RESEARCH ARTICLE

PROTEINS WILEY

Functional antibody characterization via direct structural analysis and information-driven protein–protein docking

Rafael S. Depetris[1] ⓘ | Dan Lu[1] | Zhanna Polonskaya[1] | Zhikai Zhang[1] |
Xenia Luna[1] | Amari Tankard[1] ⓘ | Pegah Kolahi[1] | Michael Drummond[2] ⓘ |
Chris Williams[2] ⓘ | Maximilian C. C. J. C. Ebert[2] | Jeegar P. Patel[1] |
Masha V. Poyurovsky[1] ⓘ

[1]Kadmon Corporation, LLC, New York, New York, USA

[2]Chemical Computing Group, ULC, Montreal, Quebec, Canada

Correspondence
Masha V. Poyurovsky, 450 E. 29th St., New York, NY 10016, USA.
Email: masha@kadmon.com

Present address
Amari Tankard, Department of Molecular Biology, Princeton University, Princeton, New Jersey, USA

Abstract

Detailed description of the mechanism of action of the therapeutic antibodies is essential for the functional characterization and future optimization of potential clinical agents. We recently developed KD035, a fully human antibody targeting vascular endothelial growth factor receptor 2 (VEGFR2). KD035 blocked VEGF-A, and VEGF-C-mediated VEGFR2 activation, as demonstrated by the in vitro binding and competition assays and functional cellular assays. Here, we report a computational model of the complex between the variable fragment of KD035 (KD035(Fv)) and the domains 2 and 3 of the extracellular portion of VEGFR2 (VEGFR2(D2-3)). Our modeling was guided by a priori experimental information including the X-ray structures of KD035 and related antibodies, binding assays, target domain mapping and comparison of KD035 affinity for VEGFR2 from different species. The accuracy of the model was assessed by molecular dynamics simulations, and subsequently validated by mutagenesis and binding analysis. Importantly, the steps followed during the generation of this model can set a precedent for future in silico efforts aimed at the accurate description of the antibody–antigen and more broadly protein–protein complexes.

KEYWORDS
antibody–antigen modeling, antibody X-ray structure function, molecular dynamics, model validation studies, protein complex modeling validation, VEGFR2 clinical antibody, VEGF signaling

1 | INTRODUCTION

Monoclonal antibodies (MAbs) development has led to many effective and successful therapeutics for the treatment of a wide spectrum of diseases, from oncology to genetic disorders. Antibodies are by far the most frequent form of protein therapies entering the clinic due in part to their high affinity, target specificity, and favorable safety profiles.[1] Structurally, antibodies are multi-domain proteins composed of beta-sheet containing immunoglobulin folds. The variable heavy chain (V_H) and the variable light chain (V_L) pair together forming the variable fragment (F_v). Six hypervariable loops, known as complementarity-determining regions (CDRs) are located in the V_H and V_L domains, are responsible for the interaction with the antigen. The binding site or the paratope is formed by solvent exposed amino

acids located on the CDRs from V_H and V_L which interact with amino acids of the antigen-binding site or the epitope.[2]

A precise description of an antibody–antigen complex can drive multiple aspects of antibody optimization, which is paramount for successful development of a therapeutic agent. Structural analysis of antibody–antigen complexes has proven to be highly useful in numerous aspects of antibody development, including structure-based affinity maturation campaigns aimed at the improvement of potency and specificity.[3,4] The information about MAbs epitopes taken from the analysis of these complexes could differentiate between agents targeting the same ligand and drive patentability and regulatory interactions.[5] Additionally, the precise structural characterization of the CDRs, principally responsible for the antibody affinity and specificity, enables a focused approach in the functional optimization and developability improvement of these therapeutic agents.[6]

X-Ray crystallography, nuclear magnetic resonance, and cryo-electron microscopy are some of the experimental methods used to ascertain detailed structural descriptions of antibody–antigen complexes.[7–9] Although highly accurate, these methods are also laborious and resource and time consuming. Computational methods developed to generate models of protein–protein complexes are becoming the faster, more accessible and less resource and training intensive alternatives to the classical tools of structural biology. Still, the accuracy of these models relies heavily on the robustness and availability of the structural information used as input, which is generated via experimental methods or by homology modeling.[10–12] In the specific case of antibody–antigen complexes, the vast pre-existing knowledge of MAbs structural features,[2] enables a reliable prediction of the amino acids involved in the interaction with the target, particularly if supported by data from antibody maturation efforts. Conversely, reliable epitope prediction remains a challenging computational exercise and an open issue in the field, making improvements of the existing modeling solutions a focus of much of the ongoing discovery efforts.[13–21]

Many of the currently available engines for the generation of models of protein–protein complexes are based on the docking of unbound structures. The output of protein–protein docking processes produces a database of numerous poses, reflecting many possible binding modes, ranked using a scoring value. Multiple approaches can be employed in the generation of these scores. One widely used strategy is the classification the poses by a scalar value related to the change in free energy derived from the formation of the complex.[22,23] Even with the application of the most sophisticated and computationally intensive scoring methods, accurate prediction of a native structure of the complex remains challenging, largely due to the vast landscape of possible contacts made by two interacting proteins.[24] Information-driven docking methods allow for the sampling of possible binding modes guided by the introduction of restraints emanating from previously available experimental data that may be used to set up modeling restraints.[25] This restraining information may result from the experimental identification of residues previously mapped to the protein–protein interface by biochemical and functional analysis,[26] the analysis of shape complementarity and electrostatics[27] or the

definition of binding site by NMR titration and mutagenesis.[28,29] Specifically, for the in silico generation of antibody–antigen complexes, it has been reported that docking algorithms that incorporate information about the antibody CDRs and even a coarse definition of the epitope are able to generate improved predictions.[28] The performance of information driven protein–protein docking strategies is evaluated continuously by the Critical Assessment of Predicted Interactions (CAPRI) initiative which is a valuable resource for the comparison of different docking algorithms and scoring methods.[30,31] Additional tools that evaluate the quality of docking models based on a single measure are a good set up for the application of machine learning methods on this assessment.[32,33]

The mammalian vascular endothelial growth factors (VEGF-A, VEGF-B, VEGF-C, VEGF-D, VEGF-E, and PIGF) and their receptors, VEGFR1, VEGFR2, and VEGFR3, are key regulators of the process of angiogenesis as well as vascular permeability and vessel dilation.[34,35] VEGF-A and VEGF-C signaling through VEGFR2 constitutes the major angiogenic signaling pathway regulating endothelial cell sprouting, migration, proliferation and survival.[36–38] The mechanism of activation of VEGFRs by VEGFs is well characterized. VEGFs are disulfide-linked homodimers which bridge two receptor molecules through two identical binding sites at opposite ends. The formation of a VEGF:VEGFR 2:2 homotetramer results in the trans-activation of the intracellular tyrosine kinase domains resulting in the activation of the downstream intracellular signaling cascades.[38,39] Available structural information on the full length VEGFR2 ectodomain (Domains 1 through 7, D1–D7) and binding analyses performed with the truncated extracellular VEGFR-1 and -2 constructs show that the VEGF-A, VEGF-C, and VEGF-E homodimers establish loop mediated interactions with the C-terminal region of D2 and the N-terminal region of D3 of the receptors. Further, the linker between D2 and D3 is positioned in close proximity to the groove formed by the VEGF interface, resulting in an additional set of contacts.[40–46] The hydrophobic nature of these interactions allows the binding with the different non-identical VEGFs. Thus, binding of the dimeric VEGFs to the VEGFR2(D2-3) extracellular segment positions the two receptor molecules at an optimal distance for the transactivation of their intracellular kinase domains and subsequent signaling cascade activation.[47–49]

Due to the therapeutic potential of the blockade of VEGF binding and VEGFR activation in oncology and other indications, several antibodies targeting VEGFR2 ectodomain have been generated.[50,51] In fact, ramucirumab (Imclone/Eli Lilly), a direct VEGFR2 antagonist antibody, is currently approved as a treatment for a variety of solid tumors. Consistent with its VEGF blocking activity, epitope mapping based on the in vitro binding assays shows that ramucirumab interacts preferentially with VEGFR2 D3; however, crystal structure of the complex shows that additional interactions with the D2 cannot be discarded.[50] To expand on the existing therapeutics, we identified and characterized another VEGFR2 antagonist antibody, KD035. KD035 binds with high affinity to VEGFR2 blocking receptor activation by the native ligands, VEGF-A and VEGF-C. The absence of the structural information describing the KD035-VEGFR2(D2–3), we developed an information-driven complex modeling strategy to characterize the

antibody-receptor interface as a part of our development pathway. Here we report a model for the KD035-VEGFR2(D2–3) complex generated using docking methods pre-informed by the experimental information from the affinity maturation process and epitope mapping. The accuracy of our model was extensively confirmed by mutagenesis and binding analysis. The model validation by site-directed mutagenesis and its evaluation by MD simulation allowed us to elucidate structural basis of the inhibition of VEGFR2 activity by KD035. Consequently, we were able to make valuable inferences about KD035's mechanism of action, useful in functional data interpretation and further affinity maturation campaigns. Additionally, the modeling paradigm set up in this work may add to the future efforts aimed at improving accuracy of computationally generated models of protein–protein complexes.

2 | MATERIALS AND METHODS

2.1 | Generation, maturation, expression, and purification of VEGFR2 antibodies

KD035 was identified by panning of Dyax FAB-310 phagemid library on immobilized human VEGFR2 (R&D systems) followed by affinity maturation using a light chain shuffling method. The constructs carrying KD035 antigen-binding fragment (Fab) domain were cloned into the mammalian expression vector pBh1 (Dyax) and transiently expressed in human 293 Expi cells (Invitrogen). Fabs were purified from cell culture supernatant by passing it several times through a protein A Sepharose HP column (GE Healthcare) using a peristaltic pump for a period of 8 h and eluting with 1 M Glycine pH 2.0. The eluate was further purified by size exclusion chromatography (SEC) using a Superdex 75 increase (GE Healthcare) column.

The variable domain amino acid sequence of IMC-1121B (which shares the amino acid sequence of ramucirumab as described in US 8057791 B2/EP 1916001 A2) was reverse-translated into DNA sequences and optimized by using Integrated DNA Technologies tools. The synthesized genes of 1121 light chain and heavy chain variable domains were cloned into a mammalian expression vector pBh1 (Dyax). The plasmid DNA containing 1121 gene was transfected and expressed by using the FreeStyle™ MAX 293 Expression System (ThermoFisher Scientific) according to the manufacture instruction. The supernatant was harvested and purified by protein A (GE) with at least 95% monomer detected by SEC-UPLC (Waters).

2.2 | VEGFR2 protein production

All the constructs including different lengths of the extracellular domain of human VEGFR2 carrying an N-terminal His-tag was expressed using the Bac-to-Bac baculovirus expression system in *Spodoptera frugiperda* Hi5 cells (ThermoFisher). These proteins were purified from conditioned media after 48 h of infection using a Nickel Sepharose HP column (GE Healthcare) followed by dialysis in PBS and a final step of purification by SEC using Superdex 200 Increase

column (GE healthcare). The integrity and aggregation of the WT and mutated or truncated proteins were evaluated by inspection of the SEC elution peak and by SDS PAGE.

2.3 | Binding ELISAs

Binding ELISAs for domain mapping studies were performed as previously described.[52,53] Binding ELISAs for mutagenesis analysis and model validation were performed by coating the surface of Immulon 2 HB plates (ThermoFisher) overnight at 4°C with His-tagged VEGFR2(D2–3) which were either wild-type, or the mutants L313R, D257A/E261A, and M197A/M211A/I213A. The plates were then washed with PBS supplemented with 0.1% Tween-20 and blocked in a solution containing PBS supplemented with 3% milk. Separately, serial dilutions of VEGF-A or KD035 Fab were prepared in the blocking solution and then incubated for 20 min at RT prior to the addition to the plate. The Fab mixture was then added to the plate and incubated for 1 h at RT, and subsequently washed and incubated with secondary antibody (anti-Fab HRP-conjugated goat polyclonal [R&D systems]). The ELISAs were developed using TMB substrate (ThermoFisher) and the reaction was stopped with 0.2 M H_2SO_4. The IC_{50} values were determined using the single site model in Prism (GraphPad).

2.4 | Surface plasmon resonance (SPR) assay

SPR experiments described in this study were performed on a BIAcore T200 instrument (GE Healthcare). KD035 (5 μg/ml) was coupled on a CM5 sensor chip (GE Healthcare) using amine coupling method under standard conditions at a flow rate of 10 μl/min. Briefly, KD035 (5 μg/ml) was diluted in Acetate buffer pH 5.0 and injected over sensor surfaces activated by a mixture of 1-Ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC)/N-hydroxysuccinimide (NHS; GE Healthcare). Prior to the immobilization, VEGFR2(D2–3) WT and the mutants L313R, D257A/E261A, and M197A/M211A/I213A were diluted in the HBS-EB running buffer to a concentration of 600 nM and further diluted in two-fold steps to a minimal concentration of 4.7 nM. The concentration series of wild-type and mutant VEGFR2(D2–3) constructs were injected for 200 s with a 600 s dissociation phase at a flow rate of 10 μl/min. Between injections, the chip was regenerated using 10 mM Glycine/HCl. Data analysis was performed using the BIAevaluation software (Biacore GE Healthcare). The SPR sensorgrams were evaluated by fitting the curves to the 1:1 binding kinetics model and the affinity of KD035 to the different VEGFR2 (D2–3) constructs were calculated.

2.5 | Crystallization

The Fab regions of B1 and KD035 were concentrated to 10 mg/ml and crystallization trays were set up using the Hampton PEG-ION screen HT (Hampton Research). Of note, 200 nl hanging drops for

each of the 96 conditions were set up mixing the protein in a 1:1 ratio with the crystallization condition using a Mosquito robot (TTP Labtech), incubated at 17°C and inspected using a Rock Imager 1000 system (Formulatrix). Both B1 and KD035 crystals were grown in a condition containing 0.2 M Na phosphate monobasic monohydrate, 20% PEG3350 (Hampton PEG-Ion screen #D3). Crystals were harvested from the tray and frozen directly in liquid nitrogen for data collection in a cryoprotectant solution containing the crystallization condition supplemented with 20% ethylene glycol.

2.6 | Data processing

All datasets were collected at the Argonne National Light Source (Argonne, IL) beamline 19-BM. Both B1 and KD035 data sets were indexed and scaled using HKL3000 (HKL Research Inc.).[54] The structures were solved using Phaser-MR from the Phenix suite[55] using the crystal structure of the Fab fragment of the human neutralizing anti-West Nile Virus Mab (PDB ID 3N9G) as the search model.[56] Refinement was done using Phenix.refine.[55] Both KD035 and B1 data sets were subjected to three cycles of rigid body refinement after finding the molecular replacement solution and subsequently to more than 10 rounds of three cycles of reciprocal and real space and B-factors and occupancies refinement. A subset of the collected dataset with a high-resolution limit of 2.40 was used in the refinement of KD035, evidenced by the lower number of unique reflections used in this step. The completeness and redundancy values for the highest resolution shell from the trimmed data are 98.56% and 3.1, and the R_{meas}, CC1/2 and CC* values are 0.486, 0.782 and 0.937. Model building was done using Coot.[57] Water molecules were placed using Coot and updated by Phenix.Refine during the refinement rounds. Data collection and refinement statistics are compiled in Table 1.

2.7 | Antibody–antigen modeling

The model for the KD035(Fv) in complex with VEGFR2(D2–3) was built using Molecular Operating Environment (MOE) (Chemical Computing Group).[58] The coordinates for KD035 Fv were taken from the crystal structure of KD035 Fab solved at Kadmon as described above. The coordinates for the VEGFR2(D2–3) domains were taken from the crystal structure of these domains in complex with VEFG-C (PDB code 2X1W—amino acids 134–319[46]).The input coordinates were processed using the protein preparation function implemented in MOE, which consists in adding hydrogen atoms, assessing partial charges using the OPLS-AA molecular mechanics force field and assigning protonation states. The complex model was obtained using the protein–protein dock tool in MOE. The docking algorithm in MOE[58] starts with a coarse-grained model which reduces the search space complexity by representing the all atom protein structures as bead models. The Hopf fibration is used to generate a set of uniformly

TABLE 1 Data collection and refinement statistics

	KD035 Fab	B1 Fab
Data collection		
Space group	P12₁1	P12₁1
Unit cell parameters		
a, b, c (Å)	53.16, 90.23, 81.52	55.69, 66.41, 66.54
α, β, γ (°)	90.0, 93.8, 90.0	90.0, 101.93, 90,0
Site of data collection	ANLS (19ID)	ANLS (19ID)
Measured reflections	166 839 (9252)[a]	59 942 (4560)[a]
Unique reflections	50 743 (4626)	16 360 (1520)
Resolution (Å)	39.46–1.99 (2.06–1.99)	42.12–2.49 (2.58–2.49)
Completeness (%)	96.50 (88.57)	97.65 (91.57)
Average redundancy	3.3 (2.0)	3.7 (3.0)
I/σ(I)	6.71 (0.76)	23.61 (7.13)
Wilson B-Factor	33.17	30.03
R_{meas}[b]	0.132 (1.05)	0.086 (0.30)
CC1/2	(0.31)	(0.90)
CC*	(0.69)	(0.97)
Refinement		
Resolution (Å)	39.46–2.40[c]	42.12–2.49
Number of atoms		
Protein	6276	3137
Solvent	139	36
Reflections used in refinement	29 562 (2945)[d]	16 358 (1520)[d]
Reflections used for R-free	1155 (119)	1636 (152)
R-work/R-free	0.20/0.28	0.19/0.25
R.M.S.D. values		
Bond lengths (Å)	0.011	0.010
Bond angles (°)	1.54	1.62
Average B factors (Å²)		
All atoms	33.93	30.56
Protein	33.92	30.53
Solvent	34.08	33.55
Ramachandran favored/ outliers (%)	94.22/0.72	92.14/2.14
Rotamer outliers	4.03	4.06
Model contents		
Protein residues	842	424
Waters	49	12
PDB accession code	7LSI	7LSB

[a]Values in parentheses are for the highest resolution shell.
[b]R_{meas} is used as a redundancy independent measurement of internal data quality.
[c]Data collection parameters for the subset of data used in refinement of KD035 are detailed in Section 2.
[d]Values in parentheses are for the highest resolution shell of the data subset used for refinement. One crystal was used for each data set.

distributed rotations[59] of the bead models and exhaustive sampling of all translations for a given rotation is performed using a Fast Fourier Transform. This generates a set of initial docked poses of the protein bead models, which are then filtered based on the placement restraints and subjected to coarse refinement. The refined bead model poses are then converted back to all atom models which are further refined with side-chain packing and molecular mechanics minimization to produce a list of all atom protein–protein docked poses. The docking process included the specification of particular residues which were candidates to be part of the antibody–antigen interface according to our alignment-based epitope mapping and the comparison of B1 and KD035 structures (see Section 3). The output database consisted in 100 poses which were generated by triangle match placement method with the side chain position refined by induced fit. The poses were scored by the estimation of the free energy of the complex (London ΔG) and rescored using the GVI/WSA function which adds an empirical term to account for the desolvation effect.[23] The protein preparation process mentioned above corrected most of the Ramachandran outliers. All the residues selected for the docking site restraint as well as those involved in the interactions were in Ramachandran favored or allowed regions.

2.8 | Molecular dynamics simulations

MD simulations were performed using NAMD.[60] Explicit solvation was performed using MOE after removal of crystallographic waters. A total of 38 561 water molecules were added in a box with periodic boundary conditions at a density of 1.032 g/cm³ and keeping a distance of 10 Å between the boundaries and the protein. The net charge of the protein was neutralized with 150 mM NaCl. For energy minimization and MD simulations, the AMBER14:EHT force field was used and the electrostatic interactions were evaluated by the particle-mesh Ewald method. The system was energy minimized for 5000 steps using the Deepest Descent and Conjugate Gradient methods. For equilibration, the system was subjected to a 100 ps simulation to gradually heating from 10 to 300 K. Next, a 100 ps NVT ensemble was generated at 300 K followed by and NPT ensemble for 200 ps at 300 K and 1 bar. Then for the complex, a 50 ns production trajectory was generated for further analysis. The database file containing the trajectory was analyzed using MOE protein–ligand interaction fingerprint (PLIF) analysis.

2.9 | VEGF-A and VEGF-C competition assays

Serially diluted aliquots of KD035 or ramucirumab were mixed with fixed concentration of VEGFR2 (final concentration was 0.5 μg/ml) and incubated in RT for about 1 h. The mixture was transferred to either VEGF-A or VEGF-C (R&D system) coated ELISA plate and incubated at room temperature for 1 h. The signal was detected by HRP-anti-human IgG Fc specific antibody (Jackson Immunoresearch).

3 | RESULTS

3.1 | Structure-enabled paratope mapping

Having generated a potent VEGFR2 targeting antibody (KD035), we wanted to further characterize the mechanism of action and the molecular drivers of binding specificity of this molecule. In order to identify key residues in KD035 responsible for ligand binding, we initiated our analysis by solving X-ray crystal structures of KD035 Fab and a B1 Fab (Figure 1A). B1 is a KD035 parental antibody that was isolated from the phage display library screen prior to the affinity maturation process. KD035 was determined to have 8.2-fold greater affinity for VEGFR2(D2–3) than B1, K_d of 0.22 nM versus 1.81 nM, respectively, as measured by SPR (Table 2). The superposition of the structures of KD035 and B1 Fv regions gave a Cα RMSD of 0.60 Å with localized structural differences in the CDRs of the light chain, which harbored the mutations generated during the B1-KD035 affinity maturation process (Figure 1B). We hypothesized that the comparison analysis of the resulting structures could help us determine the affinity maturation-generated changes in the CDRs that led to the affinity improvement in KD035. We subsequently anticipated that the identified side chains will serve as restraints in the modeling of the KD035-VEGFR2(D2–3) complex. Comparison of the protein surface maps for both Fvs identified one positively charged patch and two negatively charged patches in KD035 which were not present in B1 (Figure 1C). The inspection of the side chains mapping within these patches, highlighted the presence of three potentially relevant mutations. The large positively charged patch in KD035 resulted from the N52K mutation, while the negatively charged patches were associated with the D25E and S29D maturation changes. These three side chains were largely solvent exposed as shown on the KD035 Fv structure (Figure 1D) and when shown in a structural alignment of KD035 CDRs with their B1 counterparts (Figure 1E). The maturation related mutations S29D and N52K can easily be related with an introduction of a negative and positive charged patch on the antibody surface, respectively. The introduction of a longer side chain by the D25E mutation resulted in a negatively charged patch as a consequence of the greater solvent exposure of the carboxylate group which can also be attributed to the conformational difference of CDR L1 between the two antibodies (Figure 1E). The impact of these mutations on the antibody surface charge distribution and their correlation with and 8.2-fold improvement in affinity led us to postulate that least some of these residues constituted a part of the binding interface with the receptor. We therefore decided to include E25 (CDR L1) and D29 (CDR L1) and K52 (CDR L2) as restraints for the definition of the KD035 paratope during the antibody–antigen docking process.

3.2 | VEGFR2 epitope mapping

To improve our ability to model KD035(Fv)-VEGFR2(D2–3) interaction, we needed to map the binding area of KD035 on VEGFR2 and determine minimal receptor domain requirements for the interaction.
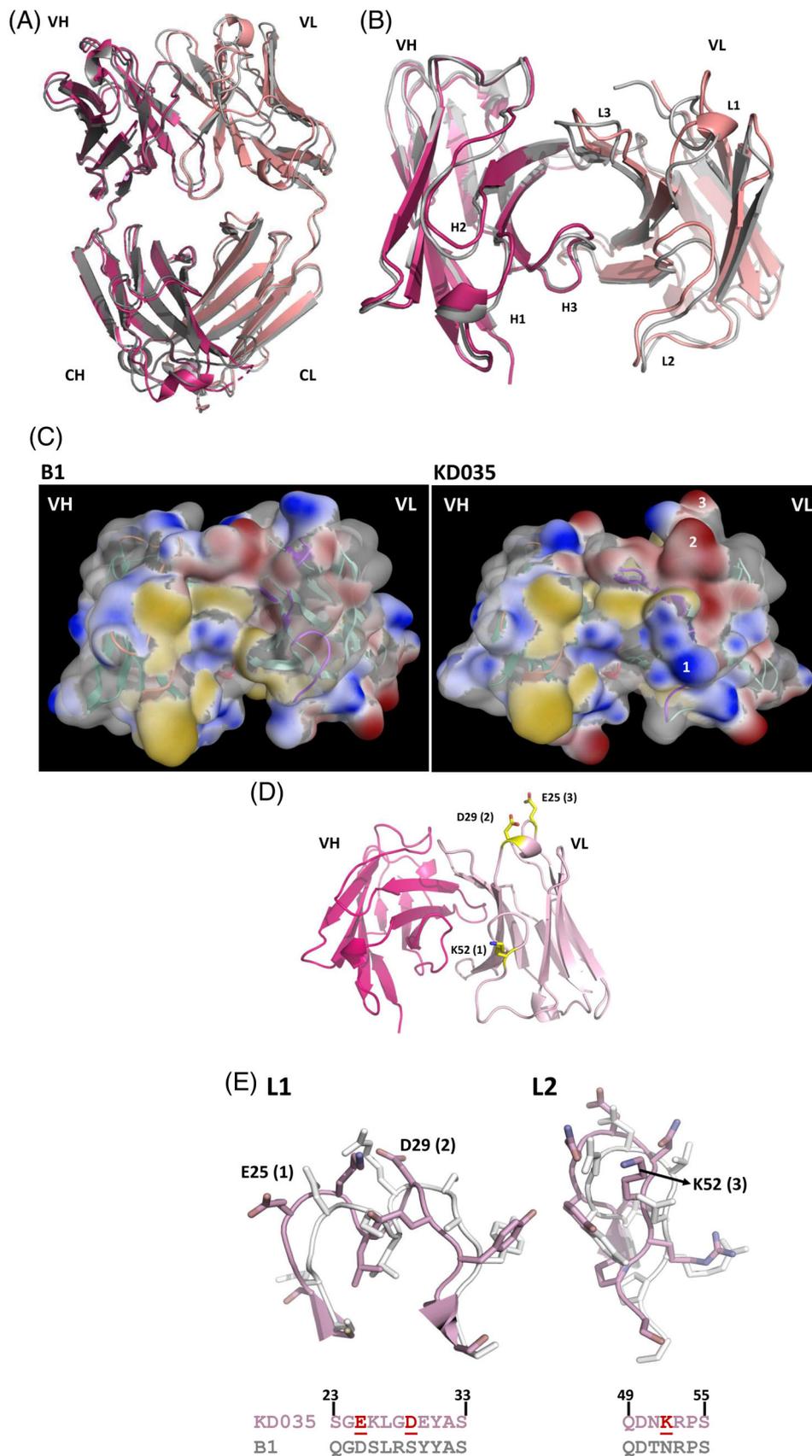
FIGURE 1 Structure of KD035 (Fab) and B1(Fab). (A) Superposition of the KD035(Fab) and B1 (Fab) structures. KD035 is shown in cartoon rendering in pink (heavy chain) and salmon (light chain) and B1 is shown in gray. CH, constant heavy; CL, constant light; VH, variable heavy; VL, variable light. (B) Superposition of the variable domains of KD035 and B1 facing the paratope region. H1-3 and L1-3 correspond to the CDRs of the heavy and light chain, respectively. (C) Surface patch analysis of B1 and KD035 Fvs, facing the paratope region. Negatively charged patches are shown in red, positive charged patches are shown in blue and hydrophobic patches are shown in yellow. VH, variable heavy; VL, variable light. The patches in KD035 VL are numbered to link them to the side chains carrying the particular charge shown in the structure. (D) Structure of KD035 (Fv) (paratope view). The residues shown in stick correspond to the surface patches numbered in (C), which were used as restrain during the antibody–antigen model generation. (E) Details of the structures of the CDRs L1 and L2of KD035 and B1. The loops for KD035 (pink) and B1 (white) are shown as cartoon representations and the loop side chains are shown as sticks. An alignment of the sequences of KD035 and B1 CDRs 1 and 2 is shown below each structure. The side chains underlined were chosen for binding site restraint during the modeling process and are also shown in the corresponding structures. The number between parentheses corresponds to the number of the electrostatic patches shown in (C)

To this end, we generated deletion constructs spanning various domains of the human VEGFR2 extracellular region (Figure 2A). Using full length VEGFR2(D1–7) extracellular region as reference, we compared the affinity of KD035 to various domain deletions of VEGFR2 by ELISA (Figure 2B). This analysis demonstrated that KD035 was able to bind to the first three N-terminal domains

VEGFR2(D1–3) as strongly as to the D1–7 construct, while a construct spanning Domains 2 and 3 only VEGFR2(D2–3) retained around 80% of the full length ectodomain interaction in this

experiment, it showed no loss of binding in subsequent assays. We also determined that both Domains 2 and 3 were required for the interaction, given that the deletion of either of these domains resulted in the loss of KD035 binding. Based on these data, we concluded that the KD035 epitope was contained principally within VEGFR2(D2–3). Additionally, based on the absence of binding to the mouse VEGFR2 (D1–7), we determined that KD035 does not recognize the mouse orthologue (Figure 2B).

To further characterize the KD035 epitope and to guide the identification of the specific residues that make direct contacts with the antibody, we took advantage of the homology between the human
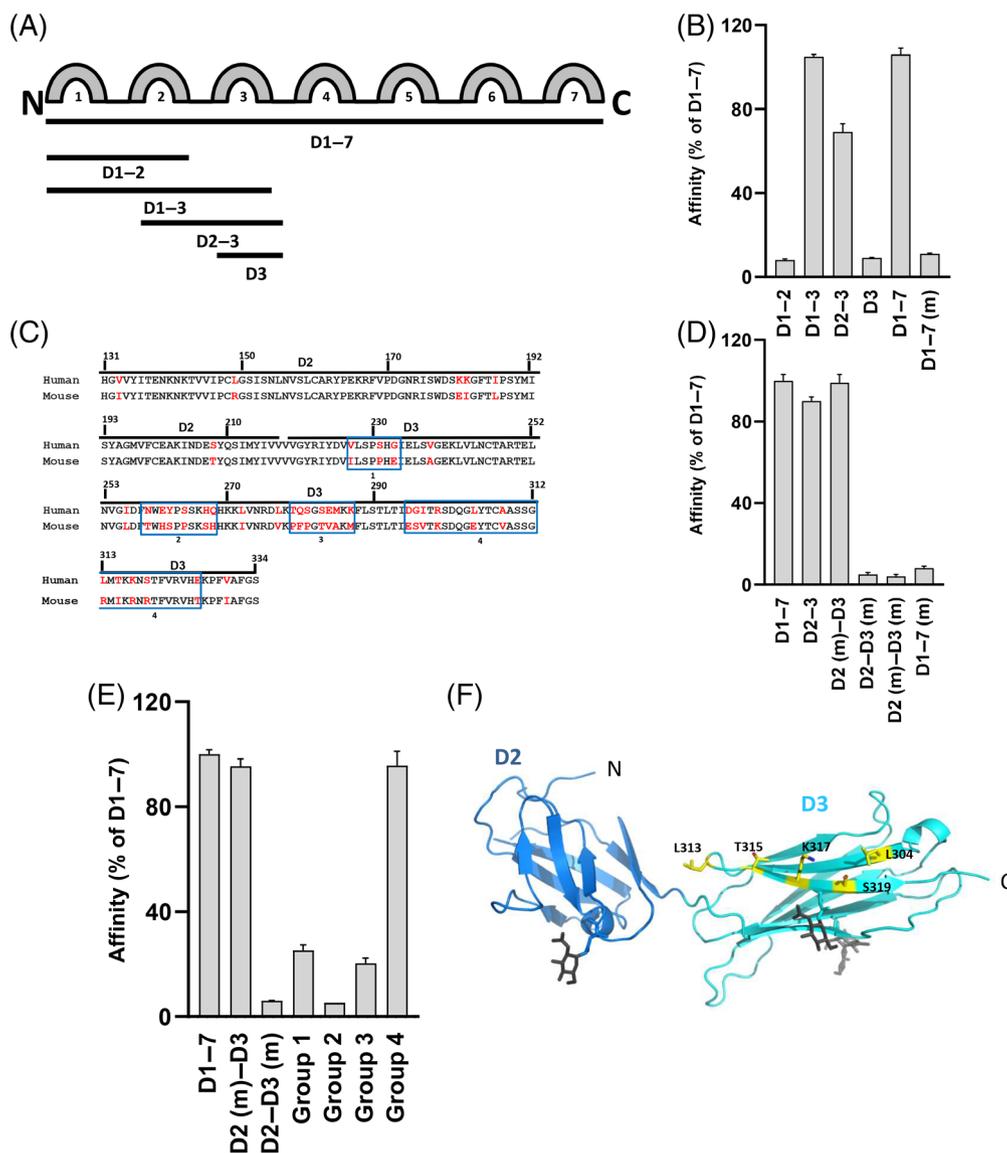
TABLE 2 Affinity values for VEGFR2(D2–3) to KD035(Fab) and B1(Fab) measured by SPR(Biacore)

MAb	$K_a \times 10^5$ (M^{-1} s^{-1})	$K_d \times 10^{-5}$ (s^{-1})	KD (nM)
KD035	2.67	5.78	0.22
B1	2.98	53.9	1.81

FIGURE 2 Mapping of the binding site of KD035 on VEGFR2 ectodomain. (A) Schematic of the organization of the extracellular domains 1–7 showing the different VEGFR extracellular fragments analyzed. (B) Binding analysis of KD035 to VEGFR2 to the different constructs schematized in (A). All the constructs refer to human VEGFR2 with the exception of D1–7(m) which includes the mouse sequence. (C) Alignment of the sequences of human and mouse VEGFR(D2–D3). Non-identical residues are shown in red. The four groups of human-to-mouse differences are shown boxed and numbered as 1 through 4. (D, E) Binding analysis of KD035 to VEGFR(D2–D3) human/mouse chimera and to mouse-to-human group back mutations. Bars represent the mean ± SD of duplicate or triplicate measurements plotted as percentage of the D1–7 binding. (F) Structure of the VEGFR(D2–D3) ectodomains (PDB:2X1W[46]). D2 is shown in blue and D3 in cyan. Solvent exposed residues from Group 4 mouse-to-human back mutations are shown in sticks and colored by element, carbohydrate decorations are shown as black sticks

and mouse orthologues of VEGFR2. The alignment of human and mouse sequences of VEGFR2(D2–3) showed that most of the differences are located in D3. More precisely, identity values were 92.5% for D2 and 70.8% for D3 (Figure 2C). This allowed us to analyze the differences in the D2–D3 construct that could account for the lack of mouse cross reactivity. To this end we tested the binding of KD035 to the chimeric VEGFR2(D2–3) constructs where one of the two human domains was replaced by the corresponding mouse sequence. Consistent with the extent of identity between the mouse and human proteins, the binding analysis of the human D2–mouse D3 and mouse D2–human D3 chimeras showed that the interaction with KD035 was retained when D2 was switched to the mouse homologue but was mostly abrogated when the mouse D3 as included (Figure 2D). To extend these data we generated an additional series of human/mouse chimeras, where systematic replacements of subsets of human-mouse non-identical amino acids (labeled as Groups 1–4) were made transferring human sequences into the mouse D3 (Figure 2C). We then analyzed the interaction of KD035 with four VEGFR2(D2–3) constructs that included a human D2 followed by a chimeric mouse D3 in which the human amino acids were introduced one group at a time. Interestingly, KD035 binding to the human D2–mouse D3 construct was completely restored only when the amino acids in Group 4 were mutated back to their human counterparts (Figure 2C,E). We therefore concluded that some of the positions in this group are responsible for the lack of KD035 interaction with the mouse VEGFR2 (as shown above in Figure 2B) and could therefore be involved in forming the KD035(Fv)/VEGFR2(D2–3) complex interface. The final set of VEGFR2 amino acids selected for setting the modeling restraint was determined after inspecting the structure of VEGFR2(D2–3), which is available in complex with VEGF-C (PDB:2X1W[46]). We included only human/mouse non-homologous amino acids from Group 4 with solvent exposed side chains in the structure (L304, L313, T315, and S319; Figure 2F).

3.3 | Modeling the KD035 Fab—VEGFR2(D2–3) interaction

Having identified the sites on the antibody and the receptor that could be part of the complex interface (Figures 1D and 2F), we set up a protein–protein docking process using the Fv part of our crystal structure of KD035 Fab and the previously reported structure of the VEGFR2(D2–3) (PDB code 2X1W[46]). These inputs were processed in advance to the docking experiment to correct for missing side chains and gaps generated during the structure building process, and to set up side chain charges at pH 7.0. To increase the accuracy of the antibody–antigen complex generation, we incorporated a placement restraint by designating the above described amino acids as being a part of the binding site for KD035(Fv) and VEGFR2(D2–3). We also included a refinement step by performing induced fit of the side chains involved in the protein–protein interface and neighboring regions. Finally the poses obtained were processed in two steps by

the calculation of free energy of binding by London dispersion (London ΔG) and subsequently by the use of implicit solvent methods (GBVI/WSA).[23] The latest step resulted in a docking score (S) which was the final scalar value used to rank the binding poses. Additional information about the docking process can be found in the methods section.

Given that the docking score is an approximation of the change in the free energy of binding, it might be possible to assume that the pose with the largest negative score is the one that most closely resembles the native structure. However, this approach can be misleading when the highest ranked poses have very similar score values. In our case, the plot of the docking score values for the first five poses (Figure 3A) showed a gap between the score of the highest ranked pose and the rest. Whereas the first and the second pose differed in 1 kCal/mol, the S value of the subsequent poses differed in 0.5 kCal/mol or less. This led us to pick the first pose for closer review and validation. The inspection of the structural features of this pose shows that KD035 CDRs in our model were positioned near the C-terminus of D2 and the N-terminus of D3, close to the linker between the D2 and D3 domains of VEGFR2, positioning the KD035-binding site in close proximity to the VEGFs binding site[44,45] (Figure 3B).

3.4 | Analysis of the KD035 Fab—VEGFR2(D2–3) model

The detailed inspection of the protein–protein interface showed that the VEGFR2 residue L313 (D3) was contributing to a network of hydrophobic contacts mediated by the aromatic rings of Y137 (D2), Y221 (D3), Y31 (CDR L1), W90 (CDR L3), and by the side chain of V219 (D3) (Figure 4A). As mentioned before, L313 belongs to the Group 4 of mouse/human non-homologous amino acids in VEGFR2 (D3). In fact, the mouse VEGFR2(D3) sequence contains an arginine at this position (Figure 2C), which could be detrimental for KD035 binding and explain the lack of interaction with the mouse D3 we observed in our ELISA experiments.

Other interactions included the side chains of E25 (CDR L1) and D29 (CDR L1), which were set up as restraints defining KD035 paratope (Figure 2) and mediated contacts with the VEGFR2(D2) in the model. E25 was located in close proximity to K183 (D2) allowing for salt bridge formation. D29, by means of its interaction with K26 (CDR L1), established an H-bond with the backbone nitrogen of V147 (D2) (Figure 4B). An additional interaction in this area involves the backbone carbonyl of G28 (CDR L1) which accepts and H-bond from the side chain of T145 (D2) (Figure 4B). The third paratope restraint, K52 (CDR L2), which introduced a positively charged patch on KD035 (Figure 2) was not involved in a polar interaction, but its alkyl group and the ones of K316 and M314 (D3) were close enough to make hydrophobic contact (Figure 4C). K52 preceding residue, N51, was positioned in close proximity to E140 (D2) and K142 (D2) and with R222 (D3) to mediate a robust net of H-bonds (Figure 4C). Interestingly, these interactions provided a potential explanation for the
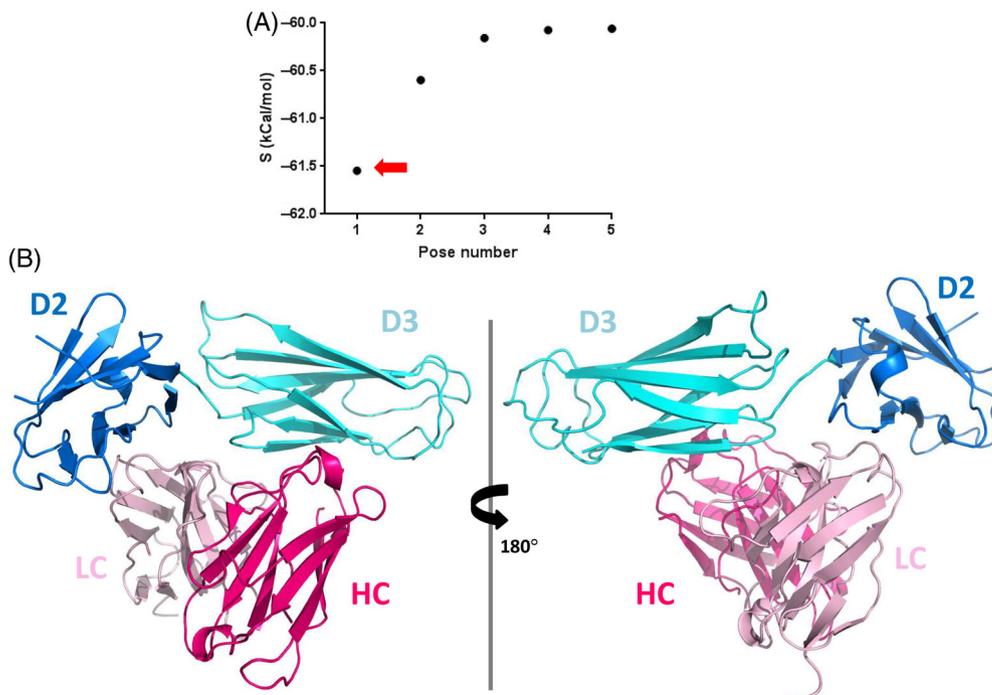
8.2-fold improvement in VEGFR2 affinity between KD035 and its less potent parental counterpart B1 (Table 2). Shorter side chains are present in B1 at all three positions (D25, S29, and T51), given the importance of these residues in the complex formation, these changes could directly contribute to the improved binding (Figure 1E). Our initial analysis focused on the changes in hydrophobic or electrostatic features, thus KD035 N51 was not included as a docking restraint. However, our model shows that the introduction of the T51N mutation during the B1 to KD035 maturation process should contribute to the observed increased affinity of KD035. To a minor extent, the interaction involving K52 and the stabilization of D29 by K26 would be negatively affected by the less suitable B1 residues N52 and S26 (Figure 1E).

Our model further suggested that residues in KD035 CDRs H1 and H3 were involved in the formation of KD035-VEGFR2(D2–3) complex. Specifically, T28 (CDR H1) formed an H-bond with N259 (D3) and E261 (D3) whereas W31 (CDR H1) formed an H-bond with D257 (D3). Additionally, the carbonyl of G26 (CDR H1) is mediating an H-bond with K317 (D3). Interestingly, the mouse homologues of N259 (D3) of E261 (D3) are threonine and histidine, respectively (Figure 2C) and these positions belong to the Group 2 of human–mouse differences in D3. We believe the mouse amino acids at these positions would be able to mediate the described interactions with KD035 CDR H1 which is in accordance with the lack of binding restoration when the human amino acids of Group 2 were introduced into the D3 chimera (Figure 2E). On the other hand, K317 (D3) belongs to Group 4 but it still possible that its interaction with KD035 CDR1 is conserved by the arginine in the mouse orthologue position. Additional interactions include K144 (D2), which establishes a salt bridge with the side chain of D50 (CDR L2) and an H-bond with the backbone carbonyl of E30 (CDR L1; Figure 4E).

3.5 | Model validation: Mutagenesis and binding assays

In order to confirm the molecular basis of the KD035(Fv)/VEGFR2 (D2–3) interaction suggested by our model, we generated a series of VEGFR2 mutants bearing substitutions in some of the residues predicted to be involved in establishing key contacts. The mutant L313R (D3) was designed to disrupt the net of hydrophobic contacts mediated by L313, by introducing a large, polar residue found in the mouse homologue (Figure 4A). We also generated a D257A/E261A (D3) double mutant, intended to disrupt the interaction of D257 and E261 with W31 and T28 on KD035 CDR H1 (Figure 4D). As an additional and negative control, we made a third mutant, M197A/M211A/I213A, representing substitutions in D2 amino acids that form a large hydrophobic patch involved in the interaction with VEGF-A; however, not predicted to be involved in KD035 binding by our model[46] (see below).

The VEGFR2(D2–3) harboring these modifications were tested for their interaction with VEGF-A, as a mean to check for deleterious effects. VEGFR2 (D2–3) carrying the M197A/M211A/I213A mutation expectedly abolished the binding to VEGF-A, while having no effect on the interaction with KD035 (Figure 5A). The L313R and D257A/E261A mutants were able to bind to VEGF-A to a similar extent as the WT VEGFR2(D2–3), confirming that the introduced substitutions did not have any deleterious effect on protein folding. Consistent with our model, where M197, M211, and I21 were not part of the KD035-VEGFR2(D2–3) interface, KD035 bound to the WT and the M197A/M211A/I213A mutants, while the interaction was completely abolished for the L313R mutant and to a large extent for the D257A/E261A double mutant construct (Figure 5A). This confirmed our model driven hypothesis that L313 was part of a net of hydrophobic
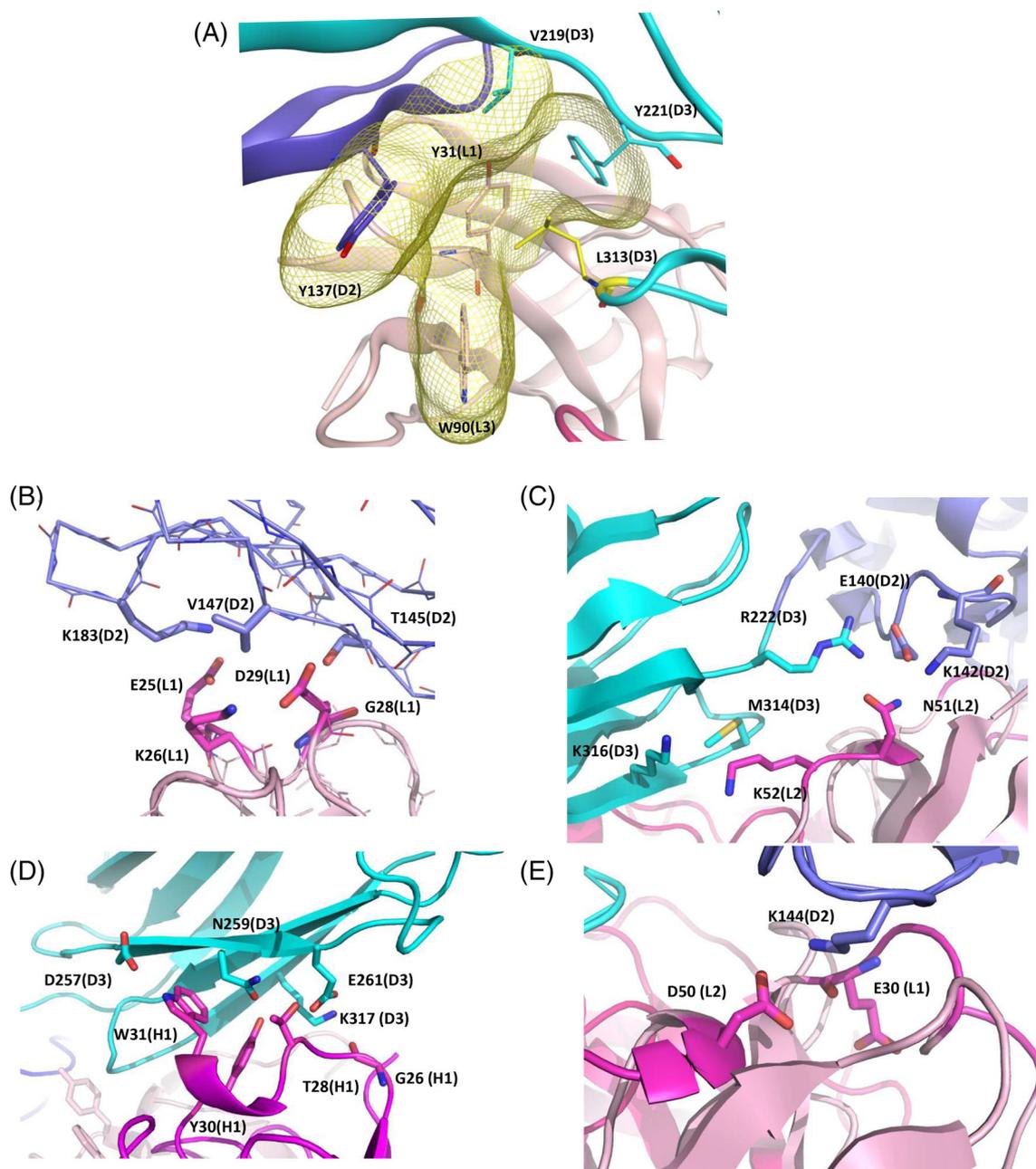
FIGURE 4 Interactions shown by the KD035-VEGFR2(D2-3) model. (A) Hydrophobic interactions mediated by L313. Side chains involved in the interactions are shown in stick format. L313 is colored yellow and the rest of the side chains are colored according to the domain or chain they belong to. The yellow netted surface depicts the extent of the hydrophobic pocket surrounding L313. (B, C) Interactions mediated by KD035 L1 and L2, respectively. (D, E) Interactions mediated by KD035 HC. VEGFR2 D2 is colored blue and the D3 is colored cyan. KD035 HC is colored red and the LC is colored pink. Amino acid labels include domain or chain information in parentheses

contacts contributing to the interaction with KD035 (Figure 4A), which could be disrupted by the introduction of a large polar residue. It also validated our model-based observations for the interaction between the D257/E261 pair and KD035 CDR H1 (Figure 4D).

In order to quantify the impact of these mutations and possibly correlate them with the strength of the related interactions, we measured the affinity of the VEGFR2(D2–3) mutants to KD035 by SPR. In line with the ELISA data, this analysis showed that the KD035 binding was completely abolished by the introduction of the L313R mutation

(Figure 5B), whereas the D257A/E261A mutation demonstrated a 27.5-fold decrease in affinity. Specifically, the resulting K_D values were 1.3 nM for VEGFR2(D2–3) WT and 35.8 nM for the D257A/E261A mutant. As anticipated, the control M197A/M211A/I213A mutant showed a modest fourfold decrease in affinity (Figure 5B). Data obtained from this mutagenesis analysis provided experimental validation for our in silico model of the KD035-VEGFR2(D2–3) complex, described in the highest ranked pose generated by the docking analysis.
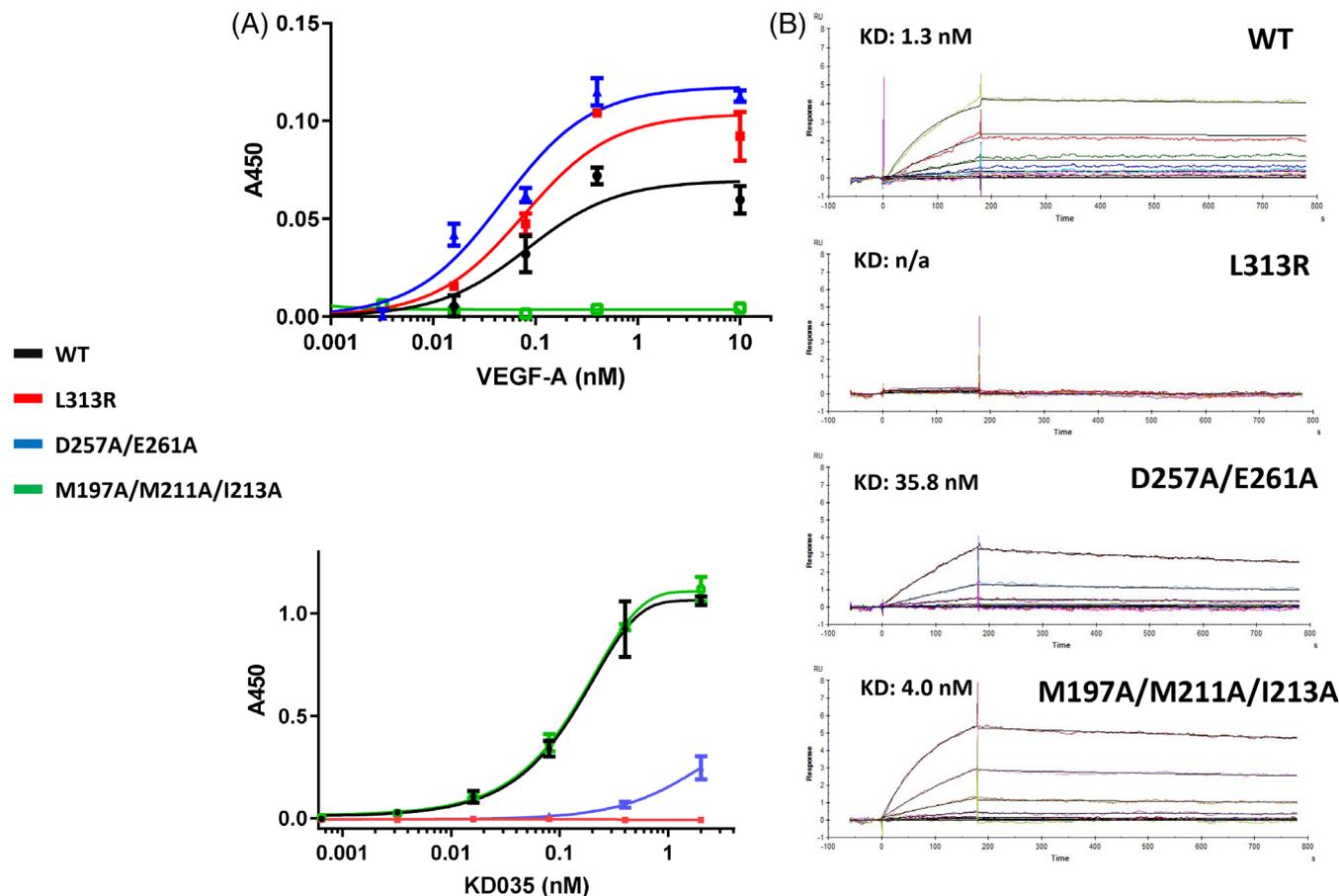
FIGURE 5 Binding analysis of KD035 to VEGFR2(D2-3) mutants. (A) ELISA binding curves to VEGF-A (top) and KD035(bottom) are shown for VEGFR2(D2-3) wild-type (WT) and the mutants L313R (red), D257A/E261A (blue) and M197A/M211A/I213A (green). Points and error bars represent mean ± SEM ($n = 3$) of absolute readout values for absorbance at 450 nm (A450). Dots represent mean ± SD of duplicate or triplicate measurements. (B) VEGFR2(D2-3) B: SPR binding analysis of KD035 (ligand) to a concentration series of the VEGFR2(D2-3) WT or mutants (analytes). Fitting traces are shown superimposed to the experimental sensorgrams with the affinity values calculated from the fitting following a single kinetic model (inset). The sensorgrams shown belong to one representative experiment of duplicate measurements

3.6 | Model validation: Molecular dynamics simulation

In order to further validate our model and the interactions suggested, we performed a 50 ns MD simulation. The superposition of the trajectory frames taken 1 ns apart allowed for the visualization of the structural variations produced by the movement of the VEGFR2(D2–3) loops and the relative stability of the antibody chains (Figure 6A). Interestingly, the bulk of the VEGFR2 variation was on the face opposite to the KD035 interface and primarily occurred in the VEGF-binding site. The extent of these variations is quite evident in the graphical representation of the calculation of the RMSD between the C-alpha chains of the trajectory frames taking the starting modeled coordinates as reference file (Figure 6B, KD035, filled circles and VGFR2(D2–3), open circles). In fact, KD035 chains show a small variation during the trajectory, with RMSD values ranging between 1.0 and 1.5 Å. On the other hand, the RMSD values for the VEGFR2(D2–3) coordinates jump up to 6.0 Å in the first 20 ns, to become stabilized at 4.0 Å throughout the rest of the trajectory (Figure 6B). The

presence of stable RMSD values indicate that an energy minimum was reached and that the 50 ns sampling time was adequate for this analysis. The analysis of the interactions present in each frame was done using the antibody–antigen interaction fingerprint tool in MOE (Figure 6C). For each frame, an occurrence of an interaction is represented by a black bar. The pairwise analysis of KD035 and VEGFR2 fingerprints allowed us to determine the specific interactions that are conserved throughout the simulation. As expected, the hydrophobic interface involving L313 (D3) is conserved, involving the interaction with the KD035 residues Y31 (CDR L1) and W90 (CDR L3; Figures 6C and 4A). The interaction involving the KD035 light chain restraints, E25 (CDR L1)—K183 (D2) and D29 (CDR L1)—V147 (D2) were maintained, alongside with the neighboring interactions involving G28 (CDR L1; Figures 6C and 4B). Additionally, the interface buried interactions between K144 (D2) and the KD035 amino acids E30 (CDR L1) and D50 (CDR L2) is also conserved during the simulation analysis (Figures 6C and 4E). The interactions surrounding N51 and K52 (CDR L2), involving side chains from VEGFR2 D2 (E140 and K142) and D3 (M314 and K316) domains were not strong enough to
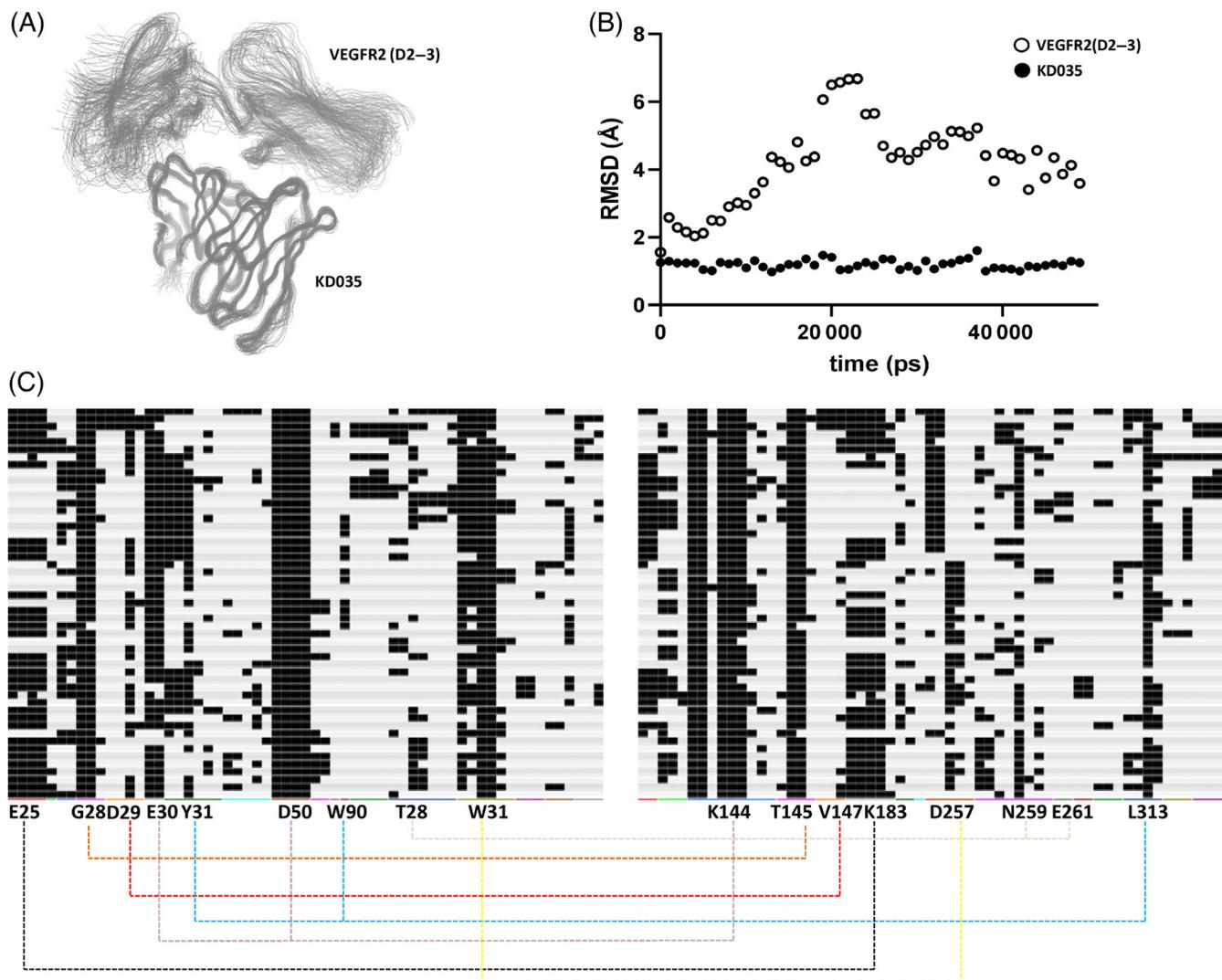
FIGURE 6 MD simulation of the KD035 (Fv)/VEGFR2 (D2-3) complex. (A) Ribbon rendering of the superposed frames of the simulation trajectory. (B) Graph of the RMSDs for the C-alpha chains of KD035 (solid circles) and the VEGFR2 (D2-3) (open circles) over the 50 ns time frame of the simulation. (C) 2D Antibody–antigen interaction fingerprint. Each black bar represents a contact occurring in a specific frame of the simulation for KD035 (left) and VEGFR2 (D2-3) (right). Interactions predicted by the model between specific amino acids of the antibody and the receptor are delineated by the dotted lines between residues listed on the horizontal axis

be maintained throughout the simulation. This set of interactions occurs at the linker between the VEGFR2 D2 and D3 domains which is expectedly an area of high flexibility (Figures 6C and 4C). Lastly, the interactions predicted for the KD035 heavy chain W31 (CDR H1)—D257 (D3) and T28 (CDR H1)—N259 (D3)/E261 (D3) are present throughout the frames (Figures 6C and 4D). This last set was also validated by the mutagenesis and binding assays where the VEGFR2(D2–3) D257A/E261A mutant demonstrated diminished binding with KD035 (Figure 5). Overall, MD analysis strongly supported the involvement of the interactions suggested by our model (particularly those involving L313) in the formation of a stable complex. MD and fingerprint analysis, further supported by Together with the mutagenesis data, 2D fingerprint of the interactions in the complex over the MD simulation provided robust validation for our model of the KD035 (Fv)/VEGFR2(D2–3) interaction.

3.7 | KD035 inhibits VEGFR2 by blocking VEGF-receptor interaction

The therapeutic effects of VEGFR2 neutralizing antibodies stem from their ability to block ligand-mediated receptor activation via disruption of complex formation between the receptor and the VEGF homo-dimer. We wanted to confirm that KD035 was able to block VEGF-receptor interaction and also extend our modeling studies to characterize the molecular mechanism of this functional antagonism. X-ray structures of VEGF-A, VEGF-C, and VEGF-E bound to VEGFR2(D2–3) have been previously reported.[45,46] Given that these three ligands share the same binding site on the receptor, we focused on VEGF-C for our comparison with the binding site of KD035. To this end, we superposed the VEGFR2(D2–3)s from the VEGF-C/VEGFR2(D2–3) structure (PDB:2X1W[46]) and from our KD035(Fv)/VEGFR2(D2–3)

model. The resulting VEGF-C/VEGFR2(D2–3)/KD035 ensemble model is shown in Figure 7A. The inspection of the relative orientation of VEGF-C and KD035-binding sites allowed us analyzes the potential for competition between of VEGF-C and KD035. The superposed structures showed that, whereas both VEGF-C and KD035 have their binding sites around the linker between D2 and D3, they bound to opposite sides of VEGFR(D2–D3) with little superposition (Figure 7A). Interestingly, the shared site of interaction included the contacts mediated by loop L3 from VEGF-C (side chains 188–196) which interacted with G312 and are close to L313 from VEGFR2 D3 (Figure 7B). Therefore, we propose that the engagement of L313 by KD035 plays a role in the ability of this antibody to compete with VEGF-C for the interaction with VEGFR2(D2–3).

Although not completely overlapping, KD035 and VEGF-C-binding sites are immediately adjacent, which could cause a steric overlap that would prevent VEGF-C (and by extension VEGF-A and VEGF-E) binding when KD035 is present. In order to confirm that the closeness of the binding sites would in fact result in a competitive displacement of the natural ligand by KD035 leading to functional receptor antagonism, we performed quantitative competition assays. We measured the extent of displacement of VEGF-A or VEGF-C from the receptor by the antibody by a competitive ELISA. KD035 capacity to compete with VEGF-A or VEGF-C was compared to ramucirumab, VEGFR2 targeting antibody, whose binding and activity has been extensively characterized.[50] KD035 effectively blocked VEGF-A and VEGF-C binding at low concentrations, with EC_{50}s of 357 pM and 620 pM, respectively (Figure 7C). KD035 was similarly potent as ramucirumab whose EC_{50}s were 470 pM and 690 pM, highlighting the potential clinical utility of KD035. The ability of KD035 to block VEGFs binding to the receptor in a dose dependent manner also validated our model where our antibody and the natural ligand were positioned to have adjacent, slightly superposed binding sites on the receptor.

4 | DISCUSSION

Data-driven approaches to computational modeling are successful in generating protein–protein structural models, which are predictive of the native complex confirmation with a high degree of accuracy.[3,26–31,34]
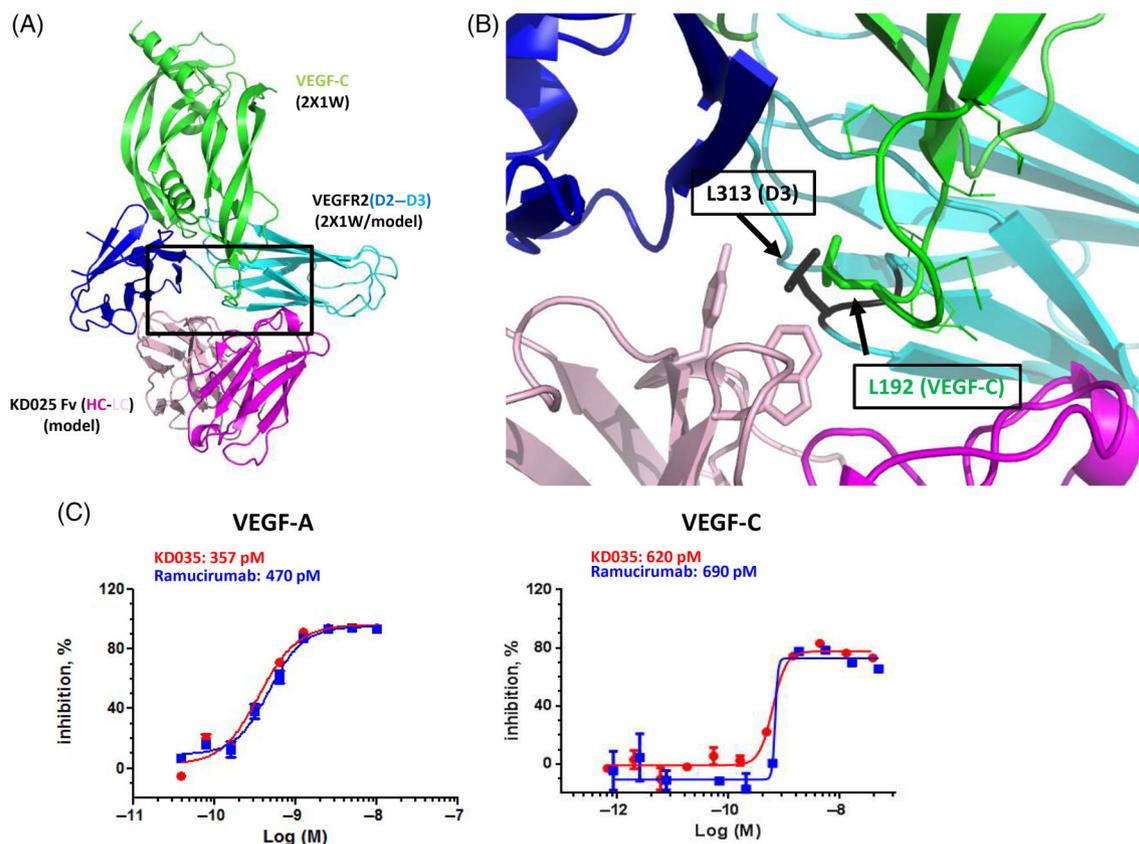

FIGURE 7 Structural basis of KD035 competition with VEGF-C. (A) Superposed ensemble resulting from overlaying of the VEGFR2(D2-3) domains from the KD035 complex model and from the VEGFR2(D2-3)/VEGF-C structure (PDB: 2X1W). VEGF-C is colored green. (B) Zoomed-in view of the region boxed in (A). L313(D3) is shown in stick format and colored black. Side chains from either KD035 or VEGF-C interacting with L313 are shown in stick format. The interaction L313(D3)-L192(VEGF-C) is depicted. The color of sticks and legends follows the rule described before for D2, D3 and VEGF-C. (C) ELISA based competition assays showing the displacement of VEGF-A (left) and VEGF-C (right) from their binding site at VEGFR2(D2-3) by KD035 and ramucirumab. KD035 and ramucirumab plots are potted as mean ± SEM of duplicate or triplicate experiments. EC_{50} values calculated using a single site fit model are shown in the inset

These methods use information about the CDRs and binding epitopes to introduce restraints that increase the accuracy of the modeling result. The data-driven approach used our study combined the comparison of the crystal structures of KD035 and its parental antibody with the data from the KD035-binding affinity changes with human and mouse VEGFR2. These data combined to form model restraints and define contacts for the antibody and the receptor. We employed the analysis of protein surfaces and enumeration of the differences in the distribution of charged patches as indicators of residues in KD035 that could be taken in consideration for paratope definition during the docking processes. A comparative analysis between the parental clone and KD035 focusing on the topological features of the paratope surface in combination with the charge distribution allowed for further definition of the binding interface.

While the modeling field is constantly improving, currently, there are still no scoring methods reliable enough to provide an adequate degree of certainty in the generated complex predictions. As with the experimentally determined structures, further validation is generally required to justify the proposed binding modes. In our study, having generated a series of possible binding poses we focused on the complex with the lowest free energy score and went on to validate the accuracy of the prediction via a mutagenesis-based analysis. In fact, our ELISA and SPR assays confirmed the role of the most relevant interactions suggested by the model. Additionally, as mentioned in Section 3, the accuracy of the selected pose is also supported by the fact that the predicted interactions provide an explanation for how changes generated by affinity maturation improved VEGFR2 affinity between the parental B1 and KD035 antibodies. Furthermore, and consistent with our initial domain mapping analysis, the model shows that only VEGFR2 D3 amino acids included in the Group 4 of human–mouse differences mediate interactions that would be severely disrupted by the replacement with the mouse orthologue. This is very remarkable for the interactions mediated by L313, which is an arginine in the mouse VEGFR2(D3) sequence. Among the remaining amino acids belonging to Group 4, which were set up as restraints, only K317 mediates an H-bond which would be conserved by an arginine in mouse VEGFR2. The remaining restraints (L304 and S319) are located away from the modeled interface.

The analysis of the model and the confirmatory mutagenesis showed that the extensive network of hydrophobic interactions mediated by KD035 L1 and L3 CDRs with L313 of VEGFR2(D3) is an important driver of affinity (Figure 4A). Existence of this network of hydrophobic interactions surrounding L313 also provides an explanation for the lack of KD035-binding to the mouse VEGFR2. This set of interactions located at the linker between the VEGFR2 domains suggests the possibility that the binding of KD035 may change their relative position, which may further contribute to the inhibitory activity if the ligand-binding site is closely located.

Related to this, our model suggests that the interactions mediated by L313 contribute to the receptor antagonist activity of KD035, destabilizing the VEFG binding, and blocking functional receptor–ligand complex formation (Figure 7B). However, other than the interactions mediated with L313, there are no additional points within the

VEGFR2(D2–3) surface where the KD035 and VEGF binding sites overlap. Based on our model and experimental data demonstrating robust competition between KD035 and VEGFs, we speculate that the displacement of the ligand from the receptor occurs primarily due to the steric hindrance exerted by the proximity of the binding sites. While not directly addressed by our data, the effect of KD035 on the relative orientation of the VEGFR2 D2 and D3 should not be discarded. It is possible that upon antibody binding, these domains veer away from the optimal orientation required for the formation of the VEGF binding site thus further impeding optimal ligand binding. The change in the relative orientation of D2 and D3 domains could be stabilized by the hydrophobic network of interactions close to the D2–D3 linker as mentioned above, and also by the described interactions of KDR035 LC and HC with D2 and D3, respectively.

Protein–protein interaction energetics are complex, involving aspects of protein folding as well as conformational changes on protein surfaces (induced fit).[61–63] Understandably, the inherent flexibility of each monomer contributes to the energy landscape of the complex.[61] MD simulations are frequently utilized as a means to complement the information obtained by static analysis of protein–protein docking[64] and also to correct errors in the crystal structure mainly due to packing forces. The inspection of the symmetry mates for our B1 and KD035 structures show no involvement in the packing interfaces of the CDR loops or areas which could otherwise affect their conformation. However, the dynamics of the interface between the light and heavy chains in the antibody influence the relative orientation of these chains. This, alongside with the inherent flexibility of the CDR H3 loop, determine the plasticity of the paratope surface[65–67] and affect the energy landscape of the antibody–antigen complex.[63] Epitopes, on the other hand are reportedly less flexible than paratopes and non-epitopic regions.[66] The stability of the suggested interactions was assessed by means of MD simulation. Previous studies showed that the trends that allow to differentiate between native and non-native complexes are already evident within the first 5 ns, and are only augmented during the rest of the simulation.[64] Our analysis of the trajectory during the 50 ns simulation showed limited variation in the epitopic region and emphasized the stability of the Van der Waal contacts established by L313, which are central for the functional KD035 complex formation. On the other hand, marginal H-bond and salt bridge mediated interactions showed a greater instability. This highlights the importance of the hydrophobic interactions, which were also previously delineated for other therapeutic antibodies.[68]

In summary, combining information from experimentally determined structures, data from binding assays and species cross reactivity we were able to generate a model of the binding site between KD035 on VEGFR2. The validation of an in silico protein–protein model via MD simulation and directed mutagenesis (a method widely used to validate experimentally determined complexes) is an acceptable proof of accuracy. Often, the validation of models similar to ours, relies largely on the macroscopic comparison with experimentally obtained structures, without validating the accuracy of the interaction surface at the level of the amino acid interactions.[3,28,29,69] Experimental validation and further analysis of the model allowed us to propose

a reliable molecular mechanism for the competitive receptor occupancy between KD035 and VEGFs. We expect that future developments in epitope mapping, possibly fueled by the improvements in the topological analyses of protein surfaces[70] together with the application of machine learning methods will greatly enhance the accuracy of in silico models.[71] In this regard, descriptions of structural and energetic properties generated by the MD simulations will undoubtedly contribute to the development of more accurate scoring methods.[64]

Overall, the reliability of our model, supported by MD simulation and mutagenesis analyses, sets up a protocol for a multidisciplinary approach where protein–protein docking output can be effectively restrained by the supplementary experimental information.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/prot.26280.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the Protein Data Bank (PDB) at http://doi.org/10.2210/pdb7LSI/pdb and http://doi.org/10.2210/pdb7LSB/pdb, reference numbers 7LSI and 7LSB.

ORCID

Rafael S. Depetris https://orcid.org/0000-0001-9707-6922
Amari Tankard https://orcid.org/0000-0003-3257-1652
Michael Drummond https://orcid.org/0000-0002-4688-3975
Chris Williams https://orcid.org/0000-0001-5428-2742
Masha V. Poyurovsky https://orcid.org/0000-0002-3024-778X

REFERENCES

1. Kaplon H, Muralidharan M, Schneider Z, Reichert JM. Antibodies to watch in 2020. *MAbs*. 2020;12:1703531. doi:10.1080/19420862.2019.1703531
2. Sela-Culang I, Kunik V, Ofran Y. The structural basis of antibody-antigen recognition. *Front Immunol*. 2013;4:302. doi:10.3389/fimmu.2013.00302
3. Cannon DA, Shan L, Du Q, et al. Experimentally guided computational antibody affinity maturation with de novo docking, modelling and rational design. *PLoS Comput Biol*. 2019;15:e1006980. doi:10.1371/journal.pcbi.1006980
4. Krawczyk K, Dunbar J, Deane CM. Computational tools for aiding rational antibody design. *Methods Mol Biol*. 2017;1529:399-416. doi:10.1007/978-1-4939-6637-0_21
5. Deng X, Storz U, Doranz BJ. Enhancing antibody patent protection using epitope mapping information. *MAbs*. 2018;10:204-209. doi:10.1080/19420862.2017.1402998
6. Wolf Perez AM, Sormanni P, Andersen JS, et al. In vitro and in silico assessment of the developability of a designed monoclonal antibody library. *MAbs*. 2019;11:388-400. doi:10.1080/19420862.2018.1556082
7. Blech M, Peter D, Fischer P, et al. One target-two different binding modes: structural insights into gevokizumab and canakinumab interactions to interleukin-1beta. *J Mol Biol*. 2013;425:94-111. doi:10.1016/j.jmb.2012.09.021
8. Bianchi M, Turner HL, Nogal B, et al. Electron-microscopy-based epitope mapping defines specificities of polyclonal antibodies elicited during HIV-1 BG505 envelope trimer immunization. *Immunity*. 2018;49:288-300 e288. doi:10.1016/j.immuni.2018.07.009
9. Ultsch M, Bevers J, Nakamura G, et al. Structural basis of signaling blockade by anti-IL-13 antibody Lebrikizumab. *J Mol Biol*. 2013;425:1330-1339. doi:10.1016/j.jmb.2013.01.024
10. Leem J, Dunbar J, Georges G, Shi J, Deane CM. A body builder: automated antibody structure prediction with data-driven accuracy estimation. *MAbs*. 2016;8:1259-1268. doi:10.1080/19420862.2016.1205773
11. Lepore R, Olimpieri PP, Messih MA, Tramontano A. PIGSPro: prediction of immunoGlobulin structures v2. *Nucleic Acids Res*. 2017;45:W17-W23. doi:10.1093/nar/gkx334
12. Weitzner BD, Jeliazkov JR, Lyskov S, et al. Modeling and docking of antibody structures with Rosetta. *Nat Protoc*. 2017;12:401-416. doi:10.1038/nprot.2016.180
13. Ansari HR, Raghava GP. Identification of conformational B-cell epitopes in an antigen from its primary sequence. *Immunome Res*. 2010;6:6. doi:10.1186/1745-7580-6-6
14. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res*. 2017;45:W24-W29. doi:10.1093/nar/gkx346
15. Kringelum JV, Lundegaard C, Lund O, Nielsen M. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS Comput Biol*. 2012;8:e1002829. doi:10.1371/journal.pcbi.1002829
16. Kunik V, Ashkenazi S, Ofran Y. Paratome: an online tool for systematic identification of antigen-binding regions in antibodies based on sequence or structure. *Nucleic Acids Res*. 2012;40:W521-W524. doi:10.1093/nar/gks480
17. Liang S, Zheng D, Standley DM, Yao B, Zacharias M, Zhang C. EPSVR and EPMeta: prediction of antigenic epitopes using support vector regression and multiple server results. *BMC Bioinformatics*. 2010;11:381. doi:10.1186/1471-2105-11-381
18. Liberis E, Velickovic P, Sormanni P, Vendruscolo M, Lio P. Parapred: antibody paratope prediction using convolutional and recurrent neural networks. *Bioinformatics*. 2018;34:2944-2950. doi:10.1093/bioinformatics/bty305
19. Olimpieri PP, Chailyan A, Tramontano A, Marcatili P. Prediction of site-specific interactions in antibody-antigen complexes: the proABC method and server. *Bioinformatics*. 2013;29:2285-2291. doi:10.1093/bioinformatics/btt369
20. Qi T, Qiu T, Zhang Q, et al. SEPPA 2.0—more refined server to predict spatial epitope considering species of immune host and subcellular localization of protein antigen. *Nucleic Acids Res*. 2014;42:W59-W63. doi:10.1093/nar/gku395
21. Ponomarenko JV, Bourne PE. Antibody-protein interactions: benchmark datasets and prediction tools evaluation. *BMC Struct Biol*. 2007;7:64. doi:10.1186/1472-6807-7-64
22. Moal IH, Torchala M, Bates PA, Fernandez-Recio J. The scoring of poses in protein-protein docking: current capabilities and future directions. *BMC Bioinformatics*. 2013;14:286. doi:10.1186/1471-2105-14-286
23. Labute P. The generalized born/volume integral implicit solvent model: estimation of the free energy of hydration using London dispersion instead of atomic surface area. *J Comput Chem*. 2008;29:1693-1698. doi:10.1002/jcc.20933
24. Halakou F, Gursoy A, Keskin O. Embedding alternative conformations of proteins in protein-protein interaction networks. *Methods Mol Biol*. 2020;2074:113-124. doi:10.1007/978-1-4939-9873-9_9

25. Ritchie DW. Recent progress and future directions in protein-protein docking. *Curr Protein Pept Sci*. 2008;9:1-15. doi:10.2174/138920308783565741

26. Loyau J, Didelot G, Malinge P, et al. Robust antibody-antigen complexes prediction generated by combining sequence analyses, mutagenesis, in vitro evolution, X-ray crystallography and in silico docking. *J Mol Biol*. 2015;427:2647-2662. doi:10.1016/j.jmb.2015.05.016

27. Gabb HA, Jackson RM, Sternberg MJ. Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J Mol Biol*. 1997;272:106-120. doi:10.1006/jmbi.1997.1203

28. Ambrosetti F, Jimenez-Garcia B, Roel-Touris J, Bonvin A. Modeling antibody-antigen complexes by information-driven docking. *Structure*. 2020;28:119-129 e112. doi:10.1016/j.str.2019.10.011

29. Dominguez C, Boelens R, Bonvin AM. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc*. 2003;125:1731-1737. doi:10.1021/ja026939x

30. Janin J, Henrick K, Moult J, et al. CAPRI: a Critical Assessment of PRedicted Interactions. *Proteins*. 2003;52:2-9. doi:10.1002/prot.10381

31. Mendez R, Leplae R, De Maria L, Wodak SJ. Assessment of blind predictions of protein-protein interactions: current status of docking methods. *Proteins*. 2003;52:51-67. doi:10.1002/prot.10393

32. Basu S, Wallner B. DockQ: a quality measure for protein-protein docking models. *PLoS One*. 2016;11:e0161879. doi:10.1371/journal.pone.0161879

33. Basu S, Wallner B. Finding correct protein-protein docking models using ProQDock. *Bioinformatics*. 2016;32:i262-i270. doi:10.1093/bioinformatics/btw257

34. Lohela M, Bry M, Tammela T, Alitalo K. VEGFs and receptors involved in angiogenesis versus lymphangiogenesis. *Curr Opin Cell Biol*. 2009;21:154-165. doi:10.1016/j.ceb.2008.12.012

35. Roskoski R Jr. Vascular endothelial growth factor (VEGF) and VEGF receptor inhibitors in the treatment of renal cell carcinomas. *Pharmacol Res*. 2017;120:116-132. doi:10.1016/j.phrs.2017.03.010

36. Grunewald FS, Prota AE, Giese A, Ballmer-Hofer K. Structure-function analysis of VEGF receptor activation and the role of coreceptors in angiogenic signaling. *Biochim Biophys Acta*. 1804;567-580:2010-2580. doi:10.1016/j.bbapap.2009.09.002

37. Shibuya M. Vascular endothelial growth factor (VEGF)-Receptor2: its biological functions, major signaling pathway, and specific ligand VEGF-E. *Endothelium*. 2006;13:63-69. doi:10.1080/10623320600697955

38. Shibuya M, Claesson-Welsh L. Signal transduction by VEGF receptors in regulation of angiogenesis and lymphangiogenesis. *Exp Cell Res*. 2006;312:549-560. doi:10.1016/j.yexcr.2005.11.012

39. Roskoski R Jr. VEGF receptor protein-tyrosine kinases: structure and regulation. *Biochem Biophys Res Commun*. 2008;375:287-291. doi:10.1016/j.bbrc.2008.07.121

40. Christinger HW, Fuh G, de Vos AM, Wiesmann C. The crystal structure of placental growth factor in complex with domain 2 of vascular endothelial growth factor receptor-1. *J Biol Chem*. 2004;279:10382-10388. doi:10.1074/jbc.M313237200

41. Iyer S, Darley PI, Acharya KR. Structural insights into the binding of vascular endothelial growth factor-B by VEGFR-1(D2): recognition and specificity. *J Biol Chem*. 2010;285:23779-23789. doi:10.1074/jbc.M110.130658

42. Markovic-Mueller S, Stuttfeld E, Asthana M, et al. Structure of the full-length VEGFR-1 extracellular domain in complex with VEGF-A. *Structure*. 2017;25:341-352. doi:10.1016/j.str.2016.12.012

43. Starovasnik MA, Christinger HW, Wiesmann C, Champe MA, de Vos AM, Skelton NJ. Solution structure of the VEGF-binding domain of Flt-1: comparison of its free and bound states. *J Mol Biol*. 1999;293:531-544. doi:10.1006/jmbi.1999.3134

44. Wiesmann C, Fuh G, Christinger HW, Eigenbrot C, Wells JA, de Vos AM. Crystal structure at 1.7A resolution of VEGF in complex with domain 2 of the Flt-1 receptor. *Cell*. 1997;91:695-704.

45. Brozzo MS, Bjelić S, Kisko K, et al. Thermodynamic and structural description of allosterically regulated VEGFR-2 dimerization. *Blood*. 2012;119:1781-1788. doi:10.1182/blood-2011-11-390922

46. Leppanen VM, Prota AE, Jeltsch M, et al. Structural determinants of growth factor binding and specificity by VEGF receptor 2. *Proc Natl Acad Sci USA*. 2010;107:2425-2430. doi:10.1073/pnas.0914318107

47. Ruch C, Skiniotis G, Steinmetz MO, Walz T, Ballmer-Hofer K. Structure of a VEGF-VEGF receptor complex determined by electron microscopy. *Nat Struct Mol Biol*. 2007;14:249-250. doi:10.1038/nsmb1202

48. Kisko K, Brozzo MS, Missimer J, et al. Structural analysis of vascular endothelial growth factor receptor-2/ligand complexes by small-angle X-ray solution scattering. *FASEB J*. 2011;25:2980-2986. doi:10.1096/fj.11-185397

49. Fuh G, Li B, Crowley C, Cunningham B, Wells JA. Requirements for binding and signaling of the kinase domain receptor for vascular endothelial growth factor. *J Biol Chem*. 1998;273:11197-11204.

50. Franklin MC, Navarro EC, Wang Y, et al. The structural basis for the function of two anti-VEGF receptor 2 antibodies. *Structure*. 2011;19:1097-1107. doi:10.1016/j.str.2011.01.019

51. Xu Y, Zhang X, Wang Y, Pan M, Wang M, Zhang J. A VEGFR2-MICA bispecific antibody activates tumor-infiltrating lymphocytes and exhibits potent anti-tumor efficacy in mice. *Cancer Immunol Immunother*. 2019;68:1429-1441. doi:10.1007/s00262-019-02379-9

52. Lu D, Kussie P, Pytowski B, et al. Identification of the residues in the extracellular region of KDR important for interaction with vascular endothelial growth factor and neutralizing anti-KDR antibodies. *J Biol Chem*. 2000;275:14321-14330.

53. Lu D, Shen J, Vil MD, et al. Tailoring in vitro selection for a picomolar affinity human antibody directed against vascular endothelial growth factor receptor 2 for enhanced neutralizing activity. *J Biol Chem*. 2003;278:43496-43507. doi:10.1074/jbc.M307742200

54. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr*. 2006;62:859-866. doi:10.1107/S0907444906019949

55. Liebschner D, Afonine PV, Baker ML, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol*. 2019;75:861-877. doi:10.1107/S2059798319011471

56. Kaufmann B, Vogt MR, Goudsmit J, et al. Neutralization of West Nile virus by cross-linking of its surface proteins with fab fragments of the human monoclonal antibody CR4354. *Proc Natl Acad Sci U S A*. 2010;107:18950-18955. doi:10.1073/pnas.1011036107

57. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of coot. *Acta Crystallogr D Biol Crystallogr*. 2010;66:486-501. doi:10.1107/S0907444910007493

58. Molecular Operating Environment (MOE), 2020.09 (2020).

59. Yershova A, Jain S, Lavalle SM, Mitchell JC. Generating uniform incremental grids on SO(3) using the Hopf Fibration. *Int J Rob Res*. 2010;29:801-812. doi:10.1177/0278364909352700

60. Phillips JC, Hardy DJ, Maia JDC, et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys*. 2020;153:44130. doi:10.1063/5.0014475

61. Ma B, Kumar S, Tsai CJ, Nussinov R. Folding funnels and binding mechanisms. *Protein Eng*. 1999;12:713-720. doi:10.1093/protein/12.9.713

62. Tsai CJ, Kumar S, Ma B, Nussinov R. Folding funnels, binding funnels, and protein function. *Protein Sci*. 1999;8:1181-1190. doi:10.1110/ps.8.6.1181

63. Tsai CJ, Ma B, Nussinov R. Folding and binding cascades: shifts in energy landscapes. *Proc Natl Acad Sci U S A*. 1999;96:9970-9972. doi:10.1073/pnas.96.18.9970

64. Jandova Z, Vargiu AV, Bonvin A. Native or non-native protein-protein docking models? Molecular dynamics to the rescue. *J Chem Theory Comput*. 2021;17:5944-5954. doi:10.1021/acs.jctc.1c00336

65. Fernandez-Quintero ML, Kraml J, Georges G, Liedl KR. CDR-H3 loop ensemble in solution—conformational selection upon antibody binding. *MAbs*. 2019;11:1077-1088. doi:10.1080/19420862.2019.1618676

66. Fernandez-Quintero ML, Loeffler JR, Waibl F, et al. Conformational selection of allergen-antibody complexes-surface plasticity of paratopes and epitopes. *Protein Eng Des Sel*. 2019;32:513-523. doi:10.1093/protein/gzaa014

67. Fernandez-Quintero ML, Pomarici ND, Math BA, et al. Antibodies exhibit multiple paratope states influencing VH-VL domain orientations. *Commun Biol*. 2020;3:589. doi:10.1038/s42003-020-01319-z

68. Horita S, Nomura Y, Sato Y, Shimamura T, Iwata S, Nomura N. High-resolution crystal structure of the therapeutic antibody pembrolizumab bound to the human PD-1. *Sci Rep*. 2016;6:35297. doi:10.1038/srep35297

69. Kilambi KP, Gray JJ. Structure-based cross-docking analysis of antibody-antigen interactions. *Sci Rep*. 2017;7:8145. doi:10.1038/s41598-017-08414-y

70. Rudden LSP, Degiacomi MT. Protein docking using a single representation for protein surface, electrostatics, and local dynamics. *J Chem Theory Comput*. 2019;15:5135-5143. doi:10.1021/acs.jctc.9b00474

71. Marks C, Hummer AM, Chin M, Deane CM. Humanization of antibodies using a machine learning approach on large-scale repertoire data. *Bioinformatics*. 2021. doi:10.1093/bioinformatics/btab434