute to antigen binding. Engineering efforts focus on reducing undesirable immunogenic responses by antibody humanization, thereby increasing the identity of non-human antibodies and scaffolds to common human antibody sequences (37–41). The main challenge in humanizing antibodies is to maintain the full biological function, which is reflected in a high binding affinity and specificity to reduce the risk of adverse side-effects (42). Therefore, to understand the role of the framework on the antigen binding site, it is crucial to biophysically characterize antibody paratope ensembles in solution in different stages of antibody humanization (38). Conformational rearrangements in the paratope, as well as antibody-antigen binding can occur in the nano-to-millisecond timescale, which exceeds routinely performed simulation times by far (43, 44). To enhance the sampling efficiency, we use metadynamics simulations to cover the relevant conformational transitions and paratope rearrangements. Figure 2A shows the sequence alignment of the parent E06 VNAR with the investigated humanized variants and highlights the importance of specific framework residues for antigen recognition. Figure 1C shows the atypical binding mode of the antigen HSA to the VNAR, as not only the CDR loops are involved in binding the antigen, but also the HV2 and extensive framework residues.

It has already been shown that framework residues can determine the binding site ensembles and consequently influence antibody-antigen binding (13, 37, 39, 45, 46). Additionally, Figure 2B depicts the surface hydrophobicity of the parent E06 VNAR compared to the human germline DPK9. We find that the parent E06 VNAR reveals a more hydrophilic interface, which is believed to be responsible for the enhanced stability and the favorable biophysical characteristics (17, 18).

Figure 3 provides an overview of the CDR loop and paratope conformational spaces and clearly shows that the shape of the paratope is not solely determined by the CDR3 loop, but is actually also influenced by the CDR1 and the HV2 loops. Compared to the parent E06 VNAR, we find that in line with the decrease in specificity reported by Kovalenko et al. (13), we observe an increase in flexibility for all investigated humanized variants and the human germline DPK9 variable domain. This increase in flexibility is reflected in the broader conformational landscapes and the increased residue-wise dihedral entropies. The residue-wise dihedral entropies provide an alignment-independent measure for flexibility and thereby facilitate the comparison of flexibility hotspots within the different variants (47). The changes in flexibility underline the challenges in rationally designing antibodies, by revealing the presence of conformational substates, which are likely to have different binding properties and may result in a high entropic cost upon binding (48). These observations presented in Figure 3 also agree with previous studies showing that affinity maturation of single-domain variable domains results in a decrease of flexibility (49–52). Apart from differences in flexibility, we observe substantial shifts of the contact per frame distributions (Figure 3B), in particular for the humanized variants huE06 v1.1, v1.2 and v1.4. Thus, we find that not only the direct interactions with the antigen are influenced by humanization (Figure 5), but also the intramolecular interaction network. Therefore, the higher flexibility of the variants can also be explained by a weaker intramolecular interaction network. The lowest number of contacts per frame can be found for the huE06 v1.4 variant, which lacks a charged critical hydrogen bond interaction of residue K64 in the HV4 loop and residue Y29 in the CDR1 loop and forms instead a less probable interaction between residues T27-Y29 in the CDR1 loop. The huE06 v1.10 variant on the other hand restores the ‘RKN’ motif located in FW2, which is crucial for antigen-binding and reveals therefore a highly similar intramolecular interaction network compared to the parent E06 VNAR. The Markov-state models in Figure 4 compare the parent E06 VNAR paratope states with the humanized variants huE06 v1.1 and huE06 v1.4. We find that the binding competent conformation is present in all investigated variants, however with reduced probability. These findings support the idea that the decrease in specificity is accompanied by a population shift, reflected in different dominant states in solution as a consequence of antibody humanization. The simulations with the antigen HSA reveal that the huE06 v1.1 displays also in complex an increase in flexibility compared to the parent E06 VNAR-HSA complex (SI Figure S2). Figure 5 shows the interaction profiles and networks of the parent E06 VNAR and the huE06 v1.1 variant and reveals a substantially reduced number of interactions for the huE06 v1.1. The higher flexibility of the huE06 v1.1 variant can be explained by the absence of the stabilizing salt bridge interaction between residue K46 in the HV2 loop with residue E230 in the antigen and the

hydrogen bond interaction of residue W91 in the CDR3 loop with residue T236.

## Conclusion

In this study, we structurally and functionally characterized the antigen-binding site of five VNARs upon antibody humanization. We observed, in line with the findings of Kovalenko et al., that not solely the CDR1 and CDR3 loops are critical for determining the shape of the antigen-binding site, but that also the HV2 loop and the FW2 are critical for antigen recognition. The germline DPK9 as well as the humanized variants reveal a higher flexibility, which is reflected in a higher conformational entropy, a broader conformational space and substantial population shifts of the dominant binding site ensembles in solution. The huE06 v1.10 variant restores three critical framework residues and retains the conformational stability and the dominant binding minimum of the parent. Additionally, the simulations in complex with the antigen reveal that the huE06 v1.1 variant shows a higher flexibility and variability in the antigen-binding site compared to the parent E06 VNAR. Apart from the higher conformational variability, the VNAR-HSA complexes differ in the duration and number of antibody-antigen interactions. The increase in flexibility upon antibody humanization can be explained by a lack of stabilizing interactions, due to changes in the framework. Thus, we conclude that the dynamic and structural characterization of VNAR binding site ensembles allows to identify the key determinants for antigen recognition and can guide antibody humanization efforts.

## Methods

A previously published method characterizing the CDR loop ensembles in solution (32, 33, 38, 50, 51, 53, 54) was used to investigate the conformational diversity of the paratope loops of VNAR humanization variants both with and without the antigen bound (13). Experimental structure information was available for the parent E06 VNAR and the humanized variant huE06 v1.1, which were crystallized with the antigen HSA. The PDB accession codes for the parent E06 and the huE06 v1.1 are 4HGK and 4HGM, respectively. For the other investigated variants and the human  $V_{\kappa}1$  germline DPK9, we used AlphaFold2 to predict the structures (28). The available X-ray structures and the models were used as starting structures for molecular dynamics simulations. The starting structures for simulations were prepared in MOE (Molecular Operating Environment, Chemical Computing Group, version 2020.09) using the Protonate3D tool (55, 56). To neutralize the charges we used the uniform background charge (57, 58). Using the tleap tool of the AmberTools20 (57, 59) package, the structures were soaked

in cubic water boxes of TIP3P water molecules with a minimum wall distance of 10 Å to the protein (60–62). For all simulations, parameters of the AMBER force field 14SB were used (63). The VNAR variants were carefully equilibrated using a multistep equilibration protocol (64).

## Metadynamics simulations

To enhance the sampling of the conformational space, well-tempered metadynamics simulations (65–68) were performed in GROMACS (69, 70) with the PLUMED 2 implementation (71). As collective variables, we used a linear combination of sine and cosine of the  $\psi$  torsion angles of the CDR1 and CDR3 loops calculated with functions MATHEVAL and COMBINE implemented in PLUMED 2 (71). As discussed previously the  $\psi$  torsion angle captures conformational transitions comprehensively (72). The decision to include the  $\psi$  torsion angles of these two loops is based on their strong involvement in the binding to the antigen as evident from the X-ray structure of the complex. The simulations were performed at 300 K in an NpT ensemble using the velocity rescaling algorithm and a Parrinello-Rahman barostat (73, 74). For the metadynamics simulations, we used a Gaussian height of 10.0 kJ/mol and a width of 0.3 radian. Gaussian deposition occurred every 1000 steps and a biasfactor of 10 was used. 1  $\mu$ s metadynamics simulations were performed for the parent E06 VNAR, the humanized variants and the DPK9 human variable domain without the antigen. As the available X-ray structures for the parent E06 and the huE06 v1.1 were crystallized with the antigen present, we also performed 1  $\mu$ s of metadynamics simulations in complex with the antigen for these two systems. The resulting trajectories were clustered in cpptraj (57, 59) by using the average linkage hierarchical clustering algorithm with a distance cut-off criterion of 1.3 Å resulting in a large number of clusters (Table 1). The cluster representatives for the parent and the humanized variants both with and without the antigen present were equilibrated and simulated for 100 ns each using the AMBER 20 simulation package (57).

## Molecular dynamics simulations

Molecular dynamics simulations were performed in an NpT ensemble using pmemd.cuda (75). Bonds involving hydrogen atoms were restrained by applying the SHAKE algorithm (76), allowing a time step of 2 fs. Atmospheric pressure of the system was preserved by weak coupling to an external bath using the Berendsen algorithm (77). The Langevin thermostat (78, 79) was used to maintain the temperature during simulations at 300 K.

Based on the backbone torsion of the CDR1, CDR3 and HV2 loops, a time-lagged independent component analysis (tICA) was performed using the python library PyEMMA 2 employing a lag time of 10 ns (80, 81). Thermodynamics and kinetics were calculated

from a Markov-state model (82, 83) by using PyEMMA 2, which uses the k-means clustering algorithm (84) to define microstates and the PCCA+ clustering algorithm (85) to coarse-grain the microstates into macrostates. PCCA+ is a spectral clustering method, which discretizes the sampled conformational space based on the eigenvectors of the transition matrix. The sampling efficiency and the reliability of the Markov-state model (e.g., defining optimal feature mappings) has been evaluated with the Chapman-Kolmogorov test (86, 87), by using the variational approach for Markov processes (88) and by taking into account the fraction of states used, as the network states must be fully connected to calculate probabilities of transitions and the relative equilibrium probabilities. To capture and quantify the kinetically relevant loop rearrangements of the VNAR variants we constructed Markov-state models based on the backbone torsions of the CDR1, CDR3 and HV2 loops, defined 100 microstates using the k-means clustering algorithm and applied a lag time of 15 ns. The images presented in this paper were created by using the PyMOL molecular graphics system (89).

## Dihedral entropies

We calculated the residue-wise dihedral entropies with the recently published X-entropy python package, which calculates the entropy of a given dihedral angle distribution (47). This approach uses a Gaussian kernel density estimation (KDE) with a plug-in bandwidth selection, which is fully implemented in C++ and parallelized with OpenMP. The obtained residue-wise dihedral entropies were projected onto the respective structures.

## Contacts

The GetContacts (90) tool was used to quantify the interactions occurring in the simulations. To generate the contact per frame histograms, we chose a bin width of 2. Inspired by their visualization tools a new script was developed. The script includes “interaction fingerprints” which facilitate the comparison of interactions among multiple systems as well as “flareplots” to visualize the interaction networks of individual systems. The “interaction fingerprint” plots are based on a hierarchical clustering on the data to compare the contact frequencies of different systems. We used as clustering criterion the contact frequencies and applied the cut-off 0.5. In the fingerprint plots interacting residues are connected *via* a line that is colored according to the frequency of the interaction. The script as well as a short introduction is provided on our GitHub ([https://github.com/riedlab/GetContacts\\_analysis](https://github.com/riedlab/GetContacts_analysis)).

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author

## Author contributions

MF-Q performed research, wrote the manuscript. A-LF performed research and analyzed data. JK performed research. FW analyzed data and contributed in writing the manuscript. CS analyzed data. KL supervised the research. All authors contributed to writing the manuscript.

## Funding

This work was supported by the Austrian Science Fund (FWF) *via* the grants P30565, P30737 and P30402, P34518 as well as DOC 30.

## Acknowledgments

The computational results presented here have been achieved (in part) using the Vienna Scientific Cluster (VSC). We acknowledge PRACE for awarding us access to Piz Daint at CSCS, Switzerland.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.953917/full#supplementary-material>

## References

- Dooley H, Flajnik MF. Antibody repertoire development in cartilaginous fish. *Dev Comp Immunol* (2006) 30(1):43–56. doi: 10.1016/j.dci.2005.06.022
- Flajnik MF, Kasahara M. Origin and evolution of the adaptive immune system: Genetic events and selective pressures. *Nat Rev Genet* (2010) 11(1):47–59. doi: 10.1038/nrg2703
- Cooper MD, Alder MN. The evolution of adaptive immune systems. *Cell* (2006) 124(4):815–22. doi: 10.1016/j.cell.2006.02.001
- Feige MJ, Gräwert MA, Marciniowski M, Hennig J, Behnke J, Ausländer D, et al. The structural analysis of shark IgNAR antibodies reveals evolutionary principles of immunoglobulins. *Proc Natl Acad Sci USA* (2014) 111(22):8155. doi: 10.1073/pnas.1321502111
- Flajnik MF. A cold-blooded view of adaptive immunity. *Nat Rev Immunol* (2018) 18(7):438–53. doi: 10.1038/s41577-018-0003-9
- Frommel D, Litman GW, Finstad J, Good RA. The evolution of the immune response. *J Immunol* (1971) 106(5):1234.
- Crisciello MF, Saltis M, Flajnik MF. An evolutionarily mobile antigen receptor variable region gene: Doubly rearranging NAR-TcR genes in sharks. *Proc Natl Acad Sci U.S.A.* (2006) 103(13):5036–41. doi: 10.1073/pnas.0507074103
- Diaz M, Stanfield RL, Greenberg AS, Flajnik MF. Structural analysis, selection, and ontogeny of the shark new antigen receptor (IgNAR): Identification of a new locus preferentially expressed in early development. *Immunogenetics* (2002) 54(7):501–12. doi: 10.1007/s00251-002-0479-z
- Matz H, Dooley H. Shark IgNAR-derived binding domains as potential diagnostic and therapeutic agents. *Dev Comp Immunol* (2019) 90:100–7. doi: 10.1016/j.dci.2018.09.007
- Hsu E. Assembly and expression of shark ig genes. *J Immunol* (2016) 196(9):3517–23. doi: 10.4049/jimmunol.1600164
- Zielonka S, Empting M, Grzeschik J, Könning D, Barelle CJ, Kolmar H. Structural insights and biomedical potential of IgNAR scaffolds from sharks. *null* (2015) 7(1):15–25. doi: 10.4161/19420862.2015.989032
- Roux KH, Greenberg AS, Greene L, Strelets L, Avila D, McKinney EC, et al. Structural analysis of the nurse shark (New) antigen receptor (NAR): Molecular convergence of NAR and unusual mammalian immunoglobulins. *Proc Natl Acad Sci U.S.A.* (1998) 95(20):11804–9. doi: 10.1073/pnas.95.20.11804
- Kovalenko OV, Olland A, Piché-Nicholas N, Godbole A, King D, Svenson K, et al. Atypical antigen recognition mode of a shark immunoglobulin new antigen receptor (IgNAR) variable domain characterized by humanization and structural analysis. *J Biol Chem* (2013) 288(24):17408–19. doi: 10.1074/jbc.M112.435289
- English H, Hong J, Ho M. Ancient species offers contemporary therapeutics: An update on shark VNAR single domain antibody sequences, phage libraries and potential clinical applications. *Antibody Ther* (2020) 3(1):1–9. doi: 10.1093/abt/taaa001
- Barelle C, Gill DS, Charlton K. Shark novel antigen receptors—the next generation of biologic therapeutics? *Adv Exp Med Biol* (2009) 655:49–62. doi: 10.1007/978-1-4419-1132-2\_6
- Clem LW, Leslie GA. Phylogeny of immunoglobulin structure and function. XIV. peptide map and amino acid composition studies of shark antibody light chains. *Dev Comp Immunol* (1982) 6(2):263–9. doi: 10.1016/S0145-305X(82)80009-8
- Hoey RJ, Eom H, Horn JR. Structure and development of single domain antibodies as modules for therapeutics and diagnostics. *Exp Biol Med (Maywood)* (2019) 244(17):1568–76. doi: 10.1177/1535370219881129
- Chiu ML, Goulet DR, Teplyakov A, Gilliland GL. Antibody structure and function: The basis for engineering therapeutics. *Antibodies (Basel)* (2019) 8(4):55. doi: 10.3390/antib8040055
- Flajnik M, Rumpf L. Early and natural antibodies in non-mammalian vertebrates. *Curr topics Microbiol Immunol* (2000) 252:233–40. doi: 10.1007/978-3-642-57284-5\_24
- Greenberg AS, Avila D, Hughes M, Hughes A, McKinney EC, Flajnik MF. A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. *Nature* (1995) 374(6518):168–73. doi: 10.1038/374168a0
- Griffiths K, Dolezal O, Parisi K, Angerosa J, Dogovski C, Barraclough M, et al. Shark variable new antigen receptor (VNAR) single domain antibody fragments: Stability and diagnostic applications. *Antibodies* (2013) 2:(1). doi: 10.3390/antib2010066
- Rumpf LL, Avila D, Diaz M, Bartl S, McKinney EC, Flajnik MF. A shark antibody heavy chain encoded by a nonsomatically rearranged VDJ is preferentially expressed in early development and is convergent with mammalian IgG. *Proc Natl Acad Sci USA* (2001) 98(4):1775. doi: 10.1073/pnas.98.4.1775
- Streltsov VA, Varghese JN, Carmichael JA, Irving RA, Hudson PJ, Nuttall SD. Structural evidence for evolution of shark ig new antigen receptor variable domain antibodies from a cell-surface receptor. *Proc Natl Acad Sci U.S.A.* (2004) 101(34):12444. doi: 10.1073/pnas.0403509101
- Steven J, Müller MR, Carvalho MF, Ubah OC, Kovaleva M, Donohoe G, et al. *In vitro* maturation of a humanized shark VNAR domain to improve its biophysical properties to facilitate clinical development. *Front Immunol* (2017) 8. doi: 10.3389/fimmu.2017.01361
- Wei L, Wang M, Xiang H, Jiang Y, Gong J, Su D, et al. Bamboo shark as a small animal model for single domain antibody production. *Front Bioengineering Biotechnol* (2021) 9. doi: 10.3389/fbioe.2021.792111
- Wimley WC, White SH. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat Struct Biol* (1996) 3(10):842–8. doi: 10.1038/nsb1096-842
- Ewert S, Huber T, Honegger A, Plückthun A. Biophysical properties of human antibody variable domains. *J Mol Biol* (2003) 325(3):531–53. doi: 10.1016/S0022-2836(02)01237-8
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature* (2021) 596(7873):583–9. doi: 10.1038/s41586-021-03819-2
- MacCallum RM, Martin ACR, Thornton JM. Antibody-antigen interactions: Contact analysis and binding site topography. *J Mol Biol* (1996) 262(5):732–45. doi: 10.1006/jmbi.1996.0548
- Di Palma F, Tramontano A. Dynamics behind affinity maturation of an anti-HCMV antibody family influencing antigen binding. *FEBS Lett* (2017) 591(18):2936–50. doi: 10.1002/1873-3468.12774
- Schmidt AG, Xu H, Khan AR, O'Donnell T, Khurana S, King LR, et al. Preconfiguration of the antigen-binding site during affinity maturation of a broadly neutralizing influenza virus antibody. *Proc Natl Acad Sci USA* (2013) 110(1):264. doi: 10.1073/pnas.1218256109
- Fernández-Quintero ML, Kraml J, Georges G, Liedl KR. CDR-H3 loop ensemble in solution – conformational selection upon antibody binding. *mAbs* (2019) 11:1077–88. doi: 10.1080/19420862.2019.1618676
- Fernández-Quintero ML, Loeffler JR, Kraml J, Kahler U, Kamenik AS, Liedl KR. Characterizing the diversity of the CDR-H3 loop conformational ensembles in relationship to antibody binding properties. *Front Immunol* (2019) 9:3065. doi: 10.3389/fimmu.2018.03065
- Fernández-Quintero ML, Loeffler JR, Waibl F, Kamenik AS, Hofer F, Liedl KR. Conformational selection of allergen-antibody complexes—surface plasticity of paratopes and epitopes. *Protein Engineering Design Selection* (2020) 32(11):513–23. doi: 10.1093/protein/gzaa014
- Fernández-Quintero ML, Pomarici ND, Math BA, Kroell KB, Waibl F, Bujotzek A, et al. Antibodies exhibit multiple paratope states influencing VH–VL domain orientations. *Commun Biol* (2020) 3(1):589. doi: 10.1038/s42003-020-01319-z
- Stocki P, Szary JM, Jacobsen CL, Demydchuk M, Northall L, Moos T, et al. High efficiency blood-brain barrier transport using a VNAR targeting the transferrin receptor 1 (TfR1). *bioRxiv* (2019) 816900:1–19. doi: 10.1101/816900
- Margreitter C, Mayrhofer P, Kunert R, Oostenbrink C. Antibody humanization by molecular dynamics simulations-in-Silico guided selection of critical backmutations. *J Mol Recognit* (2016) 29(6):266–75. doi: 10.1002/jmr.2527
- Fernández-Quintero ML, Heiss MC, Liedl KR. Antibody humanization—the influence of the antibody framework on the CDR-H3 loop ensemble in solution. *Protein Engineering Design Selection* (2020) 32:gzaa004. doi: 10.1093/protein/gzaa004
- Apgar JR, Mader M, Agostinelli R, Benard S, Bialek P, Johnson M, et al. Beyond CDR-grafting: Structure-guided humanization of framework and CDR regions of an anti-myostatin antibody. *MAbs* (2016) 8(7):1302–18. doi: 10.1080/19420862.2016.1215786
- Fransson J, Teplyakov A, Raghunathan G, Chi E, Cordier W, Dinh T, et al. Human framework adaptation of a mouse anti-human IL-13 antibody. *J Mol Biol* (2010) 398(2):214–31. doi: 10.1016/j.jmb.2010.03.004
- Teplyakov A, Obmolova G, Malia TJ, Raghunathan G, Martinez C, Fransson J, et al. Structural insights into humanization of anti-tissue factor antibody 10H10. *MAbs* (2018) 10(2):269–77. doi: 10.1080/19420862.2017.1412026
- Shankar G, Shores E, Wagner C, Mire-Sluis A. Scientific and regulatory considerations on the immunogenicity of biologics. *Trends Biotechnol* (2006) 24(6):274–80. doi: 10.1016/j.tibtech.2006.04.001
- Henzler-Wildman K, Kern D. Dynamic personalities of proteins. *Nature* (2007) 450(7172):964–72. doi: 10.1038/nature06522

44. Kern D. From structure to mechanism: Skiing the energy landscape. *Nat Methods* (2021) 18(5):435–6. doi: 10.1038/s41592-021-01140-4
45. Fernández-Quintero ML, Kroell KB, Hofer F, Riccabona JR, Liedl KR. Mutation of framework residue H71 results in different antibody paratope states in solution. *Front Immunol* (2021) 12:630034. doi: 10.3389/fimmu.2021.630034
46. Makabe K, Nakanishi T, Tsumoto K, Tanaka Y, Kondo H, Umetsu M, et al. Thermodynamic consequences of mutations in vernier zone residues of a humanized anti-human epidermal growth factor receptor murine antibody, 528. *J Biol Chem* (2008) 283(2):1156–66. doi: 10.1074/jbc.M706190200
47. Kraml J, Hofer F, Quoika PK, Kamenik AS, Liedl KR. X-Entropy: A parallelized kernel density estimator with automated bandwidth selection to calculate entropy. *J Chem Inf Model* (2021) 61(4):1533–8. doi: 10.1021/acs.jcim.0c01375
48. Löhr T, Sormanni P, Vendruscolo M. Conformational entropy as a potential liability of computationally designed antibodies. *Biomolecules* (2022) 12:1–10. doi: 10.3390/biom12050718
49. Fernández-Quintero ML, Seidler CA, Quoika PK, Liedl KR. Shark antibody variable domains rigidify upon affinity maturation—understanding the potential of shark immunoglobulins as therapeutics. *Front Mol Biosci* (2021) 8:639166. doi: 10.3389/fmolb.2021.639166
50. Fernández-Quintero ML, Seidler CA, Liedl KR. T-Cell receptor variable  $\beta$  domains rigidify during affinity maturation. *Sci Rep* (2020) 10(1):4472. doi: 10.1038/s41598-020-61433-0
51. Fernández-Quintero ML, DeRose EF, Gabel SA, Mueller GA, Liedl KR. Nanobody paratope ensembles in solution characterized by MD simulations and NMR. *Int J Mol Sci* (2022) 23(10):1–12. doi: 10.3390/ijms23105419
52. Stanfield RL, Dooley H, Verdino P, Flajnik MF, Wilson IA. Maturation of shark single-domain (IgNAR) antibodies: Evidence for induced-fit binding. *J Mol Biol* (2007) 367(2):358–72. doi: 10.1016/j.jmb.2006.12.045
53. Fernández-Quintero ML, Heiss MC, Pomarici ND, Math BA, Liedl KR. Antibody CDR loops as ensembles in solution vs. canonical clusters from X-ray structures. *mAbs* (2020) 12(1):1744328. doi: 10.1080/19420862.2020.1744328
54. Fernández-Quintero ML, Pomarici ND, Loeffler JR, Seidler CA, Liedl KR. T-Cell receptor CDR3 loop conformations in solution shift the relative  $\alpha$ - $\beta$  domain distributions. *Front Immunol* (2020) 11:1440. doi: 10.3389/fimmu.2020.01440
55. Labute P. Protonate3D: Assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins* (2009) 75(1):187–205. doi: 10.1002/prot.22234
56. *Molecular operating environment (MOE); 1010 sherbrooke st. West, suite #910. Montreal, QC, Canada: H3A 2R7* (2020).
57. Case DA, Belfon K, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham TE, et al. *AMBER 2020*. San Francisco: University of California (2020).
58. Hub JS, de Groot BL, Grubmüller H, Groenhof G. Quantifying artifacts in ewald simulations of inhomogeneous systems with a net charge. *J Chem Theory Comput* (2014) 10(1):381–90. doi: 10.1021/ct400626b
59. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput* (2013) 9(7):3084–95. doi: 10.1021/ct400341p
60. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* (1983) 79(2):926–35. doi: 10.1063/1.445869
61. El Hage K, Hédin F, Gupta PK, Meuwly M, Karplus M. Valid molecular dynamics simulations of human hemoglobin require a surprisingly Large box size. *eLife* (2018) 7:e35560. doi: 10.7554/eLife.35560
62. Gapsys V, de Groot BL. Comment on “Valid molecular dynamics simulations of human hemoglobin require a surprisingly Large box size. *bioRxiv* (2019) 563064:1–10. doi: 10.1101/563064
63. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. Ff14SB: Improving the accuracy of protein side chain and backbone parameters from Ff99SB. *J Chem Theory Comput* (2015) 11(8):3696–713. doi: 10.1021/acs.jctc.5b00255
64. Wallnoefer HG, Liedl KR, Fox T. A challenging system: Free energy prediction for factor xa. *J Comput Chem* (2011) 32(8):1743–52. doi: 10.1002/jcc.21758
65. Barducci A, Bussi G, Parrinello M. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys Rev Lett* (2008) 100(2):20603. doi: 10.1103/PhysRevLett.100.020603
66. Illott AJ, Palucha S, Hodgkinson P, Wilson MR. Well-tempered metadynamics as a tool for characterizing multi-component, crystalline molecular machines. *J Phys Chem B* (2013) 117(40):12286–95. doi: 10.1021/jp4045995
67. Barducci A, Bonomi M, Parrinello M. Linking well-tempered metadynamics simulations with experiments. *Biophys J* (2010) 98(9):L44–6. doi: 10.1016/j.bpj.2010.01.033
68. Biswas M, Lickert B, Stock G. Metadynamics enhanced Markov modeling of protein dynamics. *J Phys Chem B* (2018) 122:5508–14. doi: 10.1021/acs.jpcc.7b11800
69. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl, e. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* (2015) 1–2:19–25. doi: 10.1016/j.softx.2015.06.001
70. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* (2013) 29(7):845–54. doi: 10.1093/bioinformatics/btt055
71. Tribello GA, Bonomi M, Branduardi D, Camilloni C, Bussi G. PLUMED 2: New feathers for an old bird. *Comput Phys Commun* (2014) 185(2):604–13. doi: 10.1016/j.cpc.2013.09.018
72. Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. *J Mol Biol* (1963) 7(1):95–9. doi: 10.1016/S0022-2836(63)80023-6
73. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J Appl Phys* (1981) 52(12):7182–90. doi: 10.1063/1.328693
74. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys* (2007) 126(1):014101. doi: 10.1063/1.2408420
75. Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Walker RC. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. explicit solvent particle mesh ewald. *J Chem Theory Comput* (2013) 9(9):3878–88. doi: 10.1021/ct400314y
76. Miyamoto S, Kollman PA. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comput Chem* (1992) 13(8):952–62. doi: 10.1002/jcc.540130805
77. Berendsen H, van Postma JPM, van Gunsteren W, DiNola A, Haak JR. Molecular-dynamics with coupling to an external bath. *J Chem Phys* (1984) 81. doi: 10.1063/1.448118.
78. Doll J, E. Myers L, Adelman S. Generalized langevin equation approach for Atom/Solid-surface scattering: Inelastic studies. *J Chem Phys* (1975) 63. doi: 10.1063/1.431234.
79. Adelman SA, Doll JD. Generalized langevin equation approach for Atom/Solid-surface scattering: General formulation for classical scattering off harmonic solids. *J Chem Phys* (1976) 64(6):2375–88. doi: 10.1063/1.432526
80. Scherer M, Trendelkamp-Schroer B, Paul F, Pérez-Hernández G, Hoffmann M, Plattner N, et al. PyEMMA 2: A software package for estimation, validation, and analysis of Markov models. *J Chem Theory Comput* (2015) 11. doi: 10.1021/acs.jctc.5b00743.
81. Pérez-Hernández G, Noé F. Hierarchical time-lagged independent component analysis: Computing slow modes and reaction coordinates for Large molecular systems. *J Chem Theory Comput* (2016) 12(12):6118–29. doi: 10.1021/acs.jctc.6b00738
82. Chodera JD, Noé F. Markov State models of biomolecular conformational dynamics. *Curr Opin Struct Biol* (2014) 25:135–44. doi: 10.1016/j.sbi.2014.04.002
83. Bowman GR, Pande V, Noé F. *An introduction to Markov state models and their application to long timescale molecular simulation*. Springer (2014). doi: 10.1007/978-94-007-7606-7
84. Likas A, Vlassis N, J. Verbeek J. The global K-means clustering algorithm. *Pattern Recognition* (2003) 36(2):451–61. doi: 10.1016/S0031-3203(02)00060-2
85. Röblitz S, Weber M. Fuzzy spectral clustering by PCCA+: Application to Markov state models and data classification. *Adv Data Anal Classification* (2013) 7(2):147–79. doi: 10.1007/s11634-013-0134-6
86. Miroshin RN. Special solutions of the Chapman–kolmogorov equation for multidimensional-state Markov processes with continuous time. *Mathematics* (2016) 49:122–9. doi: 10.3103/S1063454116020114
87. Karush J. On the Chapman-kolmogorov equation. *Ann Math Statist* (1961) 32(4):1333–7. doi: 10.1214/aoms/1177704871
88. Wu H, Noé F. Variational approach for learning Markov processes from time series data. *J Nonlinear Sci* (2017) 30:23–66. doi: 10.1007/s00332-019-09567-y
89. Schrodinger. *The AxPyMOL molecular graphics plugin for Microsoft PowerPoint, version 1.8*. (2015).
90. Stanford University (adate). GetContacts. Available at: <https://getcontacts.github.io/>.