

SCIENTIFIC REPORTS



OPEN

Determinants of the assembly and function of antibody variable domains

Eva Maria Herold^{1,3}, Christine John¹, Benedikt Weber¹, Stephan Kremser², Jonathan Eras⁴, Carolin Berner¹, Sabrina Deubler¹, Martin Zacharias² & Johannes Buchner¹

The antibody Fv module which binds antigen consists of the variable domains V_L and V_H . These exhibit a conserved β -sheet structure and comprise highly variable loops (CDRs). Little is known about the contributions of the framework residues and CDRs to their association. We exchanged conserved interface residues as well as CDR loops and tested the effects on two Fvs interacting with moderate affinities (K_{DS} of $\sim 2.5 \mu\text{M}$ and $\sim 6 \mu\text{M}$). While for the rather instable domains, almost all mutations had a negative effect, the more stable domains tolerated a number of mutations of conserved interface residues. Of particular importance for Fv association are $V_L P44$ and $V_H L45$. In general, the exchange of conserved residues in the V_L/V_H interface did not have uniform effects on domain stability. Furthermore, the effects on association and antigen binding do not strictly correlate. In addition to the interface, the CDRs modulate the variable domain framework to a significant extent as shown by swap experiments. Our study reveals a complex interplay of domain stability, association and antigen binding including an unexpected strong mutual influence of the domain framework and the CDRs on stability/association on the one side and antigen binding on the other side.

In the humoral immune response, antigen recognition is mediated by immunoglobulins, specifically by the N-terminal variable domains of the light chain (V_L) and of the heavy chain (V_H) which associate non-covalently to form the so-called Fv-fragment. Three hyper-variable regions (complementarity determining regions or CDRs) in V_L and V_H comprise the residues interacting with antigens. They account for approximately 25% of the variable domains¹. CDR-H3 (i.e., the third CDR of V_H) is the most diverse of these six regions concerning length and amino acid sequence². Apart from the CDRs, both variable domains exhibit a conserved β -barrel framework stabilized by an internal disulfide bridge¹⁻⁴. They are composed of two β -sheets, one with four strands (A, B, D, E) and one with six strands (A', G, F, C, C', C''), with the strands GFC'C involved in the V_H/V_L interface. Several studies on the V_H/V_L packing geometry have shown that residues within the framework as well as interface contributing residues of the CDRs can influence the interface^{1-3,5,6}. 75% of the interface residues are constituted by framework β -sheets and 25% by the CDRs (inter-strand links between GF, BC and C'C'', respectively)^{1,3}. The interface contribution of the hypervariable loops comprises CDR1 and especially CDR3 residues^{1,3}. In this context, the CDR-H3 can affect V_H/V_L orientation due to small position re-adjustments by the varying number of interface contributing residues³. An additional influence by hypervariable loop residues is exhibited by the so called proximate zone, which is situated at the base of the antigen binding site, comprising residues that do not actively participate in the interface³. Several studies on the V_H-V_L packing geometry showed that residues within the framework as well as interface contributing residues of the CDRs can influence the interface^{1,3,5-7}. As the association of V_L and V_H is crucial for antigen binding³, understanding the underlying principles is of great importance.

For the V_H domain which is composed of about 125 residues and the V_L domain with about 110 residues, Chothia and co-workers suggested that the interface residues at positions 98, 44 and 36 in V_L and 103, 47, 45 and 37 in V_H according to Kabat numbering⁸ are conserved^{1,7}. Wang and co-workers⁹ aimed at identifying amino acid

¹Center for Integrated Protein Science Munich (CIPSM) at the Department Chemie, Technische Universität München, 85747, Garching, Germany. ²Center for Integrated Protein Science Munich (CIPSM) at the Physics Department, Technische Universität München, 85747, Garching, Germany. ³Present address: Sanofi-Aventis GmbH, Industriepark Höchst, 65926, Frankfurt am Main, Germany. ⁴Present address: ETH Zürich, Otto-Stern-Weg 5, 8093, Zuerich, Switzerland. Eva Maria Herold and Christine John contributed equally to this work. Correspondence and requests for materials should be addressed to J.B. (email: johannes.buchner@tum.de)

networks important for V_H and V_L function by covariation analysis. This multiple sequence alignment approach investigates covariations between residues at all possible positions. This allows to reveal conserved amino acids by the correlation of the presence of one particular amino acid with the presence of a second one at a particular sequence position. In their study they included more than 2000 V-class sequences of human, mouse, cow, camel, llama, macaque and chicken with a bias towards human sequences (574 out of 2432). Generally, the majority of the most strongly conserved amino acids identified in this study were positioned at the V_H/V_L interface⁹. For V_L , amino acids Y36, Q37, P44, A43, L46 and F98 were found to be highly conserved, with all residues except Q37 directly in contact with V_H . On the V_H side, amino acids V37, R38, G44, L45, E46, W47 and W103 could be identified in the V_H/V_L interface with all residues except E46 and R38 in direct contact with V_L . V_H W47 seems to be the central node based on the number and strength of its covariations with other interface residues, the same holds true for Y36 and P44 for the V_L domain. Further computational analyses revealed two main modes of interaction for V_H and V_L which is either characterized by a proline or a medium/large hydrophobic residue at position 44 in V_L ⁷. Concerning V_H , W47 seems to be essential as it was previously mutated for improved solubility and stability but none of the mutations (W47L, W47R¹⁰) were favorable. Interface residues can also affect antigen binding^{3,6,11–13} by influencing the positioning of hypervariable loops. Additional experimental studies addressed the influence of the exchange of particular conserved amino acids on the association of V_H and V_L via the stabilities of covalently linked scFv and Fab fragments^{14–18}. While these studies lay the groundwork, we are still far from a detailed and comprehensive understanding of the organization of the Fv interface. In this context, it is important to determine the affinity of the association of V_L and V_H directly. Strikingly, for the other domain interactions in IgG the K_D values differ by several orders of magnitude. For the C_H3 dimer a $K_D < 10^{-10}$ M was determined using SEC^{19,20} and the K_D for the interaction between C_H1 and C_L was 6.2 μ M obtained by the C_L induced change in intrinsic fluorescence of C_H1 ²¹. This analysis is largely lacking for the V_L/V_H interaction.

Here, we chose to use the V_L and V_H domains of the human monoclonal antibody 1HEZ (κ /IgM subclass) and the murine monoclonal antibody MAK33 (κ /IgG1 subclass) as well-studied model systems^{20–23} to analyze the contribution of different factors on V_L and V_H structure and function. The application of variable domains from different species and subclasses should reveal to which extent the mutation-associated effects are conserved. Their sequences contain all the conserved residues identified by Wang and coworkers except for an alanine at position 43 in MAK33 V_L , which is exchanged to serine and a glycine at position 44 in MAK V_H , which is an arginine in MAK33. Since the relative importance of the conserved residues for structure, stability, association and antigen binding is not clear, we mutated every conserved residue against alanine and analyzed the properties of the variants in a comprehensive manner. Additionally, we generated two MAK33 V_L double point mutations (Y36A/P44A, Y36A/S43A) to investigate the potential synergistic nature of the mutation. Importantly, we focused in the analysis on the isolated variable domains and the direct influence of point mutations on their interaction, and not on Fab or scFv fragments as in previous studies^{12,16,18} to draw conclusions concerning their stabilities and antigen binding properties. Furthermore, we performed CDR exchange experiments to address the contribution of these structural elements on domain architecture.

Our results on the effects of mutations on domain structure, stability, association and antigen binding together with CDR exchange experiments reveal complex relationships between structural and functional properties within the V_L and V_H domains.

Results

The role of conserved residues for variable domain structure and stability. To determine the influence of specific residues on the structure, stability, association and functionality of the individual domains, amino acids were selected for mutation which had been predicted to be important⁹. Based on the results of Wang and coworkers we created six single (Y36A, Q37A, S43A, P44A, L46A, and F98A) and two double point mutants for MAK33 V_L (Y36A/P44A and Y36A/S43A) and five single mutants (Y36A, Q37A, P44A, L46A, and F98A) for 1HEZ V_L as well as seven single point mutations for MAK33 V_H (V37A, R38A, R44A, L45A, E46A, W47A, and W103A) and 1HEZ V_H (V37A, R38A, G44A, L45A, E46A, W47A, and W103A). As shown in Fig. 1, these residues lie in or near the interaction interfaces of V_L and V_H . The single domains were expressed in *E. coli*, as insoluble proteins in inclusion bodies. During refolding and purification, it turned out that R38A and E46A in V_H were unstable and aggregation-prone. Therefore, they were not considered further. For the determination of structural changes, we recorded CD spectra of the variants. FUV-CD spectra report on the secondary structure of proteins. All wt domains studied show a typical β -sheet structure (minimum at 218 nm) with the exception of MAK33 V_L , which exhibits a characteristic shape in agreement with previous studies in which we investigated the folding pathway of the protein²⁴ and its amyloidogenic variants^{25–27} (Fig. 2a and b, Fig. 3a and b). In comparison to the spectra of the point mutations, a high overlap with the wt spectra is visible. Only in the case of MAK33 V_H , the point mutants exhibited more pronounced effects on the FUV-CD spectra (Fig. 2b) including shifts in the minimum (e.g. W47A). NUV-CD spectra can be regarded as specific fingerprints for each domain. They report on the tertiary structure. Deviations in the amplitude, especially in the spectra region monitoring tyrosine and phenylalanine, were observed in the NUV spectra for some V_L variants: F98A for MAK33 and P44A, L46A and F98A for 1HEZ as well as for most V_H variants (Fig. 2c and d, Fig. 3c and d). For the V_L F98A variant, the different amplitude can be explained by the lack of the Phe which contributes particularly at wavelengths from 255 to 270 nm to NUV-CD spectra. Other changes in the NUV spectra are not due to the lack of aromatic amino acids since the most significant changes in the spectra are observed from 270 to 285 nm which typically indicate changes in the environment of Tyr and Trp residues. Thus, an altered tertiary structure is expected to be the origin of amplitude deviations, indicating a key structural role for these residues. In general, the CD data show that the secondary and tertiary structure of the V_L domain seems to tolerate single point mutations of conserved residues more than the V_H domain. This holds true especially for the MAK33 variable domains which exhibit

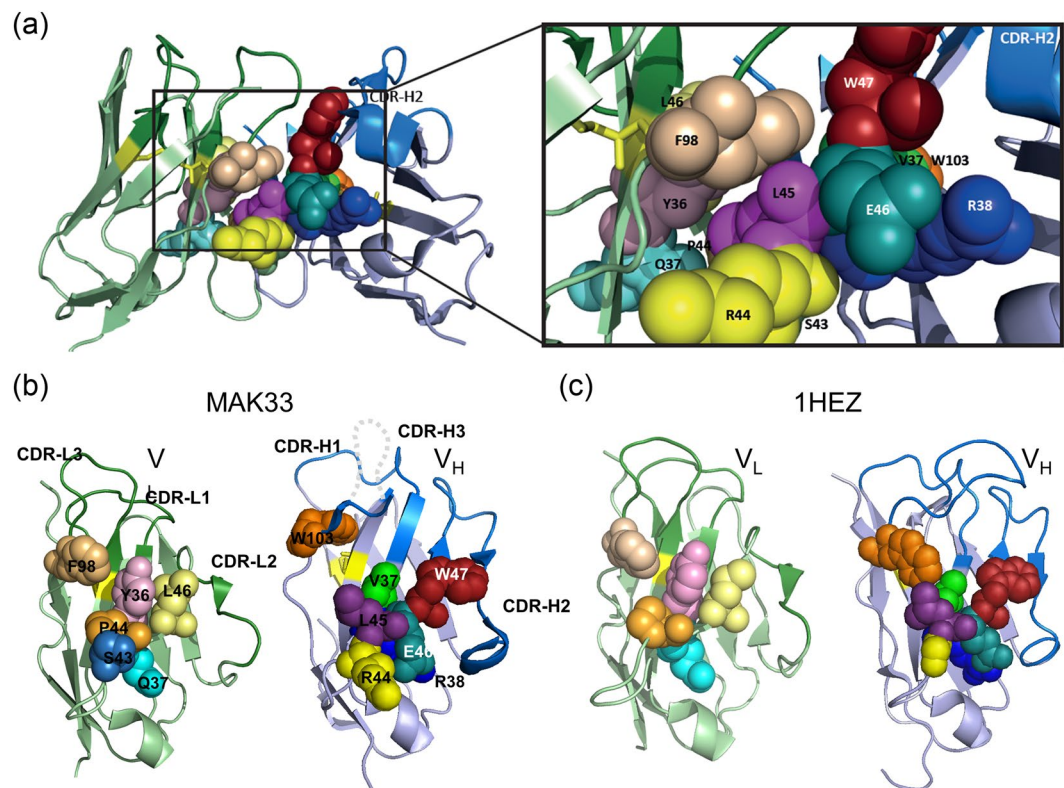


Figure 1. Conserved residues within the interface between V_L and V_H . In (a) the positions of conserved amino acids in the Fv fragment of MAK33 are shown. The V_L domain is depicted in light green with the three CDRs highlighted in dark green. V_H is shown in light blue and the CDRs are in dark blue. Conserved residues within the interface are illustrated as spheres and color-coded. On the right, the region within the rectangle is enlarged. For a better orientation CDR-H2 is indicated. The labelled residues were selected for an alanine-exchange mutational analysis. In (b) and (c) the top views of the interacting residues of the V_L (left) and V_H (right) domain of MAK33 (b) and 1HEZ (c) are shown. Structures are modified from PDB ID 1FH5 and 1HEZ. The grey dotted line indicates CDR-H3 which is not resolved in PDB ID 1FH5.

major changes in the tertiary structure of the V_H mutants. Our results support the notion that the structure of V_H is generally less tolerant than V_L against the exchange of conserved residues.

To determine whether the stability of the respective domain was affected by the mutation of conserved residues, denaturant-induced (GdmCl) unfolding transitions were performed (Figs 4 and 5, Tables 1 and 2). All transitions were fitted to a two-state model. The V_L domains showed a midpoint for GdmCl-induced unfolding of 1.2 ± 0.1 M GdmCl for MAK33 and 1.7 ± 0.1 M GdmCl for 1HEZ. Strikingly, MAK33 V_L S43A was slightly more stable than the wt against GdmCl-induced unfolding. However, the variants V_L Y36A and V_L L46A exhibited the most prominent decrease in stability compared to the two wt V_L domains (Figs 4a and 5a, Tables 1 and 2). The cooperativities of these unfolding transitions are in a comparable range for all V_L variants, from 9.1 ± 0.8 to 23.8 ± 16.2 kJ mol⁻¹ M⁻¹ (Tables 1 and 2). For the least stable V_L point mutation Y36A, two double mutations (Y36A/P44A, Y36A/S43A) were generated in the MAK33 framework to investigate the potential synergistic nature of the observed effects. Indeed when incorporating the most stable MAK33 mutation, S43A, the stability of Y36A/S43A seemed slightly increased compared to Y36A. For Y36A/P44A, where P44A alone has wt-like stability, almost no difference in stability was visible compared to Y36A (Fig. 4a, Table 1).

For the V_H wt domains, we determined a midpoint for GdmCl-induced unfolding of 0.2 ± 0.1 M GdmCl for MAK33 and 0.9 ± 0.1 M GdmCl for 1HEZ. Thus, the investigated V_H domains are less stable than their corresponding V_L domains (Figs 4 and 5, Tables 1 and 2). Moreover, the isolated 1HEZ variable domains are in general more stable than those of MAK33. Interestingly, the 1HEZ V_H domain, wt and mutants, are dimeric in the AUC size distribution analysis (Fig. S8). The only exception was the 1HEZ point mutant W47A which exhibited a 50% equilibrium between monomer and dimer fraction. The analysis of the V_H MAK33 mutants showed that all mutants and wt already start to unfold in the presence of very low GdmCl-concentrations (Fig. 4b). Thus, no distinction in stability of the MAK33 V_H mutants can be made, assuming an equally low GdmCl tolerance for all variants. For 1HEZ V_H , the V_H W47A and V_H V37A mutants showed the largest decrease in stability, followed by L45A (Table 2). The cooperativity of unfolding for the different MAK33 V_H mutations was subject to variation, from 15.3 ± 10.5 to 75.2 ± 19.4 kJ mol⁻¹ M⁻¹ suggesting that structural changes occurred²⁸. The cooperativity values for 1HEZ V_H , though, were in a similar range between 15.1 ± 4.7 to 24.1 ± 10.4 kJ mol⁻¹ M⁻¹. As there is no correlation between the cooperativity values and the structural data obtained by CD measurements (Fig. 3) this assumption cannot be confirmed. In summary, MAK33 V_H domains exhibit relatively low stabilities

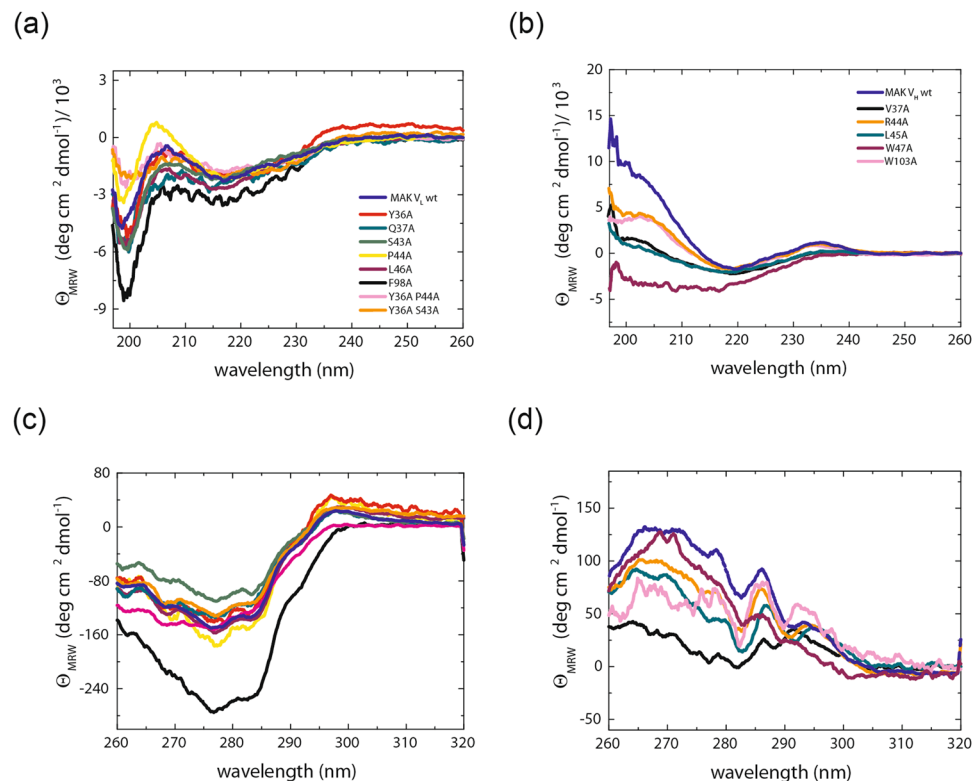


Figure 2. Secondary and tertiary structure of MAK33 V_L and V_H alanine exchange mutants. In (a) and (b), FUV-CD spectra of V_L mutants (a) and V_H mutants (b) are shown. Color code for V_L: Y36A is red, Q37A dark cyan, S43A green, P44A yellow, L45A purple, F98A black, Y36A/S43A orange, Y36A/P44A pink and V_L wild type is royal blue. Color code for V_H: V37A is black, R44A orange, L45A dark cyan, W47A purple, W103A pink and V_H wild type marine blue. In (c) and (d) the NUV-CD spectra of the V_L (c) and V_H (d) point mutants are shown. Color code for (c) and (d) is analogue to (a) and (b), respectively. For the spectra 16 accumulations each were recorded and buffer-corrected (PBS). All measurements were performed at a protein concentration of 20 μM (FUV-CD) and 50 μM (NUV-CD) in 0.5-mm (FUV) or 5-mm (NUV) quartz cuvette at 20 °C.

compared to the other variable domains analyzed. Interestingly, MAK33 V_L S43A was even slightly more stable than the wt protein. For the double mutants, an additive effect on the stability could be observed. Our analysis identified W47A and V37A for V_H and Y36A for V_L as the least stable mutations. Concerning 1 HEZ V_H, it has to be kept in mind that the wt and all mutants except for W47A (monomer dimer equilibrium) form dimers, which usually enhances its intrinsic stability. In fact, dimerization adds an additional layer of complexity. However, homo-dimerization is a typical feature of many variable antibody domains and also a number of light chains. Therefore, this additional dimension of the V_L-V_H interplay was included to this study as it competes with hetero-dimerization and thus it is interesting to determine to which extent this equilibrium influences the effects of certain mutations on the interaction of the two variable dimers in the heterodimer.

The association of V_H and V_L is particularly sensitive to mutations in V_H. Besides influencing the structure and stability of the variable domains, the conserved residues might also play a role in the association of the two domains. In previous studies, the association of the V_L and V_H domains was analyzed using scFvs or Fab fragments^{12,16,18}. This is a relatively indirect approach, as the domains were either artificially linked (scFvs) or two additional covalently linked domains were present (C_L and C_H1 in the Fab). Here, we used assays that report directly on the formation of the V_L and V_H heterodimer. For SPR measurements, the wild type domain was immobilized on the chip and different concentrations of the corresponding mutant domains were added under a constant flow. With this setup, K_Ds of 2.0 ± 0.8 to 2.9 ± 1.1 μM and 4.2 ± 1.1 to 8.2 ± 1.4 μM were determined for the MAK33 and 1HEZ wild type V_L and V_H domains, respectively (Tables 1 and 2). Thus, the variable domains interact with a similar affinity as the C_H1 and C_L²¹ domains, but have a lower affinity compared to the C_H3 homodimer^{19,20}.

Almost all MAK33 V_H mutants led to a decrease in the binding affinity, especially L45A with a 20-fold higher K_D than the wt, followed by V37A and W103A with a 2-fold change. On the 1HEZ side, only the mutants L45A and V37A exhibited a negative influence on the binding of the corresponding V_L wt domain. Generally, the MAK33 V_H mutations exerted a bigger influence on the V_H/V_L interaction with a maximum 20-fold increased K_D compared to 1HEZ with 3 to 4-fold higher K_Ds than the wt V_H domain (Table 1; Fig. 1c). The mutation of conserved residues in V_L, though, showed similar effects for the two V_L domains from different origin. In both cases a striking binding impairment with a 30-fold higher K_D was observed for the mutant P44A and a slightly decreased

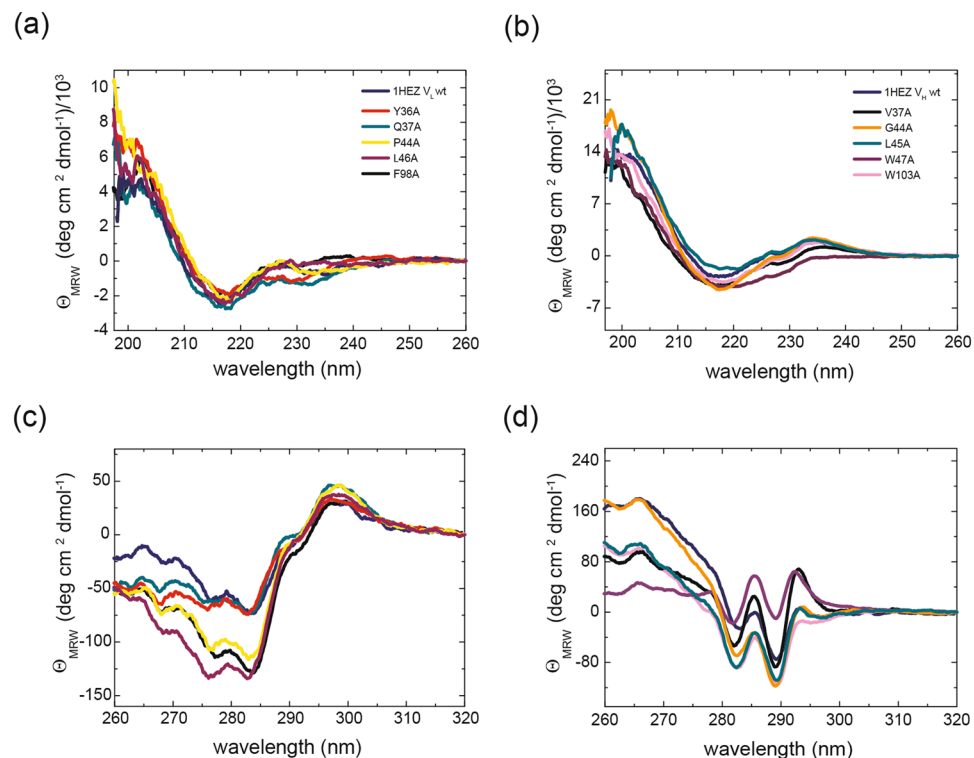


Figure 3. Secondary and tertiary structure of 1HEZ V_L and V_H alanine exchange mutants. In (a) and (b), FUV-CD spectra of V_L mutants (a) and V_H mutants (b) are shown. Color code for V_L : Y36A is red, Q37A dark cyan, P44A yellow, L45A purple, F98A black, and V_L wild type is royal blue. Color code for V_H : V37A is black, G44A orange, L45A dark cyan, W47A purple, W103A pink and V_H wild type marine blue. In (c) and (d) the NUV-CD spectra of the V_L (c) and V_H (d) point mutants are shown. Color code for (c) and (d) is analogue to (a) and (b), respectively. For the spectra 16 accumulations each were recorded and buffer-corrected (PBS). All measurements were performed at a protein concentration of 20 μ M (FUV-CD) and 50 μ M (NUV-CD) in 0.5-mm (FUV) or 5-mm (NUV) quartz cuvette at 20 $^{\circ}$ C.

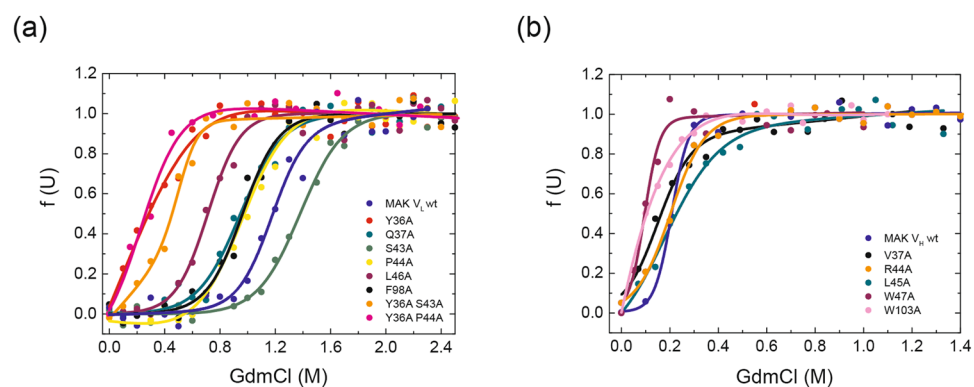


Figure 4. Influence of conserved residues on MAK33 protein stability. The stability of the V_H and V_L alanine point mutants towards GdmCl-induced denaturation (a) and (b). (a) V_L mutants: Y36A is red, Q37A dark cyan, S43A green, P44A yellow, L46A purple, F98A black, Y36A/S43A orange, Y36A/P44A pink and the wild type marine blue. (b) V_H mutants: V37A is black, R44A orange, L45A dark cyan, W47A purple, W103A pink and the wild type royal blue.

V_H/V_L interaction with a 3 fold increase in K_D for the mutation F98A (Fig. 6; Tables 1 and 2). Concerning the two MAK33 V_L double mutants (Y36A/P44A, Y36A/S43A), which were generated to investigate whether the effects observed for single mutants are additive. Y36A/S43A, which is the combination of two point mutations with no change in K_D , leads to the expected wt-like K_D . But the point mutation P44A with the highest K_D for the association with V_H ($58.0 \pm 10.4 \mu$ M), exhibited improved binding when combined with Y36A, suggesting compensatory effects (Table 1).

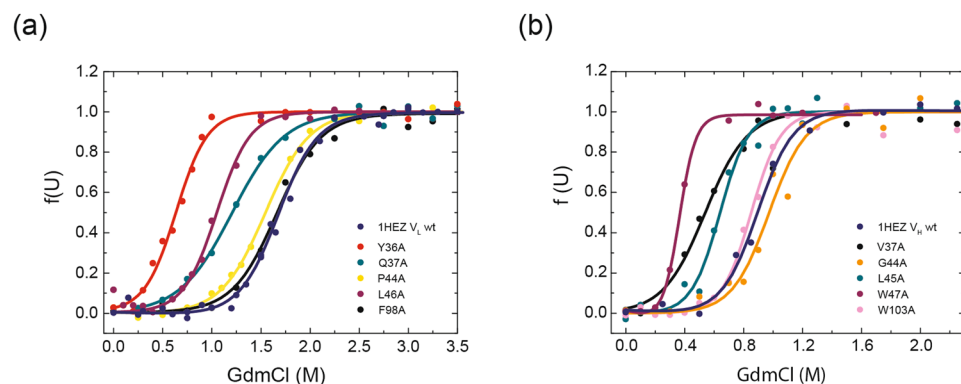


Figure 5. Influence of conserved residues on 1HEZ protein stability. The stability of the V_H and V_L alanine point mutants towards GdmCl-induced denaturation (a) and (b). (a) V_L mutants: Y36A is red, Q37A dark cyan, P44A yellow, L46A purple, F98A black and the wild type marine blue. (b) V_H mutants: V37A is black, G44A orange, L45A dark cyan, W47A purple, W103A pink and the wild type royal blue.

Domain	Mutation	Association K_D [μ M] SPR	Stability GdmCl $D_{1/2}$ [M]	Cooperativity [kJ mol ⁻¹ M ⁻¹]	Antigen binding app. K_D [μ M]
V_L	wild type	$2.0 \pm 0.8^*$	1.2 ± 0.1	17.1 ± 6.7	0.9 ± 0.1
	Y36A	$1.5 \pm 0.2^*$	0.30 ± 0.1	16.1 ± 2.3	0.6 ± 0.1
	Q37A	$1.3 \pm 0.7^*$	0.9 ± 0.5	16.0 ± 5.2	1.5 ± 0.5
	S43A	$1.6 \pm 0.4^*$	1.4 ± 1.0	15.8 ± 8.6	1.4 ± 0.1
	P44A	$58.0 \pm 10.4^*$	1.0 ± 0.2	15.0 ± 3.2	10.5 ± 0.4
	L46A	$0.4 \pm 0.1^*$	0.7 ± 0.2	21.2 ± 3.5	4.6 ± 1.2
	F98A	$5.8 \pm 0.1^*$	1.0 ± 0.2	19.0 ± 2.8	8.3 ± 0.5
	Y36A/P44A	$22.6 \pm 4.4^*$	0.2 ± 0.1	23.8 ± 16.2	4.5 ± 0.7
	Y36A/S43A	$1.5 \pm 0.5^*$	0.4 ± 0.2	17.1 ± 6.0	0.7 ± 0.1
	MAK_1DH5	$3.0 \pm 0.3^*$	1.4 ± 0.2	15.6 ± 4.8	1.2 ± 0.3
	1DH5_MAK	$42.3 \pm 9.8^*$	1.3 ± 0.2	16.5 ± 1.8	NB
	1DH5_DH5	$76.2 \pm 12.8^*$	2.4 ± 0.2	17.1 ± 6.1	NB
V_H	wild type	2.9 ± 1.1	0.2 ± 0.1	68.6 ± 23.3	1.4 ± 0.1
	V37A	6.8 ± 0.7	0.2 ± 0.1	35.3 ± 6.7	3.0 ± 0.4
	R44A	4.0 ± 2.0	0.2 ± 0.1	34.3 ± 3.7	1.5 ± 0.4
	L45A	51.0 ± 5.8	0.2 ± 0.1	21.5 ± 8.5	16.1 ± 0.9
	W47A	4.5 ± 1.3	0.1 ± 0.1	75.2 ± 19.4	NB
	W103A	6.9 ± 1.0	0.2 ± 0.1	15.3 ± 10.5	3.7 ± 0.1
	MAK_1DHU	2.5 ± 0.3	1.0 ± 0.1	4.5 ± 0.4	> 50
	1DHU_MAK	1.3 ± 0.1	1.3 ± 0.2	10.7 ± 0.9	NB
	1DHU_1DHU	± 0.5	2.5 ± 0.2	6.3 ± 3.6	NB
	MAK_MAK CDR3 1DHU	1.8 ± 0.2	0.4 ± 0.1	7.9 ± 1.4	14.7 ± 0.5

Table 1. Characteristics of IgG (MAK33) V_H and V_L mutants. K_D values were determined by SPR experiments. Titrations of the mutant domains in a concentration range between 0–200 μ M were measured with the corresponding wild type domain, which was immobilized on a CM5 chip. Experiments were performed at 20 °C; asterisks indicate that the temperature setting for the measurement was reduced to 15 °C due to low thermal stability. Stabilities against the chemical denaturation (GdmCl) of the V_L and V_H mutants. Even though most of the GdmCl-induced unfolding transitions were not reversible, data were evaluated according to a two-state equilibrium unfolding model to derive the midpoint of transitions ($D_{1/2}$), as well as the cooperativity parameter (m), for a qualitative comparison of the data. Functionality of the isolated variable domains determined by ELISA with the antigen creatine kinase. Detection of the V_H and V_L wild type domains was possible via an introduced FLAG-tag at the C-termini. As a reference, the signals of the wild type domains vs. FLAG-tagged domains after 25 minutes of incubation were chosen. Samples were corrected for the signal of the single FLAG-tagged variable domains. Apparent K_D values were obtained by a Boltzmann fit. Experiments were performed at 20 °C. NB indicates that no binding could be observed. Values for wild type proteins are highlighted in bold.

Domain	Mutation	Association K_D [μM] SPR	Stability GdmCl $D_{1/2}$ [M]	Cooperativity [$\text{kJ mol}^{-1} \text{M}^{-1}$]
V_L	wild type	8.2 ± 1.4	1.7 ± 0.1	12.1 ± 0.2
	Y36A	3.8 ± 0.9	0.6 ± 0.1	15.2 ± 1.4
	Q37A	10.0 ± 0.8	1.2 ± 0.4	9.1 ± 0.8
	P44A	189.6 ± 49.2	1.5 ± 0.1	10.7 ± 0.9
	L46A	10.6 ± 0.5	1.1 ± 0.1	14.3 ± 1.4
	F98A	20.3 ± 2.8	1.7 ± 0.1	10.8 ± 1.8
V_H	wild type	4.2 ± 1.1	0.9 ± 0.1	18.9 ± 3.9
	V37A	13.9 ± 0.4	0.5 ± 0.1	16.8 ± 9.1
	G44A	4.0 ± 0.1	1.0 ± 0.1	19.4 ± 8.7
	L45A	19.0 ± 2.0	0.6 ± 0.1	25.4 ± 8.4
	W47A	3.7 ± 1.0	0.4 ± 0.1	15.1 ± 4.7
	W103A	4.2 ± 1.3	0.9 ± 0.1	24.1 ± 10.4

Table 2. Characteristics of IgM (1HEZ) V_H and V_L mutants. K_D values were determined by SPR experiments. Titrations of the mutant domains in a concentration range between 0–200 μM were measured with the corresponding wild type domain, which was immobilized on a CM5 chip. Experiments were performed at 20 °C; Stabilities against the chemical denaturation (GdmCl) of the V_L and V_H mutants. Even though most of the GdmCl-induced unfolding transitions were not reversible, data were evaluated according to a two-state equilibrium unfolding model to derive the midpoint of transitions ($D_{1/2}$), as well as the cooperativity parameter (m), for a qualitative comparison of the data. Values for wild type proteins are highlighted in bold.

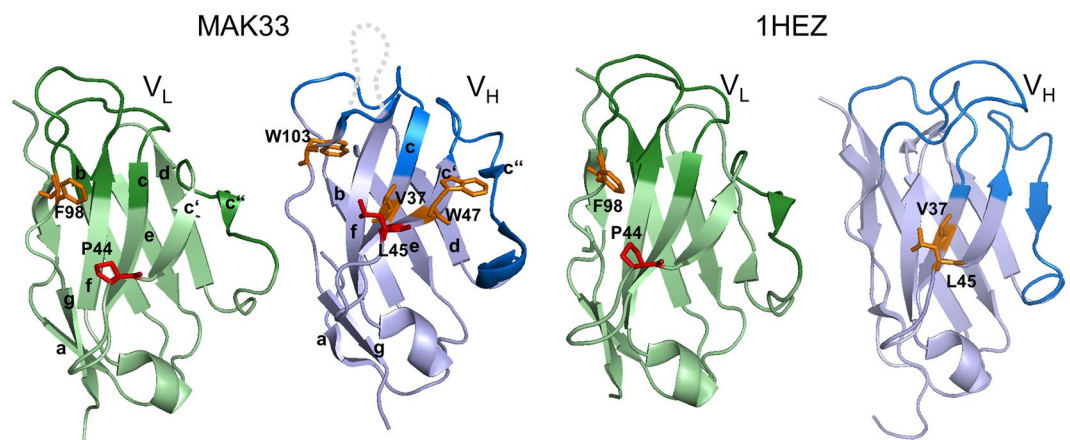


Figure 6. Conserved interface residues with importance for V_H/V_L association. The top-views for MAK33 (left) and 1HEZ (right) V_L and V_H are shown and the residues with strong influence on association highlighted in red, those with mild influence highlighted in orange. For MAK33 the numbering of the strands (a, b, c, c', c'', d, e, f, g) is also shown. Structures are modified from PDB ID 1FH5 and 1HEZ. The grey dotted line indicates CDR-H3 which is not resolved in PDB ID 1FH5.

In conclusion, P44 exhibits a strong negative impact on the association of the V_H and V_L domain, while F98 exerts only a mild influence. However, for the two different V_H domains, there is no complete overlap concerning the observed effects. While three of five conserved MAK33 V_H residues (V37, L45, W103) showed a clear influence on the association, only the mutants V37A and L45A were affected in 1HEZ V_H (Fig. 6; Tables 1 and 2). Nevertheless, a negative effect of mutant 1HEZ W47A which exists as monomer and dimer cannot be excluded. Despite the wt-like K_D , an impaired association might be compensated because in contrast to the 1HEZ wt and the other 1HEZ mutants, which are dimeric, this variant exhibited a monomer-dimer equilibrium and heterodimer formation is probably favored over the V_H homodimer. Considering the effects of the mutations on the structure and stability of MAK33 V_H , this strong effect on association for the majority of the mutants might also be caused by the low stability and changes in secondary and tertiary structure of the MAK33 V_H variants as already described. As most of the conserved interface residues possess one or several interaction partners within the other variable domain (Table S1), a similar effect on association is expected as soon as one partner is mutated. This holds true for the interaction partners P44 (V_L) and L45 (V_H) as well as F98A (V_L) and V37A (V_H) where our analysis identified the most drastic effect on association for the point mutations P44A and L45A. Consequently, those V_H point mutations in MAK33 and 1HEZ with an apparent effect on association but lacking a corresponding effect for the mutant affecting the partner residue in V_L might be due to changes in secondary (Figs 2b and 3b) and tertiary (Figs 2d and 3d) structure of V_H . For 1HEZ also V_H domain stability seems to impact the association

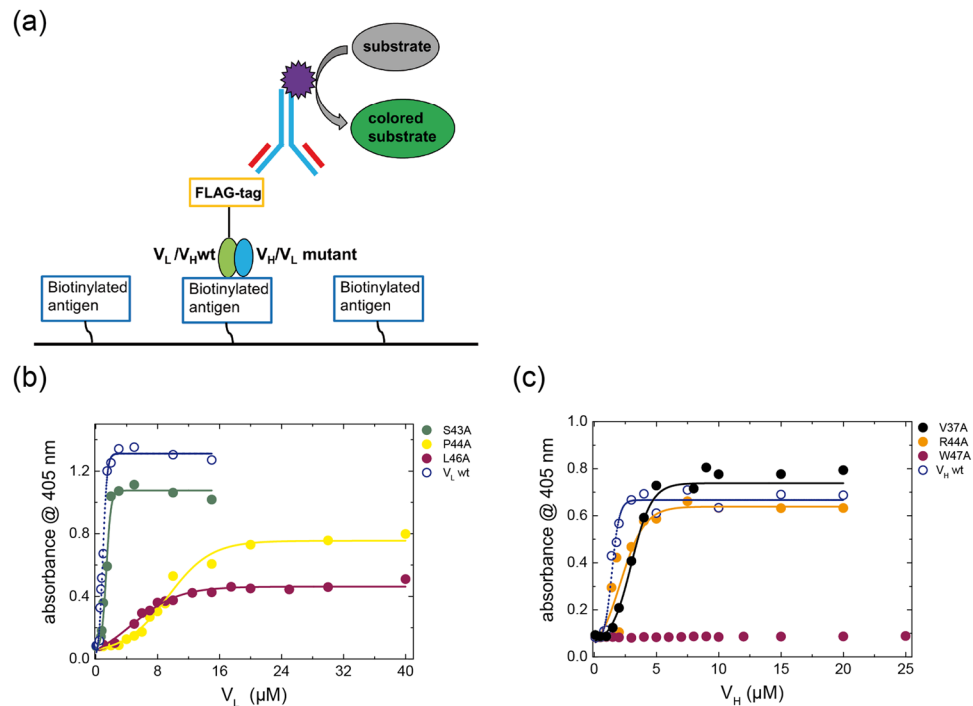


Figure 7. Influence of conserved residues on antigen binding. To determine the functionality of the mutants an ELISA with Fv fragments was performed. In (a) the ELISA setup is shown. The biotinylated antigen (creatine kinase) was immobilized on streptavidin coated 96 well plates. Subsequently, 1:1 mixtures of the FLAG-tagged wild type domain and a corresponding mutant were added. Binding was detected via a peroxidase-conjugated anti-FLAG antibody. The peroxidase mediated conversion of the substrate ABTS was measured colorimetrically at 405 nm. For V_L (a) and V_H (b), titrations of Fvs containing either the wild type, a representative of an inactive, a low active or a mutant with similar activity as the wild type is shown. The absorption at 405 nm at the signal maximum was corrected for the signal of the variable labelled domains alone. Experiments were performed at 20 °C.

since the two mutants compromised in association ($V37A$, $L45A$) show, besides 1HEZ $W47A$ which exhibits a monomer-dimer equilibrium, a significantly decreased stability compared to 1HEZ wt (Table 2, Fig. 5b).

Antigen binding of V_H and V_L is influenced by conserved interface residues. To test how the mutation of conserved amino acids influences antigen recognition by the Fv fragment, we set up an ELISA (Fig. 7a) for the MAK33 Fv-fragment and human creatine kinase as its antigen. With the corresponding set up for the 1HEZ Fv fragment with an IgG Fc part as the antigen^{29,30}, no binding was detectable. Hence, the influence of the conserved residues towards antigen binding could only be determined for the MAK33 Fv fragment. To this end, the respective wt domain was produced with a FLAG-tag at the C-terminus for detection with an anti-FLAG antibody. The tag does not negatively influence stability, folding and the interaction between the variable domains (data not shown). In this ELISA, a concentration-dependent increase in signal is only observed when V_L and V_H are combined (Fig. 7b and c). As the assay involves several protein interactions, only conclusions on an apparent K_D seem reasonable (Table 1). For the wt Fv with either V_L or V_H tagged with FLAG similar K_{Dapp} s were determined (data not shown). Antigen binding was found to be abolished for V_H $W47A$. Apart from that, the V_H mutation $L45A$ led to the strongest reduction in the affinity for the antigen with a 12-fold increased K_D compared to wt, while $V37A$ and $W103A$ exhibited a less pronounced decrease with a 2-fold higher K_D (Table 1). In general, for the V_H point mutations a low antigen binding activity correlates with a high K_D for association of the Fv-fragment (see Table 1). However, $W47A$, showed only a small decrease in the K_D but no binding to the antigen.

For the V_L domain, there is in most cases a correlation between the K_D for association and antigen binding, except for $L46A$ which shows a wt-like association to V_H but an impaired antigen binding. The V_L mutations $L46A$, $F98A$, $P44A$, as well as the double mutant $Y36A/P44A$ have a negative influence on antigen binding (Table 1), with a 5–12-fold increased K_D . As already observed for V_H/V_L association, $P44A$ exhibited the weakest binding with a 12-fold higher K_D for the antigen. Interestingly, the double mutant $Y36A/P44A$ showed only 50% of the impairment of $P44A$ alone, so there must be a compensating effect of the $Y36A$ mutation. This coincides with the data for the V_H/V_L association (Table 1), where $P44A$ exhibited the worst K_D while the double point mutation $Y36A/P44A$ showed a 2-fold higher affinity than $P44A$ alone. Surprisingly, V_L $L46A$ and $F98A$, which exhibited an affinity for the V_H domain similar to the wt, were defective in antigen binding (Table 1). Consequently, the analysis of V_L mutants supports the assumption that the affinity between V_H and V_L is not necessarily correlated with the ability to bind the antigen.

The CDR regions affect domain structure, stability and association. The CDRs of the antibody variable domains are elements of natural variations. How variations in these elements affect their association, structure and stability is therefore of special interest to obtain a comprehensive picture of the factors shaping the Fv-fragment. To address this question, we switched CDRs between MAK33 and unrelated variable domains. We chose human variable domain consensus sequences³¹ with CDRs of similar length. For V_L , we selected the 1DH5 domain, and for V_H 1DHU. These human variable domains represent a class of variable domains with a highly stable structure³².

The K_D s for the association of grafted variants with wt domains were determined by SPR. Due to the fact that the grafting constructs represent consensus sequences, no natural V_H or V_L partner domains exist. Hence, association was always determined with the corresponding MAK33 wt domain. For the association of the MAK33 V_H domain containing the 1DHU CDRs (1DHU_MAK V_H) with MAK33 V_L , a K_D of $1.3 \pm 0.1 \mu\text{M}$ was obtained. This corresponds to the value determined for the MAK33 wt domains (Table 1). Wt 1DHU V_H was a very good binding partner for MAK33 V_L , with a K_D of $1.0 \pm 0.5 \mu\text{M}$. A similar value was measured when grafting the MAK33 CDRs into the 1DHU framework (MAK_1DHU V_H), the chimera bound to MAK33 V_L with a K_D of $2.5 \pm 0.3 \mu\text{M}$. These observations lead to the conclusion that for MAK33 V_L , the MAK33 and 1DHU V_H domains are both suitable binding partners. But when grafting the MAK33 CDRs on the 1DHU framework, the association is slightly impaired, demonstrating that the CDRs can exhibit a marked influence. Concerning the V_L grafting constructs, the observations are different: the 1DH5 V_L domain binds to MAK V_H with a roughly 30-fold lower affinity than the MAK33 V_L domain. However, the 1DH5 grafting construct containing the MAK33 CDRs (MAK33_1DH5) shows an affinity similar to the MAK33 wt V_L domain. In contrast, grafting of the 1DH5 CDRs onto the V_L -MAK33 framework (1DH5_MAK33) resulted in strong binding impairment with a K_D of $42.3 \pm 9.8 \mu\text{M}$, which is only slightly lower than the 1DH5 wt K_D of $76.2 \pm 12.8 \mu\text{M}$. So for the V_L domain, the CDRs and not the framework are the determining factor for the affinity towards V_H wt.

The FUV and NUV spectra of the grafting mutants gave a similar picture. The CDR exchange (1DH5_MAK) showed a FUV-CD spectrum similar to 1DH5 V_L while the spectrum of MAK_1DH5 was different from both wts (Fig. 8b). So already on the secondary structure level the CDRs seem to be structurally important. The NUV-CD spectra exhibit a similar pattern (Fig. 8d) but here the difference in the number of aromatic amino acids, predominantly concerning the CDRs, could also play a role.

For the V_H domain, both grafting mutants showed similar FUV-CD spectra (Fig. 8c). The same was observed for the NUV-CD spectra of the grafting mutants compared to the wt domains; all spectra were similar in shape with variations in amplitude (Fig. 8e). As for V_L the observed NUV deviations might also be due to the different numbers of aromatic amino acids. For both domains the CDRs of MAK33 and 1DHU/1DH5 differ by one tryptophan and several tyrosines. Generally for MAK V_H , the CDR exchange does not exhibit the same impact as for the V_L domain.

GdmCl-induced transitions of the different V_L domains showed that, in comparison, the MAK33 V_L domain is least stable against chemical denaturation (Fig. 9a). As expected [19], the most stable domain was 1DH5 with a $D_{1/2}$ of $2.4 \pm 0.2 \text{M}$. The grafted mutants showed stabilities in-between MAK33 V_L and 1DH5. Interestingly, the stability of the MAK33 V_L framework was increased when the CDRs were exchanged against the CDRs of 1DH5. On the V_H side, 1DHU showed the highest stability. The grafting mutants were again in-between (Fig. 9b). When the CDRs of 1DHU V_H were transplanted into MAK33, the chemical stability increased, however the transition was less cooperative compared to MAK33 V_H (Fig. 9b). Exchanging only CDR-H3 in MAK33 V_H led to a slight increase in stability compared to the wt (Fig. 9b). The exchange of the CDRs of 1DHU against MAK33 CDRs decreased its stability. In summary, the analysis of the chemical stabilities allowed the same conclusion for V_L and V_H : the CDRs influence the stability of the variable domains decisively.

When antigen binding of the chimera was analyzed, the exchange of the MAK33 CDRs with that of the human consensus sequences led to a complete abolishment in both cases, as expected. For the grafting of MAK33 CDRs on the human variable domains, an interesting picture emerged. V_L (MAK_1DH5) has a slightly higher K_D of $0.3 \mu\text{M}$ for creatine kinase whereas there is almost no antigen binding detectable for the V_H grafting domain MAK_1DHU ($>50 \mu\text{M}$) (Table 1). So for binding of the V_H domain to the antigen not only the CDRs represent a determining factor.

As CDR-H3 (i.e. the third CDR of V_H) is the most flexible of these six regions concerning length and amino acid sequence², additionally a mutant carrying a swapped CDR-H3 loop was analyzed. Concerning the V_H/V_L association the CDR-H3 mutation shows a wt-like K_D of $1.8 \pm 0.2 \mu\text{M}$. Interestingly, this construct, MAK CDR-H3 1DHU V_H showed a strongly impaired binding to the antigen with a 10-fold increased K_D (Table 1). This demonstrates the importance of CDR-H3 for the antigen binding of MAK33.

Molecular Dynamics Simulations reveal mutation-induced structural alterations. For a subset of V_H and V_L mutations explicit solvent Molecular Dynamics (MD) simulations were performed for the heterodimer and for the individual domains. The set of simulations included point mutations that are associated with a significant reduction in V_H/V_L association (V_L P44A, Y36A/P44A; V_H V37A, L45A and W103A) and, as a control, also substitutions that showed only modest effects on complex affinity (V_L Y36A, S43A; V_H R44A, W47A). Simulations were started from the geometry of the wt structure (pdb-entry: 1FH5, see Methods for details). On the time scale of the simulations, none of the V_H/V_L complexes dissociated and the root-mean-square deviation (RMSD) of the complexes did not exceed 0.3 nm from the start structure (Fig. S1) and interface contacts as in the WT complex (Table S1). However, some of the mutations (e.g. V_L P44A, V_H R44A, V_H L45A and V_H W47A, see Fig. S1) resulted in overall larger final RMSDs compared to the wt indicating mutation-induced structural alterations and increased conformational freedom. Interestingly, simulations of the mutated proteins in the isolated state showed no significant differences in the RMSD (Fig. S2) except for V_H W103A (see below). The structural distortion of the V_L/V_H complexes due to some of the mutations is also reflected in overall larger root mean

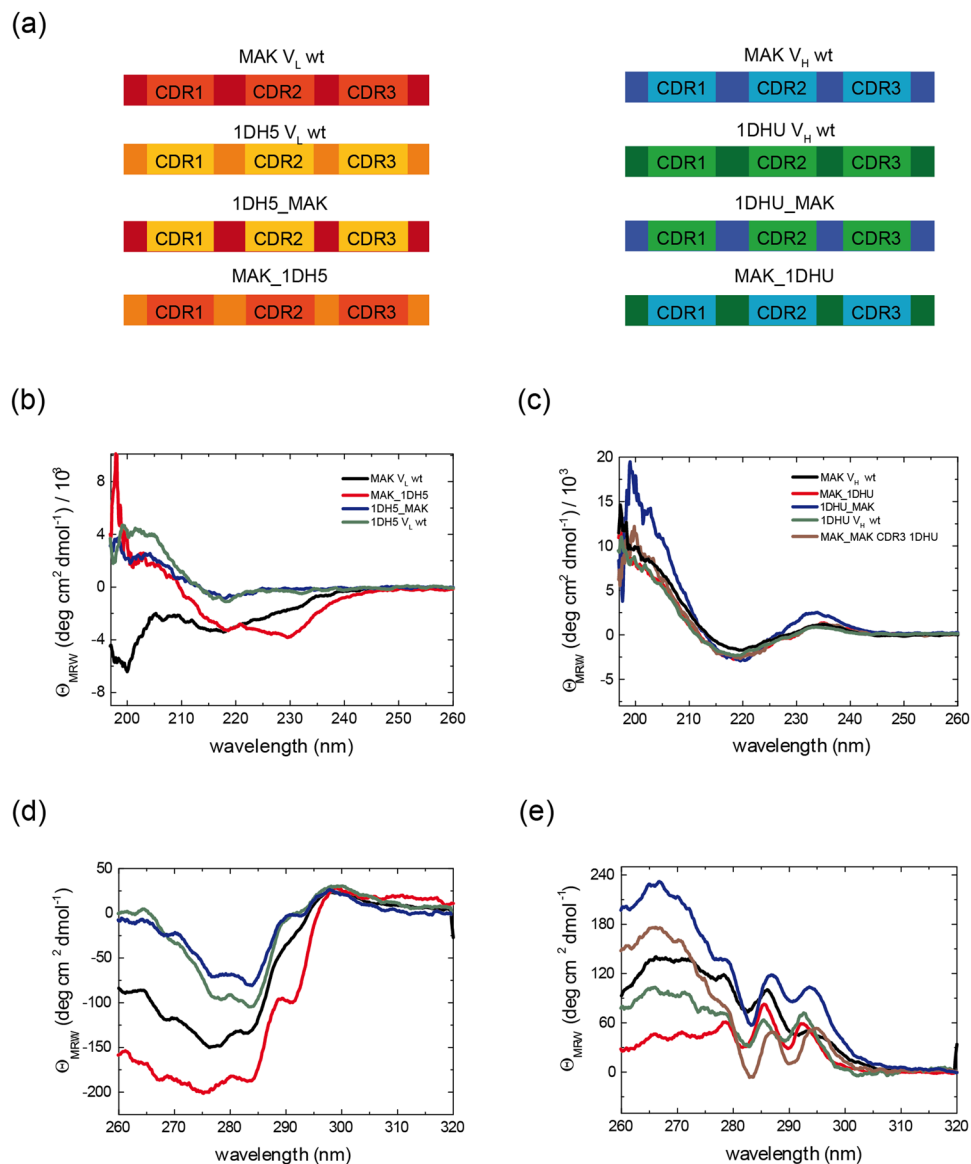


Figure 8. CDR-grafted mutants. (a) Schematic representation of CDR-grafted mutants. The framework for V_L grafting mutants (left) was 1DH5, which is the most stable human consensus sequence V_L domain (from HuCAL³¹). The framework for V_H grafting mutants (right) was 1DHU, which is the most stable human consensus sequence for V_H domains. In (b and c), FUV-CD spectra of V_L (left) and V_H (right) CDR-grafted mutants are shown. The color code for V_L is: MAK_1DH5 in red, 1DH5_MAK in blue; wild type 1DH5 is shown in green MAK33 in black. The color code for V_H is: MAK_1DHU in red, 1DHU_MAK in blue, wild type 1DHU in green, MAK33 CDR3 1DHU in brown. In (d and e) NUV-CD spectra of the V_L (left) and V_H (right) point mutants are shown.

square fluctuations of heavy atoms (RMSF) with respect to the mean structure. For the isolated mutated protein partners no significant difference to the wt was observed (Fig. S3), again with the exception of W103A.

However, in the complex the mutations with reduced binding affinity (e.g. V_L P44A, V_H R44A, V37A, L45A, W47A) showed increased fluctuations in regions at and near the binding interface but also in loops involved in antigen binding (Figs S4 and S5). Interestingly, especially for the substitutions that caused the largest drop in affinity between the V_L and the V_H domains an increased solvation at the interface (diffusion of water molecules into the space created by the introduction of a small Ala residue) was observed (illustrated in Fig. 10, Table 3). Especially for V_L P44A, V_H L45A and V_H W47A, the average number of water molecules increased near the mutation site (Table 3). The mutation V_H W103A resulted in significant changes of the backbone conformation around the mutation site, specifically the loop formed by residues 93–107 (Fig. 10), explaining the larger RMSD and RMSF observed for the isolated V_H domain in this case (Figs S2 and S3).

Besides of the effect of the mutations on the binding interface, it is interesting to investigate the changes in mobility of the CDR loops involved in antigen binding. We compared the fluctuation pattern observed in the complexes and in the individual (isolated) V_L and V_H partner domains. Even in case of the wt, the RMSF pattern

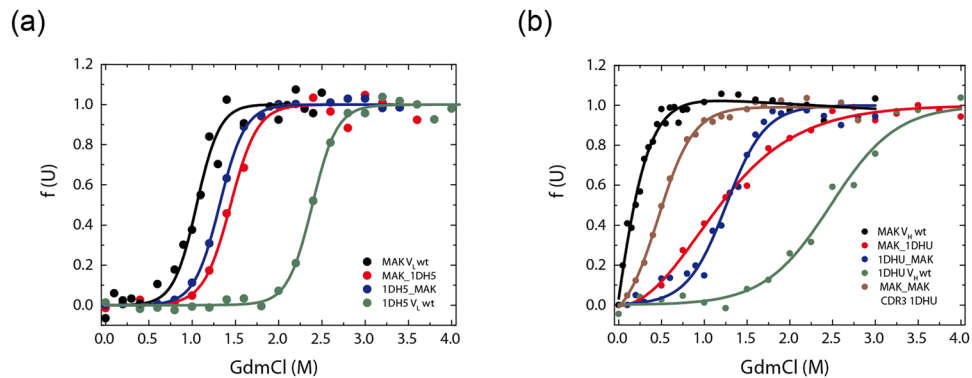


Figure 9. The influence of the CDRs on the stability of the variable domains. To assess the stability of CDR-grafted mutants, GdmCl-induced unfolding experiments were performed. Data for denaturant-induced transitions for V_L (a) and V_H (b) variants are shown. Data were evaluated according to a two-state unfolding model to obtain midpoints and cooperativity parameters of the transitions. Measurements were performed at 20 °C at a protein concentration of 1 μ M.

changes significantly in several regions that include regions involved directly in binding the partner domain but also regions involved in antigen binding (Figs S3, S4 and S5). For example, the antigen binding loop V_H : 93–107 shows large fluctuations in the absence of the V_L binding partner (Fig. S3) which drop significantly in the complex (Fig. S4). Hence, complex formation of the V_L and V_H domains appears to lock some of the antigen binding loops into distinct conformations. This effect is qualitatively also observed for several mutants, however, for some loop regions the reduction of CDR loop mobility upon binding is smaller compared to the wt. This is especially seen for the V_H : 93–107 region.

In addition to simulations of point mutations, we also studied a subset of the loop exchange constructs (MAK_1DH5, 1DH5_MAK, MAK_1DHU and 1DHU_MAK). In experiments these variants affected the binding affinity between V_H and V_L domains much less than some of the interface point mutations (see above). During simulations on the time scale of 100 ns, these variants did not show significant differences of the calculated RMSF compared to the wt (Fig. S6).

Discussion

The relationship between structure, stability and binding affinity of V_H and V_L is still unclear. This is an important aspect for understanding antibody architecture both as the basis of our immune system and also in the context of the engineering of antibodies for therapeutic purposes. In this context, it was found that in mutants an increase in affinity is often accompanied by a decrease in stability and *vice versa* - and these consequences are difficult to predict^{33–39}. In our study, we analyzed the association of the variable domains for the first time directly. This allowed us to specifically define the contribution of framework and CDR mutations on the interaction of V_H and V_L . To determine whether the observed effects for the conserved residues can be transferred between different variable domains, the analysis was performed with two distinct pairs of variable domains from a murine IgG1 (MAK33) and a human IgM (1HEZ) antibody. For the V_L wt domains, K_{DS} of $2.0 \pm 0.8 \mu$ M for the murine IgG1 and $8.2 \pm 1.4 \mu$ M for the human IgM were determined. K_{DS} for wild type V_H domains were found to be $2.9 \pm 1.1 \mu$ M and $4.2 \pm 1.1 \mu$ M for IgG1 and IgM, respectively. Consequently, the two domains interact with a similar efficiency as C_H1 and C_L , but not as efficient as the C_H3 homodimer^{19,21}. This relatively weak interaction (in the absence of the C_H1 and C_L domains) necessitates the covalent linkage of Fv fragments via a peptide and thus creating a pseudo-monomeric fusion protein (scFv) used as a therapeutic agent^{40,41}. Simulations of scFvs have shown that the stability of the interface between the two variable domains plays a critical role for the overall stability of an antibody (or fragment) as dissociation precedes unfolding⁴².

Of special interest is the nature of the domain interface. It has to support the association of the two domains but also allow accommodating different CDRs and their repositioning in the context of antigen binding. Thus, an individual interface residue may be involved in one or more of these processes: 1) formation of the immunoglobulin fold, 2) domain stability, 3) interaction between the variable domains, or 4) antigen binding. Our alanine screen of conserved residues allowed us to address each of these possibilities and differentiate between them. Consistent with the results of *in silico* analysis which showed that only very few residues (<10) are important for adopting an immunoglobulin fold⁴³, the alanine mutants in the V_L domain had only a relatively low influence on its secondary or tertiary structure compared to the V_H domain. In contrast, V_H structure is highly sensitive to the exchange of conserved residues in the interface. Two V_H interface residues, E46 and R38, were identified to be essential for the folding of both, the MAK33 and 1HEZ V_H domain. In the covariation analysis there was a very high ϕ -value for these two residues. In the crystal structure, both residues form a salt bridge⁹, R38 is buried and E46 does not interact with other V_H interface residues, but might electrostatically affect V_L binding⁹. Camelids and cartilaginous fish possess naturally occurring heavy chain antibodies lacking the light chain^{22,44,45}. Interestingly a sequence alignment of MAK33 V_H with the variable domain (V_{HH}) of the camelid VHH (PDB entry 2XT1) and the variable domain of “monomeric” shark IgNAR (V_{NAR}) (PDB entry 2I24) shows a match for the residues E46 and R38. Concerning the shark IgNAR these two residues are actually the only ones from the conserved network investigated in this study that are present at corresponding positions of this otherwise highly

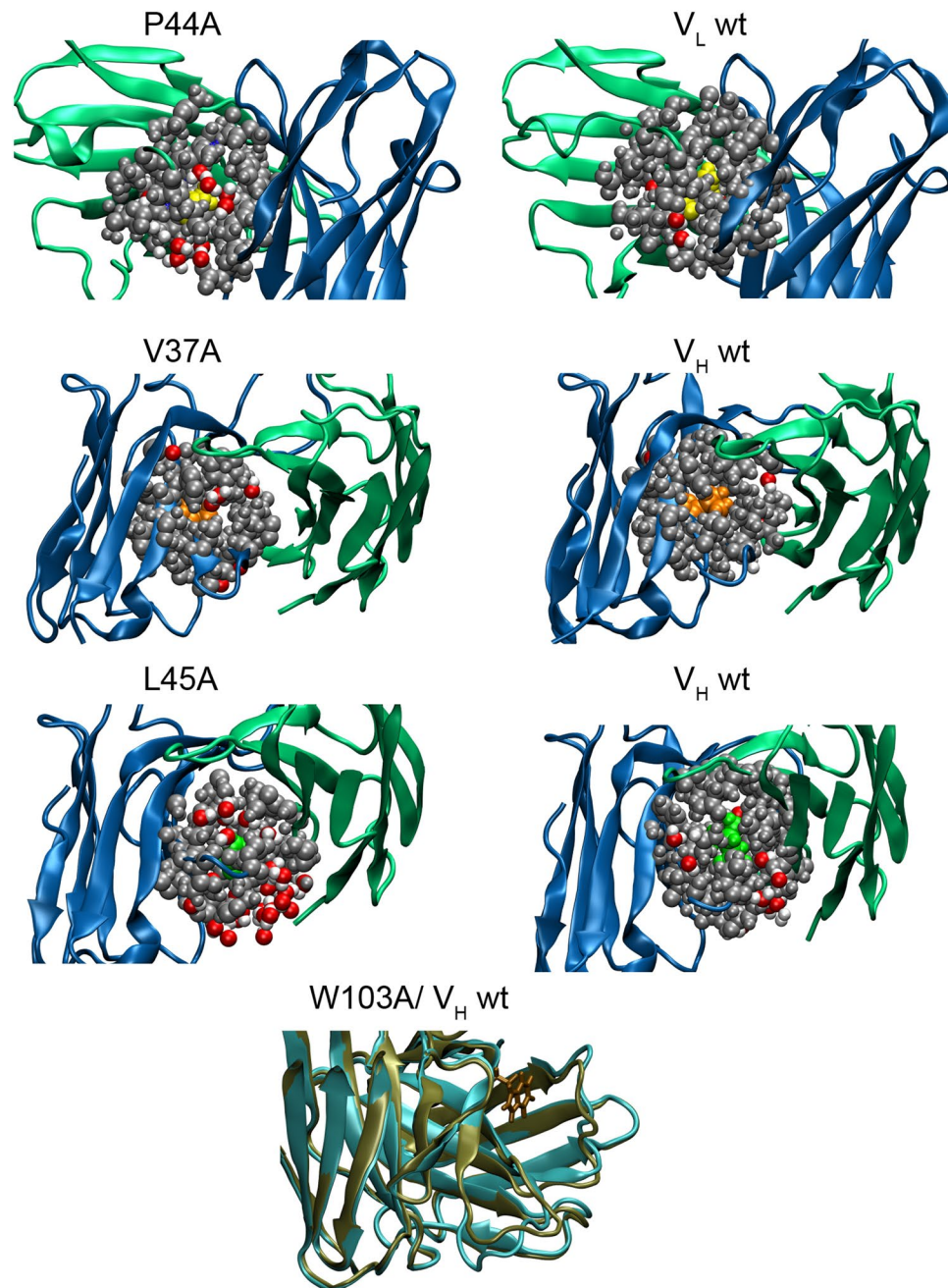


Figure 10. Interface hydration near mutation as determined by MD simulations. The snapshots illustrate the water distribution within 7 Å of the mutation site. Protein chains are shown as cartoon (V_L : green, V_H : blue). Atoms within 7 Å of the mutation site are indicated as van der Waals spheres using atom color code for water molecules (and grey for protein atoms). For comparison the same regions are also shown for the wild type case (left panels). In case of the V_H W103 mutation the average shift in backbone structure is illustrated (green cartoon) and compared to the wild type case (light blue).

divergent sequence. For camelid V_{HH} , a mutation of the hydrophobic V_H/V_L interface residues (including the tetrad: V37, G44, L45, W47) in favor of hydrophilic ones was discovered^{9,45}. The increase of hydrophilic residues in the framework also holds true for the V_{NAR} which probably evolved from a cell surface receptor^{44,46}. Both antibodies show high biophysical stability and their distinct structural patterns have by now been successfully applied to generate monomeric human V_H domains^{47–49}.

Mutations of conserved V_L residues had predominantly less impact than observed for V_H residues. As expected, MAK33 V_L S43A and MAK33 V_H R44A behaved similar to the wt. These are the only amino acid positions of the two antibodies studied which do not exhibit the conserved amino acid network identified by Wang and coworkers. In their covariation analysis an alanine is the conserved amino acid at V_L position 43 and glycine is conserved at position 44 in V_H , as can be seen for the 1HEZ variable domains.

Mutation	<water> in mutation	<water> in wild type
V _L P44A	6.8 (2)	4.3 (1)
V _L S43A	14 (4)	15 (3)
V _L Y36A	1.1 (0.3)	0.8 (0.3)
V _H V37A	1.1 (0.4)	0.9 (0.3)
V _H R44A	16 (3)	15 (3)
V _H L45A	9 (3)	5 (2)
V _H W47A	8 (3)	5 (1)
V _H W103	14 (3)	15 (3)

Table 3. Average number of water molecules around a point mutation (MAK33). Water molecules within 7 Å of the C-beta atom of a mutated amino acid were determined during a 100 ns MD simulation. The standard deviation of the number of waters is given in parenthesis.

For both V_L domains, the proline residue at position 44 showed the most prominent effects with an impaired V_H/V_L association as well as impaired antigen binding, followed by F98A. In contrast, P44 did not exhibit significant impact on V_L stability. Generally, a change in stability did not necessarily coincide with a change in functionality. Surprisingly, MAK33 V_L L46A exhibited a higher affinity for the V_H domain than the wt but it is clearly less efficient in binding the antigen (Table 1). The MAK33 double point mutations Y36A/P44A and Y36A/S43A indicate that the analyzed mutations can in principle compensate each others effects. In terms of stability, the stable mutation S43A is able to slightly improve the low stability of Y36A but, for the V_H/V_L association, Y36A in combination with the worst binder P44A doubles the affinity for V_H compared to P44A alone. This coincides with the fact that these two residues are supposed to interact and are the most important V_L interface residues in the covariation analysis based on number and strength of covariations with other interface residues.

The biophysical properties of isolated V_H domains are in general more affected by mutations compared to V_L³². MAK33 V_H domains show the lowest stabilities with a D_{1/2} value of 0.2 M GdmCl compared to 1.2 ± 0.1 M GdmCl for MAK 33 V_L and 0.9 ± 0.1 M for 1HEZ V_H. But in the case of 1HEZ V_H, it can be speculated that the higher stability might be caused by the homodimer formation, which is probably associated with a shielding effect of otherwise exposed hydrophobic V_H/V_L interface residues. The homo-dimerization of 1HEZ V_H variants (except W47A) adds another dimension to this study since it validates that the impact of highly conserved residues is preserved even in the presence of competing homo-dimerization. For the majority of the MAK33 V_H point mutations, especially L45A, an increased K_D for association with the V_L domain could be detected, which correlated in most cases with an impaired binding of the antigen. But V_H W47A, for example, showed no antigen binding while the V_H/V_L association was only slightly decreased. 1HEZ V_H exhibited an impaired binding to V_L only for two mutants L45A and V37A. According to the simulation data for MAK33 (Table S1), the assigned V_L interaction partners of these two residues are P44 and F98. This fits perfectly to the identified impairments in V_H/V_L association of the two V_L mutants in both V_L domains studied. Taken together, the MAK33 and 1HEZ data for the V_H/V_L association point toward a striking importance of the interaction partners L45(V_H)/P44(V_L) and a contribution of V37(V_H)/F98(V_L). In terms of antigen binding, it has to be kept in mind that additionally to the stabilizing effect of the V_L domain, antigen binding represents for the heterodimer an additional layer of stabilization. For the MAK33 V_H domain, in isolation some of the conserved V_H/V_L interface residues are critical for structure formation and stability. This might cause the impaired association with the V_L domain. In general, association and antigen binding do not necessarily correlate.

Surprisingly, we could not find a general consistent relationship between secondary and/or tertiary structural changes, stability, association or antigen binding. Since it is known that antibody domains share the common ‘Ig fold’, the mutations most likely do not alter the highly conserved Ig fold, but rather lead to local structural changes in agreement with the MD simulations. Due to V(D)J joining and somatic hypermutations, immunoglobulin variable domains naturally experience an extreme sequence variation, except for some rare highly conserved residues. It thus makes sense that the Ig fold has an intrinsically high tolerance to structural changes. Interestingly, our results show that this is also the case for functionally important residues. However, in some cases natural mutations can disrupt the stable Ig topology, leading to the rare but severe disease AL amyloidosis²⁷.

It seems that during antibody biogenesis the effect of CDRs on the stability of V_H domains is a decisive, so far underappreciated factor. Especially concerning the observed MAK33 V_H instability, the outcome of grafting experiments with stable human consensus sequences was interesting. The grafting constructs revealed that CDRs, in addition to antigen binding, affect variable domain structure strongly. This is especially true for CDR-H3 (Table 1). Comparing the MAK33 and 1DHU/1DH5 CDRs, CDR- H3 differed most, in terms of length (14 amino acids for MAK33 versus 11 for 1DHU) as well as charge. According to Morea and coworkers, CDR-H3 conformation does not only depend on the environment² but can be assigned to different conformation types based on some key amino acid positions in the CDRs. Since MAK33 possesses a lysine at position 94 and an aspartate at position 101, which can form a salt bridge, it is assumed to have a bulged conformation in contrast to 1DHU. As the CDR-H3 also contributes to the interface and interacts with the V_L domain, this conformational difference might not only affect antigen binding. In our case, the exchange of the MAK33 CDR-H3 impairs antigen binding (10-fold increased K_D) and slightly increases domain stability. The data for the grafting constructs show a very different picture for the V_H and V_L domain. For V_L, the CDRs seem to be important for structure and V_H/V_L association but concerning antigen binding, also within the 1DH5 framework, the affinity is wt-like. The V_H domain, though, did not exhibit such a strong CDR dependence for V_H/V_L association, only the framework had a slight

effect. However, antigen binding was almost not detectable when grafting the CDRs to the 1DHU framework. Since exchanging only CDR-H3 of the V_H MAK33 domain leads to a severe impairment of the antigen binding, this indicates a crucial role of the V_H domain and especially the CDR-H3 in this process. An explanation for these observations could be the mentioned differences between MAK33 and 1DHU V_H CDR-H3. Interestingly, grafting either the CDRs or framework from the human consensus sequences on MAK33 V_L or V_H , always led to an increase in stability. This might be an important aspect for CDR selection and the interplay with domain stability. Previous studies applying CDR graftings for antibody humanization approaches showed the importance and complexity of the influence of specific framework residues in the context of antigen binding and stability improvement^{11,36,50,51}. Here we could show that vice versa, CDRs themselves can be considered as a crucial determinant of stability.

Our MD simulations for the MAK33 variable domains indicate that the mutations altered conformational fluctuations of the isolated domains which cause structural and mobility changes at the binding interface. The non-optimal packing at the protein-protein interface leads to increased fluctuations at the interface which is also manifested in a reduction of interactions (reduced binding affinity) and also to fluctuations in the antigen binding loop regions which can reduce the binding affinity for antigens, in agreement with previous findings where it was shown that subtle changes in the interface can affect the affinity for the antigen⁵⁰. Interestingly, we found both for the wild type but also for the mutants that regions involved in antigen binding change the flexibility pattern upon complex formation. Hence, the V_L/V_H association appears to lock some of the antigen binding loops into distinct conformations and associated flexibilities. This is in line with the observed subtle influence of some interface residues on antigen binding affinity despite the location distant from the antigen binding region. In some cases also increased solvation at the interface was observed. This was especially the case for the mutations with the most impaired V_H/V_L association (V_L P44A, V_H L45A). The presence of water molecules at the interface reduces intermolecular contacts between protein partners by giving the interface an increasing non-specific character. Thus, the experimentally observed changes in binding affinity and stability of the mutations are due to a combination of effects. A quantitative correlation with the experimentally observed change in antigen binding affinity due to the mutation is, however, not observed. It should be emphasized that this is also not expected since binding to the antigen is affected by each CDR loop differently and changes in loop mobility can in principle affect antigen binding affinity in an unpredictable way.

Taken together our data indicate that multiple determinants regulate the V_H/V_L association and the affinity for the antigen. The interplay between interface interactions and CDRs turned out to be complex with mutual influences on V_H/V_L association and antigen binding.

Material and Methods

Unless otherwise stated, all experiments were carried out at 25 °C. Measurements were performed in 50 mM sodium phosphate buffer at pH 7.5

Cloning and protein expression. V_H and V_L were cloned into pET28 A (Novagen, Darmstadt, Germany) with NcoI and HindIII (NEB, Hitchin, UK) and expressed in *E. coli* BL21 star (Invitrogen, Carlsbad, USA). The transformed cells were grown in LB medium containing kanamycin at 37 °C until an OD_{600} of 0.6–0.8 was reached. The expression was induced by the addition of 1 mM isopropyl β -D-thiogalactopyranoside (IPTG). After 12 h, cells were harvested, and preparation of inclusion bodies was carried out as described previously²³. The purification was performed according to the procedure described for the V_L domain²⁴. Protein purity was checked by SDS-PAGE (Fig. S7).

Single point mutations were introduced by a quick change PCR approach using the QuikChange[®] Site-Directed Mutagenesis Kit (Agilent Technologies Inc., Santa Clara, USA) according to the manufacturer's recommendations. Primers were ordered from MWG Operon (Ebersberg, Germany). Intact protein was verified by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.

CD and Fluorescence spectroscopy. CD measurements were carried out using a Jasco J-720 spectropolarimeter (Jasco, Grossumstadt, Germany) equipped with a Peltier element. Far-UV CD spectra were measured using 10 μ M protein in 1.0-mm quartz cuvettes between 260 nm and 198 nm and near-UV CD spectra between 320 nm and 260 nm using 50 μ M protein in 5-mm quartz cuvettes. The spectra were accumulated 16 times and buffer corrected.

For denaturant-induced unfolding transitions, structural changes were monitored by fluorescence spectroscopy at 355 nm. Excitation wavelength was 280 nm and slit widths were 1 nm (excitation) and 3 nm (emission) for V_H and 2 nm and 5 nm for V_L , respectively. All measurements were performed with 1 μ M protein in a 1-cm quartz cuvette. The samples were incubated overnight at 20 °C at the different GdmCl concentrations prior to measurements.

Data evaluation was performed with Origin 8 G (OriginLab, Northampton, USA); for GdmCl-transitions a two-state model was applied⁵².

SPR. V_H/V_L association was measured by SPR experiments, performed with a Biacore \times 100 (GE Healthcare, Freiburg, Germany). The wild type domain was immobilized on a CM5 chip by amine coupling chemistry. Subsequently, multi cycle runs with titrations of the mutant domains, ranging from 0–200 μ M, were measured at 20 °C with an injection time of 60 seconds and a constant flow rate of 10 μ l/min. After each cycle, bound protein was removed with 2 M NaCl regeneration solution.

ELISA. Binding of the variable domains of MAK33 to the antigen creatine kinase was analyzed by ELISA. Assay components and microwell plates were from Roche (Mannheim, Germany). Samples were prepared in 10 μ l volume. Different mutants were tested within a concentration range of 100 nM up to 50 μ M against wild type V_L

or V_H with a C-terminal Flag-tag for detection. After the addition of 90 μ l reaction mix I, the sample was incubated in a streptavidine-coated microwell plate to immobilize human biotinylated creatine kinase. Incubation was performed for 45 minutes with constant agitation at 20 °C or 10 °C for V_L and V_H mutants, respectively. Afterwards, the samples were washed with sterile pure water for three times. Then 100 μ l/well of reaction mix II were added together with the detection antibody for the Flag-tag coupled to horseradish peroxidase in a 1:15.000 dilution. Afterwards, the samples were washed three times with water again and then 100 μ l of reaction mix III was added. The product of the enzymatic reaction was monitored at 405 nm in a GENios plate reader (Tecan, Männedorf, Switzerland) for 0.5–3 hours until a plateau was reached.

Molecular dynamics simulations. Start structures of MAK33 V_L/V_H complexes as well as individual V_L and V_H domains were obtained by extracting the corresponding coordinates from the crystal structures 1FH5⁵⁰. Missing residues were added/corrected using the program PyMol⁵⁰ with a final sequence corresponding exactly to the wild type sequence used in the experiments. Model start structures of all mutants were generated based on the wild type structure with residue substitutions generated *in silico*. All Molecular Dynamics (MD) simulations as well as the analysis of root-mean square deviation (RMSD) and fluctuations (RMSF) were performed using the Gromacs4.6 package^{53,54} in combination with the AmberSB99_ILDN force field⁵⁵. Proteins were solvated in dodecahedral boxes including explicit ions (Na^+ and Cl^-) and explicit (TIP3P) water molecules⁵⁶. The simulation systems were first energy-minimized (until the maximum force was smaller than 500 kJ/mol) followed by heating up to 300 K at a constant volume with position restraints on the protein. Subsequently, a pressure equilibration at 1 bar with position restraints on the protein was carried out. All production simulations were performed at a temperature of 300 K and a pressure of 1 bar and extended to 100 ns. Root mean square deviation (RMSD) and root mean square fluctuations (RMSF) with respect to the mean structure were calculated with `g_rms` and `g_rmsf` modules of Gromacs. Snapshots were created using VMD⁵⁷.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

- Chothia, C., Novotny, J., Bruccoleri, R. & Karplus, M. Domain association in immunoglobulin molecules. The packing of variable domains. *Journal of molecular biology* **186**, 651–663 (1985).
- Morea, V., Tramontano, A., Rustici, M., Chothia, C. & Lesk, A. M. Conformations of the third hypervariable region in the VH domain of immunoglobulins. *Journal of molecular biology* **275**, 269–294 (1998).
- Vargas-Madrado, E. & Paz-Garcia, E. An improved model of association for VH-VL immunoglobulin domains: asymmetries between VH and VL in the packing of some interface residues. *Journal of molecular recognition: JMR* **16**, 113–120 (2003).
- Bork, P., Holm, L. & Sander, C. The immunoglobulin fold. Structural classification, sequence patterns and common core. *Journal of molecular biology* **242**, 309–320 (1994).
- Abhinandan, K. R. & Martin, A. C. Analysis and prediction of VH/VL packing in antibodies. *Protein engineering, design & selection: PEDS* **23**, 689–697 (2010).
- Narayanan, A., Sellers, B. D. & Jacobson, M. P. Energy-based analysis and prediction of the orientation between light- and heavy-chain antibody variable domains. *Journal of molecular biology* **388**, 941–953 (2009).
- Chothia, C., Gelfand, I. & Kister, A. Structural determinants in the sequences of immunoglobulin variable domain. *Journal of molecular biology* **278**, 457–479 (1998).
- Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Miller, M. & Perry, H. In National Institutes of Health, Bethesda (1983).
- Wang, N. *et al.* Conserved amino acid networks involved in antibody variable domain interactions. *Proteins* **76**, 99–114 (2009).
- Perchiacca, J. M., Bhattacharya, M. & Tessier, P. M. Mutational analysis of domain antibodies reveals aggregation hotspots within and near the complementarity determining regions. *Proteins* **79**, 2637–2647 (2011).
- Nakanishi, T., Tsumoto, K., Yokota, A., Kondo, H. & Kumagai, I. Critical contribution of VH-VL interaction to reshaping of an antibody: the case of humanization of anti-lysozyme antibody, HyHEL-10. *Protein science: a publication of the Protein Society* **17**, 261–270 (2008).
- Chatellier, J., Van Regenmortel, M. H., Vernet, T. & Altschuh, D. Functional mapping of conserved residues located at the VL and VH domain interface of a Fab. *Journal of molecular biology* **264**, 1–6 (1996).
- Khalifa, M. B. *et al.* Effects on interaction kinetics of mutations at the VH-VL interface of Fabs depend on the structural context. *Journal of molecular recognition: JMR* **13**, 127–139 (2000).
- Honegger, A. & Pluckthun, A. The influence of the buried glutamine or glutamate residue in position 6 on the structure of immunoglobulin variable domains. *Journal of molecular biology* **309**, 687–699 (2001).
- Honegger, A., Malebranche, A. D., Rothlisberger, D. & Pluckthun, A. The influence of the framework core residues on the biophysical properties of immunoglobulin heavy chain variable domains. *Protein engineering, design & selection: PEDS* **22**, 121–134 (2009).
- Rothlisberger, D., Honegger, A. & Pluckthun, A. Domain interactions in the Fab fragment: a comparative evaluation of the single-chain Fv and Fab format engineered with variable domains of different stability. *Journal of molecular biology* **347**, 773–789 (2005).
- Hugo, N., Lafont, V., Beukes, M. & Altschuh, D. Functional aspects of co-variant surface charges in an antibody fragment. *Protein science: a publication of the Protein Society* **11**, 2697–2705 (2002).
- Tan, P. H., Sandmaier, B. M. & Stayton, P. S. Contributions of a highly conserved VH/VL hydrogen bonding interaction to scFv folding stability and refolding efficiency. *Biophysical journal* **75**, 1473–1482 (1998).
- Iseman, D. E., Lancet, D. & Pecht, I. Folding pathways of immunoglobulin domains. The folding kinetics of the Cgamma3 domain of human IgG1. *Biochemistry* **18**, 3327–3336 (1979).
- Bertz, M., Buchner, J. & Rief, M. Mechanical stability of the antibody domain CH3 homodimer in different oxidation states. *Journal of the American Chemical Society* **135**, 15085–15091 (2013).
- Feige, M. J. *et al.* An unfolded CH1 domain controls the assembly and secretion of IgG antibodies. *Molecular cell* **34**, 569–579 (2009).
- Feige, M. J. & Buchner, J. Principles and engineering of antibody folding and assembly. *Biochimica et biophysica acta* **1844**, 2024–2031 (2014).
- Feige, M. J., Walter, S. & Buchner, J. Folding mechanism of the CH2 antibody domain. *Journal of molecular biology* **344**, 107–118 (2004).
- Simpson, E. R., Herold, E. M. & Buchner, J. The folding pathway of the antibody V(L) domain. *Journal of molecular biology* **392**, 1326–1338 (2009).
- Nokwe, C. N. *et al.* The Antibody Light-Chain Linker Is Important for Domain Stability and Amyloid Formation. *Journal of molecular biology* **427**, 3572–3586 (2015).

26. Nokwe, C. N. *et al.* A Stable Mutant Predisposes Antibody Domains to Amyloid Formation through Specific Non-Native Interactions. *Journal of molecular biology* **428**, 1315–1332 (2016).
27. Nokwe, C. N. *et al.* A residue-specific shift in stability and amyloidogenicity of antibody variable domains. *The Journal of biological chemistry* **289**, 26829–26846 (2014).
28. Englander, S. W., Mayne, L. & Rumbley, J. N. Submolecular cooperativity produces multi-state protein unfolding and refolding. *Biophysical chemistry* **101–102**, 57–65 (2002).
29. Graille, M. *et al.* Crystal structure of a Staphylococcus aureus protein A domain complexed with the Fab fragment of a human IgM antibody: structural basis for recognition of B-cell receptors and superantigen activity. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 5399–5404 (2000).
30. Bonagura, V. R. *et al.* Mapping IgG epitopes bound by rheumatoid factors from immunized controls identifies disease-specific rheumatoid factors produced by patients with rheumatoid arthritis. *Journal of immunology* **160**, 2496–2505 (1998).
31. Knappik, A. *et al.* Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *Journal of molecular biology* **296**, 57–86 (2000).
32. Ewert, S., Huber, T., Honegger, A. & Pluckthun, A. Biophysical properties of human antibody variable domains. *Journal of molecular biology* **325**, 531–553 (2003).
33. Nieba, L., Honegger, A., Krebber, C. & Pluckthun, A. Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved *in vivo* folding and physical characterization of an engineered scFv fragment. *Protein engineering* **10**, 435–444 (1997).
34. Kugler, M. *et al.* Stabilization and humanization of a single-chain Fv antibody fragment specific for human lymphocyte antigen CD19 by designed point mutations and CDR-grafting onto a human framework. *Protein engineering, design & selection: PEDS* **22**, 135–147 (2009).
35. Honegger, A. Engineering antibodies for stability and efficient folding. *Handbook of experimental pharmacology*, 47–68 (2008).
36. Ewert, S., Honegger, A. & Pluckthun, A. Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering. *Methods* **34**, 184–199 (2004).
37. Weidenhaupt, M. *et al.* Functional mapping of conserved, surface-exposed charges of antibody variable domains. *Journal of molecular recognition: JMR* **15**, 94–103 (2002).
38. Worn, A. & Pluckthun, A. Stability engineering of antibody single-chain Fv fragments. *Journal of molecular biology* **305**, 989–1010 (2001).
39. Jung, S. *et al.* The importance of framework residues H6, H7 and H10 in antibody heavy chains: experimental evidence for a new structural subclassification of antibody V(H) domains. *Journal of molecular biology* **309**, 701–716 (2001).
40. Chen, L. *et al.* Anti-hepatoma human single-chain Fv antibody and adriamycin conjugates with potent antitumor activity. *International immunopharmacology* **18**, 20–26 (2014).
41. Nicholson, I. C. *et al.* Construction and characterisation of a functional CD19 specific single chain Fv fragment for immunotherapy of B lineage leukaemia and lymphoma. *Molecular immunology* **34**, 1157–1165 (1997).
42. Wang, T. & Duan, Y. Probing the stability-limiting regions of an antibody single-chain variable fragment: a molecular dynamics simulation study. *Protein engineering, design & selection: PEDS* **24**, 649–657 (2011).
43. Galitsky, B. Revealing the set of mutually correlated positions for the protein families of immunoglobulin fold. *In silico biology* **3**, 241–264 (2003).
44. Feige, M. J. *et al.* The structural analysis of shark IgNAR antibodies reveals evolutionary principles of immunoglobulins. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 8155–8160 (2014).
45. Ewert, S., Cambillau, C., Conrath, K. & Pluckthun, A. Biophysical properties of camelid V(HH) domains compared to those of human V(H)3 domains. *Biochemistry* **41**, 3628–3636 (2002).
46. Streltsov, V. A. *et al.* Structural evidence for evolution of shark Ig new antigen receptor variable domain antibodies from a cell-surface receptor. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 12444–12449 (2004).
47. Kovalenko, O. V. *et al.* Atypical antigen recognition mode of a shark immunoglobulin new antigen receptor (IgNAR) variable domain characterized by humanization and structural analysis. *The Journal of biological chemistry* **288**, 17408–17419 (2013).
48. Muyldermans, S. Nanobodies: natural single-domain antibodies. *Annual review of biochemistry* **82**, 775–797 (2013).
49. Davies, J. & Riechmann, L. ‘Camelising’ human antibody fragments: NMR studies on VH domains. *FEBS letters* **339**, 285–290 (1994).
50. Fransson, J. *et al.* Human framework adaptation of a mouse anti-human IL-13 antibody. *Journal of molecular biology* **398**, 214–231 (2010).
51. Lapidot, G. D. *et al.* AbDesign: An algorithm for combinatorial backbone design guided by natural conformations and sequences. *Proteins* **83**, 1385–1406 (2015).
52. Bolen, D. W. & Santoro, M. M. Unfolding free energy changes determined by the linear extrapolation method. 2. *Incorporation of delta G degrees N-U values in a thermodynamic cycle. Biochemistry* **27**, 8069–8074 (1988).
53. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* **4**, 435–447 (2008).
54. Berendsen, H. J. C., Vanderspoel, D. & Vandrunen, R. Gromacs - a Message-Passing Parallel Molecular-Dynamics Implementation. *Comput Phys Commun* **91**, 43–56 (1995).
55. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins-Structure Function and Bioinformatics* **78**, 1950–1958 (2010).
56. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* **79**, 926–935 (1983).
57. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *Journal of molecular graphics* **14**(33–38), 27–38 (1996).

Acknowledgements

This study was supported by grants of the DFG to J.B. and the SFB1035 to M.Z.

Author Contributions

The *in vitro* experiments were conceived, conducted and analyzed by E.M.H., C.J., B.W. and C.B. with the exception of the ELISA and the grafting construct experiments, which were conducted and analyzed by S.D. and J.E., respectively. The molecular dynamics simulations were performed by S.K. and M.Z. E.M.H., C.J., B.W. and J.B. wrote the manuscript. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-12519-9>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017