Molecular Immunology 90 (2017) 143–149

Contents lists available at ScienceDirect

Molecular Immunology

journal homepage: www.elsevier.com/locate/molimm

# An unusual cysteine $V_L87$ affects the antibody fragment conformations without interfering with the disulfide bond formation

Carolina Attallah[a], María Fernanda Aguilar[a], A. Sergio Garay[b], Fernando E. Herrera[b], Marina Etcheverrigaray[a], Marcos Oggero[a], Daniel E. Rodrigues[b,c,*]

[a] UNL, CONICET, FBCB, Cell Culture Laboratory, Ciudad Universitaria UNL, Pje. "El Pozo" - C.C. 242, S3000ZAA Santa Fe, Argentina
[b] UNL, CONICET, FBCB, Departamento de Física, Ciudad Universitaria UNL, Pje. "El Pozo" - C.C. 242, S3000ZAA Santa Fe, Argentina
[c] INTEC, CONICET-UNL, Predio CONICET Santa Fe, Pje. "El Pozo", S3000 Santa Fe, Argentina

ARTICLE INFO

*Keywords:*
Single-chain variable fragment
Cysteine contiguous residues
Molecular modeling
Mutating cysteine

ABSTRACT

The Cys residues are almost perfectly conserved in all antibodies. They contribute significantly to the antibody fragment stability. The relevance of two natural contiguous Cys residues of an anti-recombinant human-follicle stimulation hormone (rhFSH) in a format of single-chain variable fragment (scFv) was studied. This scFv contains 5 Cys residues: $V_H22$ and $V_H92$ in the variable heavy chain ($V_H$) and $V_L23$, $V_L87$ and $V_L88$ in the variable light chain ($V_L$). The influence of two unusual contiguous Cys at positions $V_L87$ and $V_L88$ was studied by considering the wild type fragment and mutant variants: $V_L$-C88S, $V_L$-C87S, $V_L$-C87Y. The analysis was carried out using antigen-binding ability measurement by indirect specific ELISA and a detailed molecular modeling that comprises homology methods, long molecular dynamics simulations and docking. We found that $V_L$-C87 affected the antibody fragment stability without interfering with the disulfide bond formation. The effect of mutating the $V_L$-C87 by a usual residue at this position like Tyr caused distant structural changes at the $V_H$ region that confers a higher mobility to the $V_H$-CDR2 and $V_H$-CDR3 loops improving the scFv binding to the antigen.

## 1. Introduction

scFvs are recombinant antibody fragments consisting of the variable light chain ($V_L$) and the variable heavy chain ($V_H$) domains connected via a short flexible peptide (linker) (Huston et al., 1988). Each variable domain has a characteristic tertiary structure consisting of two β-sheets that contain three hypervariable loops, known as complementary determining regions (CDRs), evenly distributed between four less variable framework regions (FRs). The antigen binding site is primarily composed by the six CDRs located in the variable domains that are close to each other (Morea et al., 2000).

The Cys residues at positions $V_H22$ and $V_H92$ in $V_H$, and $V_L23$ and $V_L88$ in $V_L$ (numbering according to Kabat et al. (1991)) are almost perfectly conserved in all antibodies. They form buried intra-domain disulfide bridges which contribute significantly to the antibody fragment stability (Langedijk et al., 1998). Only 0.3% of the 6042 mouse $V_L$ sequences listed in the Kabat database (http://www.bioinf.org.uk/abysis2.7/), contains a residue different from Cys at position $V_L88$. The conserved Cys residue at position $V_L88$ locates as the flanking residue of $V_L$-CDR3. The $V_L87$ position contains the Phe residue in 28% of $V_L$ sequences of the database and Tyr in 69% of the instances. Less than 1% of sequences contain a Cys at position $V_L87$, including some which present a residue different from Cys at $V_L88$. Although these unusual sequences are available in the Kabat database, there is no information concerning the function of the corresponding antibodies. Some studies demonstrated that only few antibody molecules were able to tolerate the loss of the referred disulfide bonds. Nevertheless for some sequences it has been proven that additional mutations stabilized the antibody fragment structures (Frisch et al., 1996; Langedijk et al., 1998; Proba et al., 1998; Rudikoff and Pumphrey, 1986). The instability of these antibodies with missing disulfide bonds might be the reason that limits the occurrence of Cys at unusual positions. (Duan et al., 2012) reported the presence of 5 Cys residues in a scFv anti-rabies virus G protein. The unusual Cys was found at the first position after the Cys that flanks the $V_L$-CDR3 (i.e., inside the CDR3). Despite of the presence of the free thiol residue, this scFv bound to the target molecule. These authors also proposed that the free thiol might form a disulfide bond mismatch in the scFv structure based on the evidence that mutating the unusual Cys to Ser improved the antibody fragment performance (Duan et al., 2014).

In this work we study the influence of contiguous Cys residues at positions $V_L87$ and $V_L88$ present in an anti-recombinant human follicle

* Corresponding author at: Ciudad Universitaria UNL. Pje. "El Pozo" - C.C. 242, S3000ZAA Santa Fe, Argentina.
*E-mail address:* danielr@fbcb.unl.edu.ar (D.E. Rodrigues).

http://dx.doi.org/10.1016/j.molimm.2017.07.008
Received 26 April 2017; Received in revised form 10 July 2017; Accepted 15 July 2017

stimulating hormone (rhFSH) scFv, derived from a monoclonal antibody (mAb) that naturally contains 5 Cys residues. For the present sequence, the unusual Cys residue was found at the position immediately before the $V_L$-C88 that flanks the $V_L$-CDR3 (i.e., inside the FR3), at variant with the case studied by Duan et al. (2014, 2012). The aim of the present work is to analyze the role of the unusual Cys residue at position $V_L$87 and the changes that produce the mutation of this amino acid. Different mutants: $V_L$-C88S, $V_L$-C87S and $V_L$-C87Y were experimentally developed and their binding ability to rhFSH were evaluated. 3D structures of the wild type, mutants $V_L$-C87S and $V_L$-C87Y of the scFv were obtained by comparative modeling and long molecular dynamics simulations (MD) were run to evaluate the equilibrium properties of the fragment. In silico docking analysis of the scFv over rhFSH structure were conducted to identify the interacting residues. We found that the Cys residue at $V_L$87 is not involved in the disulfide bond formation and therefore a disulfide bridge mismatch is not feasible as it was described by Duan et al. (2014) for their scFv. The unusual Cys residue in the scFv anti-rhFSH affects the antibody fragment stability and its ability to bind the antigen.

## 2. Materials and methods

### 2.1. scFv expression vectors

The wild type scFv anti-rhFSH gene was amplified by overlap extension PCR (SOE-PCR). The $V_H$ and $V_L$ cDNAs templates were previously obtained from hybridoma cells producing the anti-rhFSH mAb. A flexible peptide linker $(Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser)_3$ coding DNA was used to connect $V_H$ and $V_L$. The reactions were catalyzed by Platinum™ Pfx DNA Polymerase (Invitrogen, USA). $V_H$, linker and $V_L$ products were amplified with primers 1-2, 3-4 and 5-6 (listed in Table 1-SI, see supplementary information), respectively. The wild type scFv final product was amplified with primers 1 and 6. Site-directed mutagenesis was employed to obtain the mutants $V_L$-C87Y, $V_L$-C88S and $V_L$-C87S by SOE-PCR using the following sets of primer pairs 1-8 and 7-6, 1-10 and 9-6, and 1-12 and 11-6, listed in Table 1-SI (supplementary information), where the UGC, UGU and UGC codons were changed for UAC, UCU and UCC, respectively. The final mutated scFv products were amplified with primers 1 and 6. All constructs were clone into modified pET22b(+) vector (Novagen®, USA) through the removal of the N-terminal pelB signal peptide (pET22b-pelB) to express the scFv variants in the cell cytoplasm of E. coli SHuffle® T7 strain (NEB Inc., USA). This strain possesses a chromosomal copy of the disulfide bond isomerase DsbC (de Marco, 2009) to assure a proper folding of disulfide bridges. All constructs were confirmed by sequencing (Macrogen Inc, Korea).

### 2.2. scFv expression and extraction

The plasmids pET22b-pelB containing the wild type scFv or each of the mutated scFvs ($V_L$-C87Y, $V_L$-C88S and $V_L$-C87S) were used to transform *E. coli* SHuffle® T7 competent cells (NEB Inc., USA). To express the wild type scFv and the mutated ones, single colonies were used to inoculate 10 ml LB medium supplemented with 0.1 mg/ml ampicillin and grown at 30 °C and 200 rpm. After overnight (ON) incubation, the cultures were inoculated into 200 ml LB medium supplemented with 0.1 mg/ml ampicillin and 1% (w/v) glucose, and grown at 30 °C and 200 rpm until they reached an $OD_{600}$ of 0.8. Protein expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and cells were allowed to grow at 20 °C at 150 rpm for another 16 h. Cells were harvested by centrifugation at 5000 rpm for 15 min and then resuspended in 20 mM phosphate-buffered saline (PBS: 10 mM $Na_2HPO_4$, 1.8 mM $KH_2PO_4$, 14 mM NaCl, 2.7 mM KCl, pH 7.4), 500 mM NaCl and 0.1 mM EDTA, and lysed using sonication 8 × 30 s. The cell lysates were centrifuged at 10,000 rpm for 30 min and the supernatants containing the soluble scFv were stored at −20 °C.

### 2.3. scFv quantification by competitive ELISA

A competitive ELISA was performed in 96-well plates (Greiner, Austria) coated with 50 ng per well of rabbit anti-6xHisTag polyclonal antibodies (ab125265, Abcam, UK) prepared in 50 mM carbonate/bicarbonate buffer (pH 9.6). Plates were incubated for 1 h at 37 °C and then ON at 4 °C. After blocking 1 h at 37 °C with 1% (w/v) bovine serum albumin in PBS (PBS-BSA), plates were incubated with 2-fold serial dilutions of the test samples, followed by the addition of equal volume of 50 ng/ml of a biotinylated C-terminal 6xHis-tagged irrelevant protein. This protein was trivial for the system and it was only used as a His-fused protein to compete with the His-tagged scFv to bind to the coated anti-6xHisTag antibody. After incubating for 1 h at 37 °C, horseradish peroxidase (HRP)-conjugated streptavidin (Amdex™, GE Healthcare) at a dilution of 1:10,000 was added to each well. After 1 h at 37 °C, plates were incubated for 10 min with substrate solution (3 mg/ml *o*-phenylenediamine, 0.12% (v/v) $H_2O_2$ in 50 mM phosphate-citrate buffer). The reaction was stopped by the addition of 2 N $H_2SO_4$. Absorbance was measured at 492 nm with a microtiter plate reader (Labsystems Multiskan MCC/340, Finland). Between every step, plates were washed 6 times with PBS containing 0.05% (v/v) Tween 20 (PBS-T). Dilutions of tested samples and antibodies were prepared in PBS-T containing 0.1% (w/v) BSA. The C-terminal 6xHis-tagged irrelevant protein was previously purified by immobilized-metal affinity chromatography (IMAC) and used as standard to prepare a 2-fold serial dilution curve from 400 ng/ml to 3.125 ng/ml. Also, its biotinylated form was used as tracer for the competition step. The assay was reproduced in triplicate.

### 2.4. scFv-rhFSH binding determination by indirect specific ELISA

Indirect ELISA was performed in 96-well plates. Coating was achieved by incubating the plates with 50 ng per well of rhFSH prepared in 50 mM carbonate/bicarbonate buffer (pH 9.6). Plates were incubated for 1 h at 37 °C and then ON at 4 °C. After blocking 1 h at 37 °C with 1% (w/v) bovine serum albumin in PBS (PBS-BSA), plates were incubated with 2-fold serial dilutions of the test samples for 1 h at 37 °C. Then, plates were incubated with appropriately diluted rabbit anti-6XHisTag polyclonal antibodies for 1 h at 37 °C. Finally, an horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (DAKO, Denmark) at a dilution 1/2000 was added to the wells. After 1 h at 37 °C, plates were incubated for 10 min with substrate solution and then, the reaction was stopped by the addition of 2 N $H_2SO_4$. Absorbance was measured at 492 nm with a microtiter plate reader. Washing between steps and dilutions were carried out using the buffers mentioned in 2.3.

### 2.5. Molecular modeling

#### 2.5.1. Homology modeling

Template identification was performed using Psi-Blast (Jones and Ward, 2003) and HHblits (Remmert et al., 2012), which were able to find very good templates to build the wild type scFv model from Protein Data Bank (http://www.rcsb.org/pdb/)(Berman et al., 2000). The best of them, 4h0g (Tapryal et al., 2013), had 71.5% and 80.2% of identical and similar residues, respectively with 4% of gaps; these values guarantee the quality of the model. The alignment was obtained running HHblits and it was used to build the model using Modeller software (Eswar et al., 2001) after visual inspection of the alignment to eliminate any possible mistake. The initial model quality was verified using Qmean6 composite score (Benkert et al., 2008) and the stereochemical quality of the model was monitored with PROCHECK (Laskowski et al., 1993). The Qmean6 Z-score of −1.20 indicates that our model fell at −1.20 standard deviation from a distribution of scores calculated for reference structures solved by X-ray crystallography with the same number of residues. Taking into account that the model was going to be

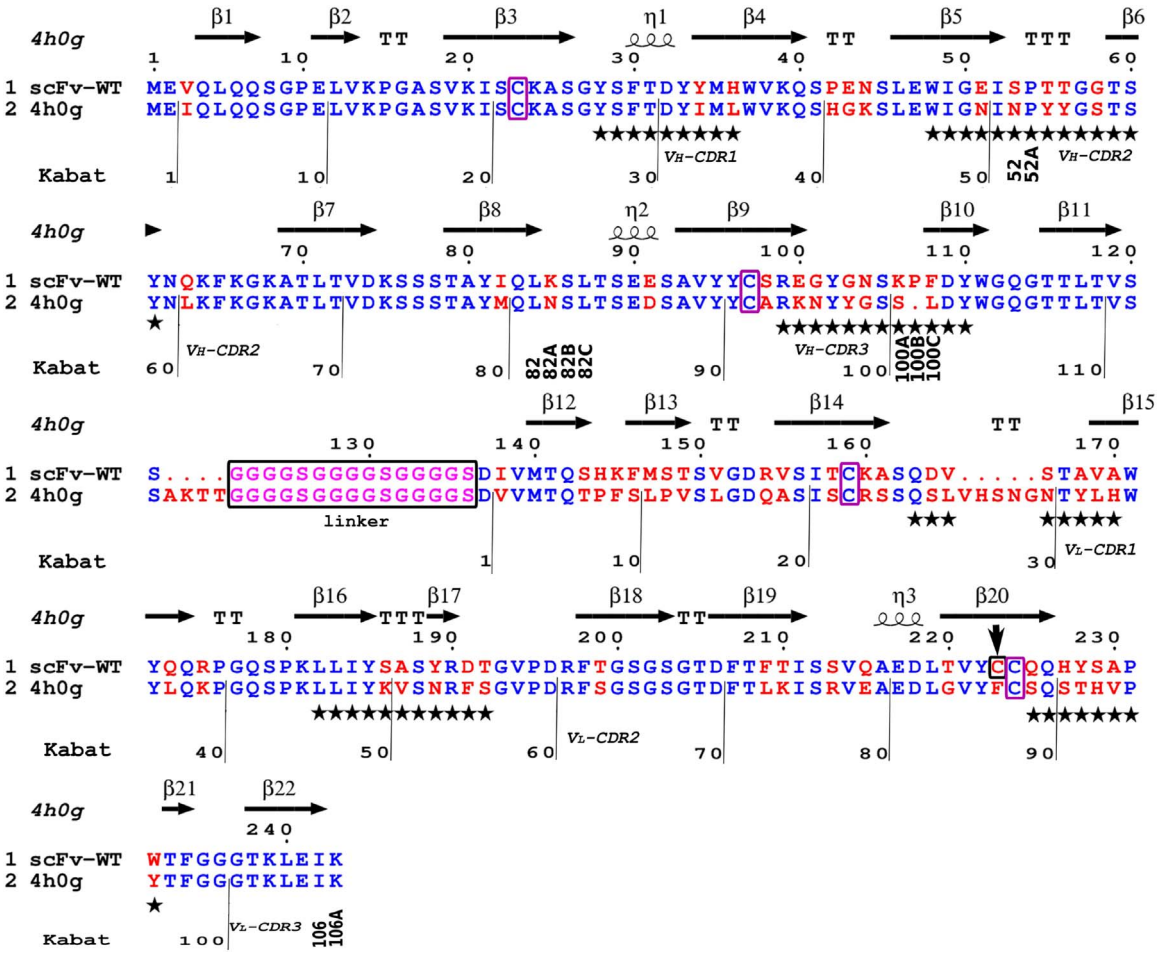


**Fig. 1.** Sequence Alignment between template structure 4h0g and scFv obtained with Hhblits and used to build the homology model. Identical residues are shown in blue color, linker residues in purple color and non-dentical residues in red. The CDR residues are tagged by black stars and a black arrow is used to mark off the $V_L$-C87 residue involved in the mutation analysis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

used for running a long Molecular Dynamics simulation, not more efforts to improve it were made, hopping that the last procedure drive the loops of the protein to a correct position.

### 2.5.2. Molecular dynamics simulations (MD)

Molecular dynamics (MD) simulations were carried out using the program package GROMACS 5.0.4 (Abraham et al., 2015) along with AMBER99SB-ILDN (Lindorff-Larsen et al., 2010) force field. Initially wild type model was solvated with TIP3P (Jorgensen et al., 1983) water embedded in an octahedral box. The minimum distance between the protein and the simulation box was 1.2 nm, to avoid any inappropriate behavior of the water molecules. There was no need to add contraions to the system because the protein charge was 0. Subsequently, the system was subjected to a steepest descent energy minimization until reaching a tolerance of 100 kJ/mol. After the solvent molecules were equilibrated with the fixed protein at 300 K (27 °C), the entire system was gradually relaxed and heated up to 300 K. Finally, 300 ns MD simulations were performed under the room temperature (300 K) and pressure with coupling time constant 1.0 ps. The particle mesh Ewald method (Essmann et al., 1995) was used to treat long-range Coulombic interactions. The LINCS algorithm (Hess et al., 1997) was used to constrain bond lengths of protein atoms and SETTLE for the water molecules (Miyamoto and Kollman, 1992). Van der Waals forces were considered up to distances of 1.5 nm, and Coulomb interactions were truncated at 1.5 nm. Two mutations from the wild type model, C223S ($V_L$-C87S in Kabat nomenclature) and C223Y ($V_L$-C87Y), were introduced to the structure taken from last trajectory frame of the wild type structure. A MD simulation for 300 ns in the same conditions (300 K, 1 atm) were performed for the three systems (wild type, $V_L$-C87S and $V_L$-C87Y mutants). After this equilibration period each system was simulated another 300 ns at 310 K (37 °C) which was the temperature of the experimental work (Fig. 1-SI shows a summary of the equilibration process).

### 2.5.3. Docking antibody-antigen

Potential antigenic sites were identified to reduce the spatial docking search considering that hFSH is a heterodimer which contains a common α−subunit (chain A) with human chorionic gonadotropin (hCG), luteinizing hormone (LH) and thyrotropin (TSH) and a unique β−subunit (chain B). Two different servers were used: ElliPro (Ponomarenko et al., 2008) and Discotope (Kringelum et al., 2012). The residues predicted as potential antigenic sites of hFSH (1FL7) are shown in Table 2-SI (see supplementary information). Both servers identified groups of residues in the same region as antigenic sites, mainly in the chain A. In chain B, they only coincided in the identification of the antigenic determinant composed by residues 122–130 named to as Det-2.

Considering that CDR identification methods may miss > 20% of the residues that actually bind the antigen (Nucleic Acids Res. 2012 Jul; 40 (Web Server issue): W521–W524), the server Paratome (Kunik et al., 2012) was used to check the coincidence between the antigen predicted sites and the previous identified loops (see Fig. 2-SI in supplementary information).

Based on the hFSH antigenic determinants and the mAb paratopes

predicted and described above, Protein-Protein docking were run using Haddock web server (de Vries et al., 2010). Initially, three docking simulations were run for each antigen (1FL7, 1HRP), one for each antigenic determinant. A dimeric structure of 1FL7 and 1HRP were used, excluding the carbohydrates moieties pending on residues: 76, 102 (chain A, in both structures), 25, 42 (1FL7 - chain B) and 33, 50, 141, 147, 152, 158 (1HRP - chain B) because Haddock web server is not able to treat modified amino acid residues. The best initial model of the scFv was chosen as the second interacting protein.

The residues predicted as paratopes and the corresponding residues of each antigenic determinant were defined as "active" sites. Passive residues were defined as those remaining around the active ones.

## 3. Results and discussion

### 3.1. Cys at position $V_L$87 is not involved in the intra-domain disulfide bond

#### 3.1.1. Homology model

Based on the good quality sequence alignment (see Fig. 1) obtained between scFv wild type variant and the best template 4h0g structure we can assume that $V_H$-C22, $V_H$-C92, and $V_L$-C23, $V_L$-C88 must be involved in disulfide bridge respectively. Accordingly $V_L$-C87 is not able to form an intra-domain disulfide bridge because is next to $V_L$-C88 and both belong to the same beta strand (Beta 20) which presents side chain opposing orientations. The secondary structure prevents the correct orientation of the side chain of $V_L$-C87 to form an intradomain disulfide bridge with $V_L$-C23. We found for our studied scFv that a disulfide bond mismatch cannot occur as it was reported in the case of two contiguous Cys in the scFv from the work of Duan et al. (2014, 2012).

A 3D rendering of the model obtained from the above alignment is shown in Fig. 2. The representation illustrates the above conclusion about the reasons that preclude the $V_L$-C87 residue to bridge to $V_L$-C23.

#### 3.1.2. scFv-C87S and scFv-C88S binding to rhFSH

Two mutants of the studied scFv were designed and produced: $V_L$-C87S and $V_L$-C88S. The Ser residue was used to replace Cys due to their physicochemical similarity. The presence of the hydroxyl instead of the sulphydryl group avoids the occurrence of a de-stabilizing or reactive free thiol.

Considering that non potential N-/O-glycosylation sites were found into the scFv sequence, wild type and mutated scFvs were expressed intracellularly in bacteria and the functionality to bind rhFSH was assayed by indirect ELISA in the cell extracts. Previously, scFv-containing supernatants were quantified by competitive ELISA. In order to determine an apparent affinity constant ($KA_{app}$), the half effective concentration (EC50) to bind the rhFSH was obtained from the indirect ELISA. The $KA_{app}$ values are shown in Table 1. The free thiol group of the wild type scFv did not preclude its binding ability to rhFSH. However, the $V_L$-C88S mutant completely lost the ability to bind the rhFSH. This evidence confirmed that for the present case, the $V_L$-C88 residue is critically involved in the ability of the scFv to bind rhFSH and that the $V_L$-C87 was not able to replace its function in the mutant. Based on the 3D model for scFv we conclude that the $V_L$-C88S mutation precluded the formation of the disulfide bond of the $V_L$-C88 residue and that its structure did not rearrange to allow the $V_L$-C87 residue to build a functional replacement bond.

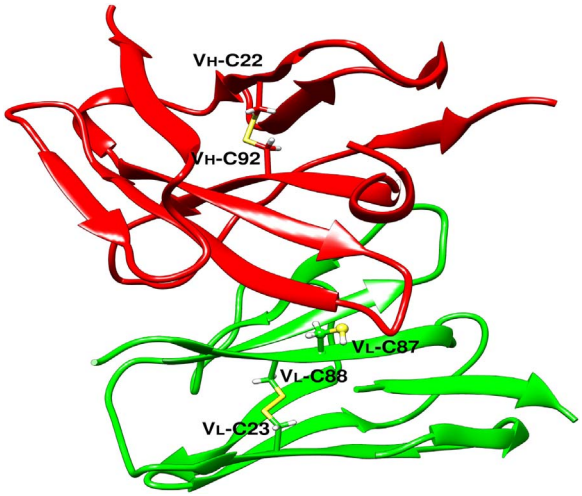


**Fig. 2.** Representation of the homology model structure obtained for the scFv. The Cys residues involved in disulfide bridges are rendered with sticks and $V_L$-C87 with Ball and Sticks.

**Table 1**
scFvs apparent affinity constants.

| scFv | $KA_{app}$ ($M^{-1}$) |
|---|---|
| Wild type | $(3.5 \pm 0.2) \times 10^8$ |
| $V_L$-C87S | $(9.6 \pm 0.2) \times 10^8$ |
| $V_L$-C88S | n/b |
| $V_L$-C87Y | $(4.5 \pm 0.4) \times 10^9$ |

n/b: no binding.

### 3.2. Effect of exchanging the amino acid residue at $V_L$87

#### 3.2.1. Binding ability of $V_L$-C87S and $V_L$-C87Y

Less than 1% of sequences listed in the Kabat database contain Ser at position$V_L$87 and there is no information concerning the function of these antibodies. Nevertheless, the mutation $V_L$-C87S improved the ability of the mutant scFv to bind rhFSH (2.7-fold increment was observed, see Table 1).

Considering that 69% of the sequences listed in the Kabat database contain a Tyr residue at $V_L$87, the $V_L$-C87Y mutant of scFv was generated and its binding ability to rhFSH assayed to compare with the above variants.

The $V_L$-C87Y mutant improved 4.7 and 12.7 times the ability to bind the hormone with respect to the $V_L$-C87S mutant and wild type scFv, respectively (see Table 1). Although the $V_L$-C87 residue does not participate of the disulfide bond formation, its presence had a negative impact on the scFv ability to bind the hormone.

Besides, integrity of the molecules in terms of protein degradation was analyzed by western blot, detecting only one band with a molecular mass of 28,300 Da that corresponds to the entire scFv variants (Fig. 3-SI in supplementary information).

#### 3.2.2. Molecular dynamic simulations

Molecular dynamics simulations were performed to evaluate the role of each mutated residue at $V_L$87 on the equilibrium structure of these variants. The wild type scFv was also simulated to compare the results. Time evolution of backbone Root Mean Square Deviation (RMSD) and radius of gyration of all scFv variants were used to check the convergence to equilibrium. Fig. 4-SI in supplementary information, shows the RMSD calculated taking the first trajectory frame as reference. The RMSD showed that after 200 ns of time evolution all the simulated variants reached a plateau that is indicative that each system has reached the thermal equilibrium. The gyration radii temporal evolutions also show the same condition. Therefore, the last 40 ns of the simulated time for each variant were taken as the production stage for equilibrium properties analysis. The asymptotic value of the RMSD for the $V_L$-C87Y mutant was larger than that for the wild type and $V_L$-C87S which is indicative that the former variant undergoes the largest structural changes taking as reference the wild type.

Root Mean Square Fluctuation (RMSF) profiles as a function of the residue number is a measure of the spatial fluctuations that undergoes each amino acid in the equilibrium condition and was presented in

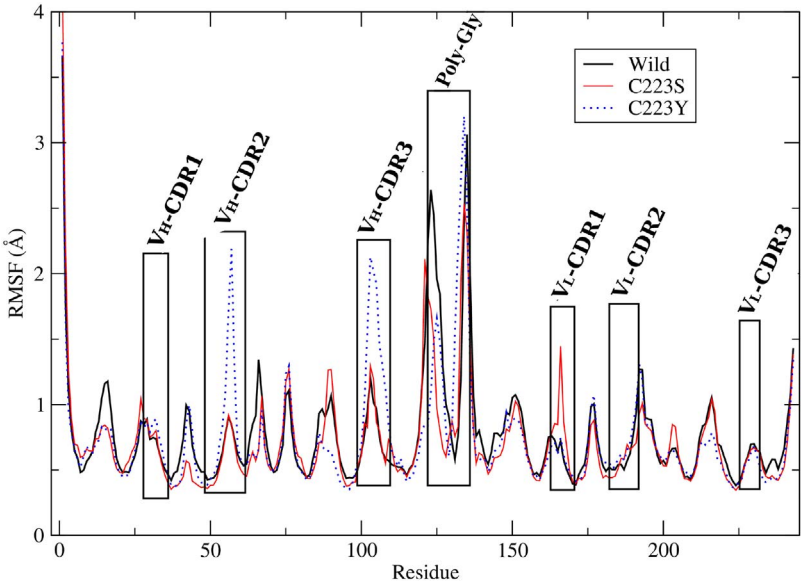


**Fig. 3.** Root Mean Square Fluctuations (RMSF) of the backbone atoms respect to the average backbone structure obtained in the last 40 ns (production stage), for each system. Residues belonging to each CDR loop and the linker are outlined with square boxes. C223S is $V_L$-C87S in Kabat nomenclature and C223Y is $V_L$-C87Y.

Fig. 3 for all the variants. Fig. 3 shows that the largest fluctuations correspond to the poly-Gly linker, whose profiles exhibit a dip in the middle due to interactions with antibody chains. The $V_L$-C87Y mutant presents notably larger spatial fluctuations at the region of the $V_H$-CDR2 and $V_H$-CDR3 loops. These features indicates a higher mobility of the loop in the C87Y mutant or at least the existence of two or more conformational states.

Fukuda et al. (2017) demonstrated that a proper mobility in the CDR loops is important for high affinity binding via conformational rearrangements. The work of Li et al. (2014) showed that the effects of mutations on proteins could propagate through the molecular network leading to long-range changes. Studying an anti-lymphotoxin-β receptor scFv, they found that a stabilizing mutation could induce a local increased rigidity and simultaneously lead to a higher flexibility in a portion of the protein distant from the mutation site in agreement with the Le Châtelieŕs principle. They also conclude that the formation or rupture of hydrogen bonds (HB) are responsible of the observed changes. Therefore we decided to study the changes in interactions between the wild type and mutant scFvs that could affect the structural stability of different regions.

The average number of HB during the production stage of the simulations between residues of the heavy and light chain domains of scFv ($V_H$: residues 1–121; $V_L$: residues 137–243) were calculated. Table 2 shows the list of the most persistent inter-chain HB's residue pairs calculated in the production stage. At the top of the table, the average total numbers of inter-chain HB are quoted. The averaged total number of inter-chain HB of the $V_L$-C87Y mutant (5.2 ± 1.2) is larger than the number for $V_L$-C87S (3.7 ± 1.3) and that for the wild type scFv (3.5 ± 1.6). The inter-chain HBs between the side chains of Gln40-Gln174 ($V_H$39 and $V_L$38 in Kabat notation) appears recurrently in $V_L$-C87Y (85% + 92 %) and $V_L$-C87S (73 %) mutants. The larger value of persistence for the $V_L$-C87Y mutant is a consequence of the building of two HB's by the side chains of these residues (Gln). The wild type scFv did not exhibit HB's between the same residue pair, which is an evidence of the local change in the structures due to the mutations. The wild type protein built a less frequent (50%) interchain HB but in this case between Gln45-Gln174 ($V_H$44 and $V_L$38 in Kabat notation). The locations of the involved residues are sketched in a molecular representation in the supplementary information (see Fig. 5 -SI). The fact that $V_H$44 residue (Gln) participate in the building of the HB in the wild type scFv instead of $V_H$40, is indicative of the local structure changes in this region. The appreciably larger total number of inter-chain HB's for the $V_L$-C87Y mutant signals the higher inter-domain stability of this variant.

We have noted above the increased spatial fluctuations of the regions of loops $V_H$-CDR2 and $V_H$-CDR3 in the $V_L$-C87Y variant (see Fig. 3). A clustering analysis of the geometry of these loops extended to the three variants (wild type, $V_L$-C87Y and $V_L$-C87S mutants) was performed to categorize different conformations. The clustering algorithm serves to assemble several molecular structures into different

**Table 2**
Interchain most persistent hydrogen bonded residue pairs for each studied variant. The persistence was defined as the percentage of frames in which the HB exists referred to the total number of frames in the production stage. HB pairs with persistence larger than 50% were highlighted with a yellow background. At the top of the table the average total number of inter-chain HB's are quoted. We used the geometrical criteria for the HB existence that the distance between the donor and acceptor atoms is less than 3.5 Å and the angle between the acceptor-donor hydrogen less than 30°. (For interpretation of the reference to colour in this table, the reader is referred to the web version of this article.)

| 37°C Wild inter-chain HB (avg. Hbond=3.5±1.6) | | | 37°C C223S inter-chain HB (avg. Hbond=3.7±1.3) | | | 37°C C223Y inter-chain HB (avg. Hbond=5.2±1.2) | | |
|---|---|---|---|---|---|---|---|---|
| donor | aceptor | persist. | donor | aceptor | persist. | donor | aceptor | persist. |
| Phe108-main | Tyr172-side | 56.7 | Phe108-main | Tyr172-side | 74.3 | Ser45-side | Gly235-main | 99.0 |
| Ser45-side | Gln174-side | 51.1 | Gln40-side | Gln174-side | 72.9 | Gln174-side | Gln40-side | 91.6 |
| Trp232-side | Lys106-main | 36.1 | Asn44-side | Phe234-main | 69.5 | Ser179-side | Gly112-main | 87.0 |
| Tyr96-side | Gln178-main | 26.4 | Ser45-side | Gly236-main | 40.7 | Gln40-side | Gln174-side | 85.4 |
| Gln174-side | Ser45-side | 24.8 | Trp232-side | Glu100-side | 25.2 | Lys106-side | Asp191-side | 80.2 |
| Lys145-side | Pro42-main | 22.9 | Tyr96-side | Gln174-side | 23.6 | Trp111-side | Tyr172-side | 11.2 |

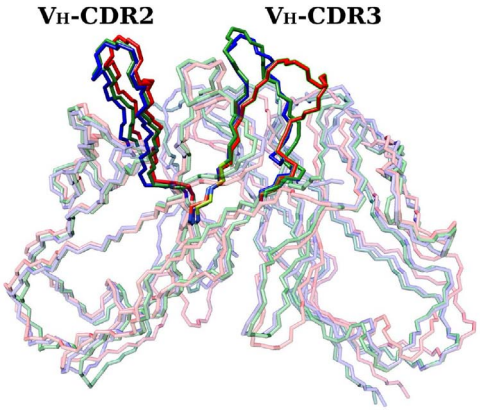


**Fig. 4.** Superposition of the average backbone structure for the three variants of scFv (the last 40 ns of the production run). $V_L$-C87S, $V_L$-C87Y and wild type scFv structures are colored in green, red and blue, respectively. The loops are shown in opaque color while the rest of the structures with transparency to enhance the visualization. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

groups according to their structural similarity measured by a RMSD criterion. A RMSD cutoff of 0.4 Å was used for the calculation. The analysis for the loop $V_H$-CDR2 showed that there is only one principal conformational motif for the three antibody variants. The clusterization of the structures of the loop $V_H$-CDR3 showed only one shared main conformational motif for the wild type scFv and the $V_L$-C87S mutant. The $V_L$-C87Y mutant exhibits two main conformational motifs of the $V_H$-CDR3 loop, that are both different from those of the other variants. One of the conformational motifs of the $V_L$-C87Y mutant is correlated with the formation of the persistent (50%) Lys106-Asp191 inter-chain HB ($V_H$100A and $V_L$55 in Kabat nomenclature, see Table 2), that links loops $V_H$-CDR3 and $V_L$-CDR2. It is relevant to remark that these features are associated with a global relative rotation of the two chain domains of approximately 11° ± 3. There is also a relative rotation of the chain domains for the case of $V_L$-C87S mutant but it amounts to only 5° ± 2. The different structural motifs of the loops $V_H$-CDR3 and $V_L$-CDR2 are graphically illustrated in Fig. 4, that shows the wild type and mutant structures superimposed, each with the average backbone structure of the motifs.

### 3.2.3. Molecular docking and experimental assays of binding

We have performed an in silico docking simulation of scFv on the hFSH and hGCH to gain some insight into the potential binding sites of the antibody. In a first step the antigenic regions of both targets were identified using two different algorithms implemented in servers ElliPro (Ponomarenko et al., 2008) and Discotope (Kringelum et al., 2012) (see Table 2-SI in supplementary information for details). This step was conducted to restrict the search space of the docking simulations.

Table 3 summarizes the scores obtained for the docking simulations of scFv on the three antigenic determinants of both hormones. The results showed that the region of Det-2 of both hormones has a higher score than the other antigenic determinants and therefore a higher binding probability. These results also showed that the binding to hFSH has higher score values than to hCGH. Both hormones differ in their chain B sequence. The energetic analysis beyond the docking procedure of scFv was able to recognize a difference between hFSH and hGCH. Analyzing the structure of the docking configurations (results not shown), we can conclude that the Det-2 antigenic determinant is the one that exposes a region of chain B in its closest environment to scFv.

The binding of the original mAb was analyzed by indirect ELISA coating plates with hFSH or hCGH. A binding value 2.5-fold higher with hFSH than hCGH was obtained (data not shown). Also, western blot experiments of both hormones were carried out in dissociated and non-dissociated conditions. The original mAb equally recognized chain A when subunits of hFSH or hCGH were heat-dissociated but it only recognized heterodimeric hFSH when hormone were assayed in non-dissociated conditions (data not shown). Then considering that the scFv binds chain A, it is probably that the paratope recognize subtle differences between them.

The docking simulations and the experimental evidence showed that scFv is mainly directed against antigen chain A, which is identical in both hormones. Nevertheless, the antigen binding to scFv is affected by the presence of chain B. These facts give support to the hypothesis that the Det-2 epitope (sequence KTMLVQ) is the target for the scFv.

Assuming as valid the hypothesis that Det-2 is the antigenic determinant recognized by the antibody, we used this fragment as a hFSH derived target in a docking simulation on the three variants of scFv (wild type, $V_L$-C87S and $V_L$-C87Y mutants), to test if the algorithm is able to detect mutation effects. The six potential binding loops of scFv were used as active regions in the scanning over the receptor. Table 4 summarizes the results of these docking simulations. The scFv variant $V_L$-C87Y showed the highest score (2-fold higher than scFv wild type). These results are in agreement with the above immunochemical experiments which showed that $V_L$-C87Y mutant had the best binding ability among the three variants to bind the hormone. From the analysis of the representative structures of the docking on $V_L$-C87Y mutant, we observed that some residues of $V_H$-CDR2 (6 residues) and $V_H$-CDR3 (7 residues) are in close proximity to the Det-2 epitope. Some of these residues also built HB to Det-2 antigen determinant. From the MD simulations we had concluded that $V_H$-CDR3 loop exhibits for the $V_L$-C87Y variant conformational motifs different from the wild type or $V_L$-C87S mutant and the larger relative rotation of the chain domains (see above).

## 4. Conclusions

We have shown that the extra $V_L$-C87 of an anti-rhFSH antibody, naturally containing 5 Cys residues, affected the antibody fragment ability to bind the antigen without interfering with the disulfide bond formation. In the $V_L$-C88S mutant, the $V_L$-C87 was not able to replace the $V_L$-C88 role and the antibody lost its ability to bind hFSH. This fact is a consequence of the structure of the scFv in the environment of these residues that could prevent the disulfide bond formation. Nevertheless, the $V_L$-C87 is related to the inter-chain stability and the conformations of scFv loops as it was shown by the analysis of the wild type, $V_L$-C87S and $V_L$-C87Y mutants. Both mutants increased the ability of the antibody to bind hFSH. From the molecular dynamics study we found that the mutant $V_L$-C87Y caused a conformational change in the $V_H$-CDR3 loop, increased flexibility of the $V_H$-CDR2 and $V_H$-CDR3 loops and a relative rotation of the $V_H$ and $V_L$ domains. These changes also comprised the increment in the total number of inter-chain hydrogen bonds

**Table 3**
Haddock Docking score in arbitrary units. All the complex quality assessment obtained by Haddock web server (RMSD-l (< 5 Å), RMSD-i (< 2 Å), FCC(> 0.3)) indicates that they could be classified as "Medium" quality.

| | hFSH(1FL7) | | | hCG(1HRP) | | |
|---|---|---|---|---|---|---|
| scFv | Det-1 | Det-2 | Det-3 | Det-1 | Det-2 | Det-3 |
| wild type Initial Model (Haddock score*) | −74 | −112 | −51 | −70 | −106 | −45 |

**Table 4**
Haddock scores obtained using only Det-2 as Ligand target of hFSH (active residues), and three variants of scFv as receptors (six loops for each variant).

| | hFSH(1FL7)-Det-2 | | |
|---|---|---|---|
| scFv | wild type | $V_L$-C87S | $V_L$-C87Y |
| Haddock score* | 941 | −54 | −82 |

that brings more stability to the multi-domain structure. For the case of the $V_L$-C87S mutant, the molecular dynamics simulations showed only a moderate increment in the number of inter-chain HB but no appreciable change of the mobility of epitope loops. The docking simulations of scFv on hFSH and hGCH showed that the antibody binds preferentially to the former antigen. This result is in agreement with the measured best performance of $V_L$-C87Y mutant to bind the α-subunit of the heterodimeric hFSH or of the related hormone hCGH, suggesting the antigenic determinant Det-2 (sequence KTMLVQ) as the binding site. The docking of the wild type scFv, $V_L$-C87S and $V_L$-C87Y mutants on hFSH proposed that the last variant has the larger probability to bind the target. The docking studies also suggested that the $V_H$-CDR2 and $V_H$-CDR3 are the loops that mainly interact with the Det-2 epitope and therefore this interaction is prone to be affected by their conformations or mobility.

## Acknowledgements

Authors acknowledge the use of the computational facilities of the FaCAP (Facilidad de Computación de Alta Performance) at the Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, Argentina, and would like to thank the financial support of CONICET (PIP-2012-0895); UNL-CAID (PI 501 201101-00040 LI, PIC 504 201501-00063 LI, PI 501 201101 00313). CA, FEH, ME, MO and DER are member of the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). We also thank to Zelltek S.A. to provide rhFSH for immunochemical tests, Guillermina Forno to allow working with the original anti-rhFSH antibody and Ma. de los Milagros Bürgi for providing information about properties of the mentioned antibody.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.molimm.2017.07.008.

## References

Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., Lindahl, E., 2015. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1, 19–25.
Benkert, P., Tosatto, S.C., Schomburg, D., 2008. QMEAN: A comprehensive scoring function for model quality assessment. Proteins 71, 261–277.
Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The protein data bank. Nucleic Acids Res. 28, 235–242.
Duan, Y., Gu, T.J., Jiang, C.L., Yuan, R.S., Zhang, H.F., Hou, H.J., Yu, X.H., Chen, Y., Zhang, Y., Wu, Y.G., Kong, W., 2012. A novel disulfide-stabilized single-chain variable antibody fragment against rabies virus G protein with enhanced in vivo neutralizing potency. Mol. Immunol. 51, 188–196.
Duan, Y., Gu, T., Zhang, X., Jiang, C., Yuan, R., Li, Z., Wang, D., Chen, X., Wu, C., Chen, Y., Wu, Y., Kong, W., 2014. Negative effects of a disulfide bond mismatch in anti-rabies G protein single-chain antibody variable fragment FV57. Mol. Immunol. 59, 136–141.
Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., Pedersen, L.G.A., 1995. Smooth particle mesh Ewald method. J. Chem. Phys 103, 8577–8593.
Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M.-y., Pieper, U., Sali, A., 2001. Comparative protein structure modeling using MODELLER. Current Protocols in Protein Science. John Wiley Sons, Inc.
Frisch, C., Kolmar, H., Schmidt, A., Kleemann, G., Reinhardt, A., Pohl, E., Uson, I., Schneider, T.R., Fritz, H.J., 1996. Contribution of the intramolecular disulfide bridge to the folding stability of REIv, the variable domain of a human immunoglobulin kappa light chain. Fold. Des. 1, 431–440.
Fukuda, N., Suwa, Y., Uchida, M., Kobashigawa, Y., Yokoyama, H., Morioka, H., 2017. Role of the mobility of antigen binding site in high affinity antibody elucidated by surface plasmon resonance. J. Biochem. 161, 37–43.
Hess, B., Bekker, H., Berendsen, H.J.C., Fraaije, J.G.E.M., 1997. LINCS: A linear constraint solver for molecular simulations. J. Comput. Chem. 18, 1463–1472.
Huston, J.S., Levinson, D., Mudgett-Hunter, M., Tai, M.S., Novotny, J., Margolies, M.N., Ridge, R.J., Bruccoleri, R.E., Haber, E., Crea, R., et al., 1988. Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. 85, 5879–5883.
Jones, D.T., Ward, J.J., 2003. Prediction of disordered regions in proteins from position specific score matrices. Proteins 53 (Suppl. 6), 573–578.
Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L., 1983. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79, 926–937.
Kabat, E.A., National Institutes of H, Columbia U, 1991. Sequences of Proteins of Immunological Interest. U.S. Dept. of Health and Human Services, Public Health Service, National Institutes of Health, Bethesda, MD.
Kringelum, J.V., Lundegaard, C., Lund, O., Nielsen, M., 2012. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. PLoS Comput. Biol. 8, e1002829.
Kunik, V., Ashkenazi, S., Ofran, Y., 2012. Paratome: an online tool for systematic identification of antigen-binding regions in antibodies based on sequence or structure. Nucleic Acids Res. 40, W521–4.
Langedijk, A.C., Honegger, A., Maat, J., Planta, R.J., van Schaik, R.C., Pluckthun, A., 1998. The nature of antibody heavy chain residue H6 strongly influences the stability of a VH domain lacking the disulfide bridge. J. Mol. Biol. 283, 95–110.
Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291.
Li, T., Tracka, M.B., Uddin, S., Casas-Finet, J., Jacobs, D.J., Livesay, D.R., 2014. Redistribution of flexibility in stabilizing antibody fragment mutants follows Le Chatelier's principle. PLoS One 9, e92870.
Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J.L., Dror, R.O., Shaw, D.E., 2010. Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 78, 1950–1958.
Miyamoto, S., Kollman, P.A., 1992. Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. J. Comput. Chem. 13, 952–962.
Morea, V., Lesk, A.M., Tramontano, A., 2000. Antibody modeling: implications for engineering and design. Methods 20, 267–279.
Ponomarenko, J., Bui, H.H., Li, W., Fusseder, N., Bourne, P.E., Sette, A., Peters, B., 2008. ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinf. 9, 514.
Proba, K., Worn, A., Honegger, A., Pluckthun, A., 1998. Antibody scFv fragments without disulfide bonds made by molecular evolution. J. Mol. Biol. 275, 245–253.
Remmert, M., Biegert, A., Hauser, A., Soding, J., 2012. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat. Methods 9, 173–175.
Rudikoff, S., Pumphrey, J.G., 1986. Functional antibody lacking a variable-region disulfide bridge. Proc. Natl. Acad. Sci. U. S. A. 83, 7875–7878.
Tapryal, S., Gaur, V., Kaur, K.J., Salunke, D.M., 2013. Structural evaluation of a mimicry-recognizing paratope: plasticity in antigen-antibody interactions manifests in molecular mimicry. J. Immunol. 191, 456–463.
de Marco, A., 2009. Strategies for successful recombinant expression of disulfide bond-dependent proteins in Escherichia coli. Microb. Cell Fact. 8, 26.
de Vries, S.J., van Dijk, M., Bonvin, A.M., 2010. The HADDOCK web server for data-driven biomolecular docking. Nat. Protoc. 5, 883–897.