Deciphering cross-species reactivity of LAMP-1 antibodies using deep mutational epitope mapping and AlphaFold

Tiphanie Pruvost,[1,2] Magali Mathieu,[3] Steven Dubois,[1] Bernard Maillère,[1] Emmanuelle Vigne,[2] Hervé Nozach [1*]

[1] Université Paris-Saclay, CEA, INRAE, Medicines and Healthcare Technologies Department, SIMoS, Gif-sur-Yvette, France.

[2] Sanofi, Large Molecule Research, Vitry-sur-Seine, France.

[3] Sanofi, Integrated Drug Discovery, Vitry-sur-Seine, France.

Correspondence should be addressed to:

Hervé Nozach
herve.nozach@cea.fr
Service d'Ingénierie Moléculaire pour la Santé (DRF/JOLIOT/DMTS/SIMoS)
Bâtiment 152 Point Courrier 24 CEA/SACLAY
91191 Gif-sur-Yvette cedex, France
Tel. (33) (0)1 69 08 42 15

Keywords:

Monoclonal antibodies, deep mutational scanning, yeast surface display, epitope mapping, cross-species reactivity, LAMP-1

Abbreviations:

DMS: Deep Mutational Scanning

YSD: Yeast Surface Display

NGS: next-generation sequencing

cryo-EM: cryogenic electron microscopy

BLI: biolayer interferometry

FACS: Fluorescence-activated cell sorting

CDRs: complementarity-determining regions

HDX-MS: Hydrogen deuterium exchange mass spectrometry

RMSD: root-mean-square deviation

Abstract

Delineating the precise regions on an antigen that are targeted by antibodies has become a key step for the development of antibody therapeutics. X-ray crystallography and cryogenic electron microscopy are considered the gold standard for providing precise information about these binding sites at atomic resolution. However, they are labor-intensive and a successful outcome is not guaranteed. We used deep mutational scanning (DMS) of the human LAMP-1 antigen displayed on yeast surface and leveraged next-generation sequencing to observe the effect of individual mutants on the binding of two LAMP-1 antibodies and to determine their functional epitopes on LAMP-1. Fine-tuned epitope mapping by DMS approaches is augmented by knowledge of experimental antigen structure. As human LAMP-1 structure has not yet been solved, we used the AlphaFold predicted structure of the full-length protein to combine with DMS data and ultimately finely map antibody epitopes. The accuracy of this method was confirmed by comparing the results to the co-crystal structure of one of the two antibodies with a LAMP-1 luminal domain. Finally, we used AlphaFold models of non-human LAMP-1 to understand the lack of mAb cross-reactivity. While both epitopes in the murine form exhibit multiple mutations in comparison to human LAMP-1, only one and two mutations in the Macaca form suffice to hinder the recognition by mAb B and A, respectively. Altogether, this study promotes a new application of AlphaFold to speed up precision mapping of antibody-antigen interactions and consequently accelerate antibody engineering for optimization.

Introduction

Antibodies bind to antigens in a multitude of ways, defining a wide range of possible interacting antigenic surfaces called epitopes. Generally, epitopes cover a surface of 600-900 square Å and involve one to several dozen amino acids of the antigen protein.[1] These interacting amino acids are grouped in a continuous three-dimensional (3D) surface and can be carried by a stretch of linear sequence or, on the contrary, scattered over the primary sequence of the protein. Delineating epitopes can help to understand antibody functions or to facilitate the selection of antibodies that target specific regions of the antigen.

Epitopes can be mapped by various experimental processes.[2] Over the years, a wide range of techniques have been used to determine which areas of the antigens are recognized by the antibodies. These include structural methods,[3] peptide-based approaches,[4] mutagenesis methods[5, 6] and mass spectrometry.[2, 7] More recently, computational modeling has enabled prediction of the antigen/antibody interface.[8, 9] The field of protein structure prediction has seen unprecedented progress, notably with AlphaFold and RoseTTAFold.[10] X-ray crystallography and more recently cryogenic electron microscopy (cryo-EM) are still considered as gold standards for providing precise information on interaction sites with near atomic resolution. More precisely, 3D structures of complexes of antibodies with their antigens reveal amino acids from both sides of the interacting partners (namely structural epitope for the antigen and structural paratope for the antibody) that are close to each other and the chemical bonds that contribute to stability of the complex. However, the exact role of each amino acid present in the interacting surface can be difficult to decipher. Indeed, not all amino acids within a 4-4.5 Å radius from the other partner are necessarily important contributors to the binding free energy or to the specificity of the interaction.[2]

In recent years, deep mutational scanning (DMS) approaches have considerably accelerated the pace of mutational studies, which can now explore every possible single amino acid substitution in a selected protein.[11, 12] By combining high-throughput screening methods such as display techniques (e.g., phage display, yeast surface display) with deep sequencing, an increasing number of studies have analyzed the mutational landscape to understand the modalities of interaction between protein partners.[12] More specifically, several studies have allowed the identification of the epitope of monoclonal antibodies (mAbs) against prion protein,[13] *S. aureus* alpha toxin,[14] nerve growth factor[15] and

Salmonella antigens.[16] DMS identifies the functional epitope as the key interacting amino acids that cannot be replaced without causing a major loss in binding activity. By extension, DMS has recently proved useful in predicting antigen mutations that allow escape from the action of therapeutic mAbs. This is a known mechanism of resistance against natural or therapeutic antibodies targeting viral antigens that are subject to high selective pressure and for which substitutions can reduce antibody-mediated neutralization.[17] Comprehensive escape maps were notably established for selected viral antigens of Zika,[18] HIV,[19] influenza[20, 21] and SARS-CoV-2[22].

Beyond uncovering the mode of action of therapeutic antibodies and potential resistance mechanisms, detailed knowledge of epitopes can be useful in understanding the cross-reactivity of antibodies to antigens, including selectivity towards proteins belonging to a same family or cross-reactivity between species. Species cross-reactivity is very valuable in evaluating antibody therapeutic potential in preclinical animal models, such as mice or non-human primates.

In this study, we explored the molecular determinants of the binding of two LAMP-1-specific antibodies. While LAMP-1 comprises 50% of all lysosomal membrane proteins and is widely used as a cell surface marker of lymphocyte activation and degranulation, its exact role remains uncertain.[23, 24] LAMP-1 is a physiologically essential protein involved in stabilizing lysosomes and regulating autophagy to prevent embryonic lethality. Previous studies have demonstrated limited cell surface expression of LAMP-1 in normal tissues and moderate to high membrane expression in a number of breast, colorectal, gastric, prostate, lung, and ovarian tumors,[24] making it a target of interest for oncology applications. Some evidence point to a role for LAMP-1 in tumor progression.[25, 26]

LAMP-1 is a type I transmembrane protein comprising two heavily glycosylated luminal domains with 18 potential N-glycosylation sites and 6 O-linked oligosaccharides, a transmembrane domain, and a small cytoplasmic tail. The LAMP-1 protein is highly conserved between human and cynomolgus (97.2% sequence identity), resulting in a difference of 10 amino acids in the luminal part of the protein, which counts 352 amino acids. In sharp contrast, the human LAMP-1 protein is relatively distant from its murine ortholog with a sequence identity of 64.3%. The 3D structure of human LAMP-1 has not been described to date, while a structure of the second luminal domain of the murine LAMP-1 protein is available.[23]

This report describes how combining a DMS approach with structural modeling enabled by AlphaFold successfully uncovered why both mAbs display nanomolar affinity for human LAMP-1, but fail to bind similarly to non-human primate LAMP-1 despite a very high identity between the two species proteins.

Results

Identification of mutations affecting mAb binding to human LAMP-1

We first determined the affinity constants of the two LAMP-1 mAbs for human LAMP-1 and their non-human primate and mouse counterparts. Affinity measurements using biolayer interferometry (BLI) demonstrated the high affinity of mAbs A and B for the human LAMP-1 antigen with apparent K_D values of 0.8 and 12 nM, respectively (Fig. 1). The affinity of mAb A for the cynomolgus antigen was approximately 20-fold lower at 18 nM, while no binding signal was observed with mAb B at the concentration of 200 nM, revealing the lack of cross-reactivity of this antibody (Fig. 1). Furthermore, both antibodies failed to bind the murine antigen at the maximum tested concentration (200 nM). Fluorescence-activated cell sorting (FACS) experiments confirmed the binding profile of the two mAbs to human and cynoLAMP1 presented on the surface of engineered cell lines (data not shown). To link these biochemical data to sequence information, we performed DMS of the human LAMP-1 antigen using the yeast surface display (YSD) technique, thereby expressing mutants of the extracellular domain of human LAMP-1 on the surface of yeast cells. The two luminal domains of human LAMP-1 linked together by its hinge region were anchored in the yeast cell wall through a C-terminal fusion with the Aga2p protein, itself attached to Aga1p by two disulfide bonds (Fig. 2A). We first demonstrated that wild-type human LAMP-1 can be expressed on the yeast surface and retains binding to each antigen-binding fragment (Fab). We further showed that the two mAbs bound simultaneously to LAMP-1, and therefore target independent epitopes. We then generated single mutant libraries of the human LAMP-1 antigen by SOE-PCR using libraries of primers each carrying a single degenerate codon (Fig. 2A). The resulting linear DNA fragment libraries were then transformed into the yeast *S. cerevisiae*. Given the large size of the luminal domain of human LAMP-1 and to facilitate next-generation sequencing (NGS) procedures, we generated 5 sub-libraries encompassing each luminal domain and the hinge region (Fig. 2B).

The five yeast libraries were then simultaneously labeled with Fabs A and B for FACS sorting. Both antibodies were found to bind without affecting the binding of the other molecule, indicating the independence of their two epitopes. We preferred Fab to IgG to avoid experimental bias related to avidity phenomena. Fabs were used at concentrations close to their K_D affinity constants, to allow the most sensitive discrimination between mutants and isolate those for which a loss of recognition by either of the two Fabs is observed. Flow cytometry showed that most LAMP-1 variants displayed strong fluorescence signals with both Fabs (Fig. 2C), demonstrating that corresponding mutations in LAMP-1 had no effect on Fab binding. More interestingly, some subpopulations within libraries 1 and 2 lost binding to Fab B, but not to Fab A (red gates, upper panel of Fig. 2C). Symmetrically, some cells in libraries 4 and 5 expressed LAMP-1 mutants that were no longer recognized by Fab A, but still by Fab B (red gates, lower panel of Fig. 2C). Finally, mutations in the hinge had no effect on the binding of either Fab (Library 3, data not shown). These results not only confirmed that both Fabs bind LAMP-1 at independent epitopes, but also demonstrated that Fab A binds the second luminal domain (libraries 4 and 5) and Fab B binds the first luminal domain (libraries 1 and 2). To identify the amino acid substitutions responsible for the loss of recognition by either of the two Fabs, corresponding cells were sorted before bulk sequencing of their human LAMP-1 mutant sequence.

NGS data were comprehensively tabulated with the enrichment score for each substitution on each position of human LAMP-1 (Fig. 3, Supp. Fig. 2 and 3). Most substitutions had limited influence on the binding of Fabs, and are therefore not detected in the sorted populations. In contrast, mutations with an enrichment score greater than two (i.e., frequencies four times higher in the sorted over unsorted populations) are those that most markedly affect the binding of either Fab to human LAMP-1 (bright red, Fig. 3). For each position, we determined an index by counting the number of substitutions with an enrichment score higher than two. This index and associated mutational pattern were thoroughly analyzed to identify the positions necessary for the recognition of LAMP-1 by the two Fabs.

Positions with an index greater than five were localized in libraries 4 and 5 for Fab A (26 and 7 positions, respectively) and in libraries 1 and 2 for Fab B (9 and 22 positions, respectively) (Fig. 3 and Supp. Fig. 1 and 2). These positions are discontinuously distributed along the primary LAMP-1 sequence, with several motifs consisting of a few consecutive amino acids.

Multiple positions were particularly intolerant to substitutions (index ≥ 15). This is notably the case for positions R254, E281, G282, I309 and P311 for which many substitutions had a deleterious influence on the recognition by Fab

A, while most substitutions in positions R106, A108, I149, D150, Q176, R187 and G188 suppressed Fab B binding (Fig. 3). The DMS data therefore suggest that these positions are critically involved in LAMP-1/Fab binding.

A second class of positions with indexes between five and 15 were also affected by substitutions. Some of these positions are close to key positions with an index higher than 15 in the primary sequence of LAMP-1. They form motifs of 3 - 6 consecutive amino acids in the vicinity of E281-G282 (280-284) and of P311 (308-312) for Fab A, and around R106 (106-108), I149/D150 (149-151), Q176 (175-180) or R187/G188 (185-188) in Fab B. We also observed in this second category some hydrophobic amino acids that were relatively dispersed throughout the primary sequence of LAMP-1. They consisted essentially of leucine, methionine, phenylalanine, or isoleucine residues (e.g., L232, M236, L240, L242, I258, L286, F288 or F290 in libraries 4 and 5, and M43, A44, F46, F50, V52, F94, L100, L102, F128, I175 in libraries 1 and 2) (Fig. 3).

3D modeling to guide the fine determination of the functional epitope

We decided to generate structural and 3D modeling data to distinguish positions directly involved in the epitope from those affecting the overall conformation of the antigen and its folding, and ultimately refine the epitopes. We first solved the crystallographic structure of the complex between Fab B and an aglycosylated form of the first luminal domain of human LAMP-1 (Fig. 4A). This domain adopts the same overall β-prism fold as murine LAMP-1 [23] and DC-Lamp3.[27] Most of the interaction between Fab B and LAMP-1 is mediated by amino acids in the heavy chain complementarity-determining regions (CDRs). Briefly, loop 82-86 of LAMP-1 interacts with CDRH1 and the FR3 loop from the Fab heavy chain. Loop 106-109 interacts with all heavy chain CDRs and loop 149-151 is in contact with CDRH3 and CDRL1. Lastly, loop 178-187 contacts both CDRH1 and CDRH3, along with CDRL1 and CDRL2. All LAMP-1 amino acids at the interface, i.e., at less than 4.5 Å from the Fab molecule are represented in yellow in Fig. 4A and constitute what might be termed the 'structural epitope'.

Next, we examined the localization of the amino acids identified by DMS for the Fab B within the crystallographic structure. We colored in red the 15 positions for which at least 10 substitutions were deleterious to Fab B binding and in yellow the 16 positions for which 5 to 9 substitutions were not tolerated (Fig. 4B). We observed that 11 of the 15 positions with an index higher than 10 were accessible to the solvent, from which 10 positions were in direct contact with the Fab molecule. In contrast, none of the 16 positions with an index of 5 - 9 were within a 4.5 Å radius

from the antigen, 11 of these positions being non-exposed on the protein surface. The upper table of Fig. 5A summarizes these findings and highlights the functional epitope of Fab B. Overall, combining structural information with DMS data enabled fine-tuning of the Fab B functional epitope, by discarding buried positions mutation of which may affect global domain folding.

Given the difficulty of obtaining structures of antigen-antibody complexes, we also sought to use structure models to refine the DMS data. Considering the unparalleled accuracy recently demonstrated by the AlphaFold 2 algorithms, we retrieved the model of the human LAMP-1 first luminal domain from AlphaFold DB (Fig. 4C). The data overlay showed a very good alignment of the AlphaFold model with the crystallographic structure of the LAMP-1 domain, with an root-mean-square deviation (RMSD) of atomic positions of 1.27 Å for all main-chain atoms, demonstrating the good quality of the model. The only noticeable difference between the model and structure lies in the LAMP-1 loop 82-86, which is part of the structural epitope. However, the loop 82-86 defined by AlphaFold was found to form a steric clash with the CDRH1 of Fab B when superimposed with the crystallographic structure (circled in red, Fig. 4B). It should be noted that this loop appears to have limited structural constraints and the model confidence scores for this part of the protein are not very high (Supp. Fig. 3A). Importantly, the introduction of mutations in the 82-86 loop of LAMP-1 was not identified as important for Fab B recognition. Altogether, this showed that the AlphaFold model of the first LAMP-1 luminal domain can be used to refine the functional epitope of Fab B and, more broadly, gives confidence to use of the algorithm for predicting the 3D structure of the second LAMP-1 luminal domain for which no structural data is available. On this basis, we filtered out the positions identified as buried in the second luminal domain of human LAMP-1 using the AlphaFold model and mapped the functional epitope of Fab A (lower table in Fig. 5A). We have summarized the different steps of the method, from cell sorting to the different data processing steps in Fig. 6.

Identification of LAMP-1 positions implicated in the low cross-reactivity with murine and cynomolgus antigens

Finally, we sought to use these epitope mapping data to understand the low cross-reactivity of Fab A and the lack of recognition of Fab B for cynomolgus and murine antigens. We retrieved the murine LAMP-1 antigen available on AlphaFold DB [28] and generated the model for cynomolgus LAMP-1 with ColabFold.[29] Fig. 5C highlights the amino

acids located similarly to those identified in the human epitope on the surface of the cynomolgus and mouse models and the amino acids that diverge from the human sequence in the corresponding species.

We observed that two positions differed in cynomolgus LAMP-1 within the mAb A epitope, with substitutions T283S and I309T, and only one in mAb B epitope, namely G187E. Consistently, both I309 and G187 were identified as positions critical for binding by DMS (Fig. 5A); more specifically, I309T and G187E mutations resulted in loss of binding to Fab A and B, respectively (Fig. 3). The sequences of the human and murine antigens within the considered zones diverge quite significantly, with 10 and 5 differences in the Fab A and B epitopes, respectively. These differences likely alter dramatically both topology and charges of the epitopes, explaining the lack of recognition of the murine form of LAMP-1 by both Fabs.

Discussion

This report promotes the systematic use of the most recent structural modeling algorithms such as AlphaFold combined with DMS data to expedite the parallel fine mapping of antibody/antigen interfaces in antibody discovery programs. In the absence of pre-existing structural data, AlphaFold models of the antigen turned out to be essential for the three-dimensional representation of high-resolution DMS data. It proved very useful to finely identify surface amino acids of the antigen and thus differentiate substitutions influencing protein folding from those directly involved in the antibody/antigen interface. It ultimately enabled refining of the functional epitopes of two mAbs and explains their interactions with their antigen orthologs.

In recent years, many studies have sought to determine the epitope of different therapeutic antibodies. In addition to understanding mechanisms of action and selecting antibodies that target specific areas of proteins, epitope determination is also valuable in strengthening intellectual property and patent protection.[30] Few methods are capable of identifying conformational epitopes with high resolution at the amino acid level.[31]

All methods for epitope mapping have limitations. X-ray crystallography or cryoEM can reveal simultaneously both the epitope and paratope of a mAb/antigen complex. However, they are dependent on the quality of the complex and its capacity to crystallize at high enough resolution or generate high-quality images, respectively. Hydrogen deuterium exchange mass spectrometry (HDX-MS) is a fast and cost-effective alternative approach enabling

parallelized epitope mapping. However, its accuracy and precision can be compromised by insufficient peptide coverage for large complexes or highly glycosylated antigens, or the inability to discriminate between direct binding interface and allosteric conformational change. [32] The Ala mutagenesis technique can provide some answers on the areas of the antigen involved in the interaction, but is far less precise than DMS, which scans the 20 proteinogenic amino acids. In our dataset, we observe that Ala substitutions would not have identified some important positions, such as K151 for mAb B or L310 for mAb A, and of course not A108, which is already an alanine residue.

The nature of the antigen can also be a challenge for some methods. Unlike cryoEM and HDX-MS, Alascan and X-ray crystallography are applicable to soluble proteins or protein domains, but these approaches prove to be technically very complex for integral membrane proteins such as G-protein coupled receptors or transporters. These targets can be studied with DMS expressed on the surface of yeast or mammalian cells, opening new possibilities for such challenging targets.

Finally, structure-based methods and HDX-MS do not provide information on the impact of single mutations. They can be combined with predictive methods such as *in silico* **ΔΔ**G mutagenesis to propose which mutations in an already known epitope/paratope region would result in a gain or a loss of binding affinity, which then requires additional experimental validation.

Here, we show that YSD/DMS combined with AlphaFold 2 can successfully and rapidly map epitopes in a parallelized manner. Importantly, DMS goes beyond epitope mapping by generating data on the effect of single substitutions in the antigen on its binding to the antibody, and thereby contributes to the understanding of antibody escape mutants or in our case, of lack of species cross-reactivity.

Yeast cells are known to be capable of expressing a large variety of proteins on their surface. [33] It is remarkable that, despite the presence of several disulfide bridges and numerous glycosylation sites, the cellular machinery of *S. cerevisiae* allows surface expression of the full extracellular domain of the human LAMP-1 protein and its proper recognition by the two studied mAbs. The probable presence of mannose-rich glycans typical of yeast glycosylation machinery[34] in place of mammalian glycosylation patterns did not affect antibody recognition. This is consistent with the successful complex formation between Fab B and the aglycosylated form of the first luminal domain of human LAMP-1 used for the crystallography study. The functional and structural epitopes uncovered in this study ultimately

corroborate that the two mAbs do not recognize LAMP-1 glycotopes. While N-glycosylation sites are distant from the functional epitope of Fab A, they lie at the periphery of the Fab B binding site with an orientation not hindering its binding (Supp. Fig. 4).

In our experimental setup, the LAMP-1 mutant libraries were almost comprehensive and only rare substitutions could not be screened. Of the 354 positions considered, no variant was detected for four of them (Supp. Fig. 2 and 3), which suggests a problem during the synthesis of the oligonucleotides rather than during the screening process. For all other positions, the use of degenerate NNS codons provided very good coverage (greater than 99%) of possible mutations. These data provide new evidence that DMS approaches are not limited to small proteins and can be applied to larger proteins. The most time-consuming part of the DMS consists in the generation of the libraries, while their sorting in FACS and subsequent high-throughput sequencing are rapid. Therefore, this DMS approach is fully parallelizable in determining the epitopes of multiple antibodies targeting the same antigen. We report here a parallel study of two mAbs, but it could be scaled up to a few dozen antibodies in a cost- and time-effective manner, as previously discussed in other studies.[31]

By design, we performed this epitope mapping approach by simultaneous labeling of the libraries with both Fabs shown to be non-competitors, aiming at controlling surface expression and potentially folding of each LAMP-1 mutant. The DMS demonstrated that each Fab binds one of the two LAMP-1 luminal domains known to be separated by a proline-rich linker region.[35] Consistently, substitutions affecting the recognition of one Fab, even when occurring in a hydrophobic core, did not affect the binding of the other, which indicates a high structural independence of the two luminal domains. AlphaFold model analysis further shows that the predicted aligned error scores are large for pairs of amino acids located in the two distinct domains (Supp. Fig. 3B), thereby confirming their independence. Had the study been performed with antibodies targeting the same luminal domain, it is possible that mutations affecting the hydrophobic core would have influenced the binding of both molecules, as reported in other studies.[16]

Substitutions introduced into human LAMP-1 affected antibody recognition in at least two distinct ways. The first was by directly disrupting the interaction with the Fab via the introduction of a mutation in the epitope. The second affected antigen structure in such a way that it distorted the epitope and prevented Fab binding, with longer range effects at distances typically greater than 5 Å from the interface with the antibody. By distinguishing the amino acids present on the surface of LAMP-1 from those embedded in the hydrophobic core of the antigen, the structural

information helped discriminate these two types of effects and allowed rapid identification of the 'functional epitopes' when adopting the terminology previously proposed by van Regenmortel.[36] The DMS data were compared with the crystal structure of the complex between Fab B and the human LAMP-1 first luminal domain or with the AlphaFold model of the full human LAMP-1 protein for both Fabs. This showed that the two Fabs have a conformational epitope with amino acids spread in the primary sequence of the antigen which assemble into a continuous and discrete entity on the surface of the antigen, strongly suggesting the accuracy of the epitope.

A close examination of the nature of the affected positions in LAMP-1 and tolerated substitutions led to definition of DMS patterns governing the outlines of the epitope. We first chose to focus on positions with an index greater than 15, i.e., those for which a very small number of substitutions were tolerated. The threshold of 15 is quite stringent, but allows the selection of positions for which conservative mutations are sometimes tolerated. All five Fab A positions and six of the eight Fab B positions with such a high index are exposed on the LAMP-1 surface and belong to the functional epitopes. However, two residues did not follow this rule in the Fab B DMS map. C155 displayed an index of 17, mirroring the index of 11 for C191, showing that the abolition of the disulfide bridge between these two cysteines of the first luminal domain was very unfavorable for the recognition of mAb B. Each LAMP-1 luminal domain has four cysteine residues that form two disulfide bonds and are conserved among the family of lysosome-associated membrane proteins LAMP-1, 2 and 3 and across species.[23] These disulfide bridges likely play a critical role in the overall assembly and stability of these proteins, which explains why disruption of the C155-C191 bond alters the epitope while being buried in the protein core. Remarkably, Q176 is the only hydrophilic amino acid highly intolerant to substitution and being buried. With A177, it is located at the base of the $Y_{178}L_{179}S_{180}$ triplet in the epitope and can be considered as scaffolding residues having no direct interaction with the antigen. While the C155/C191 cysteine pairs could be a priori discarded from the Fab B epitope, Q176 was ruled out due to the antigen structure, whether experimental or predicted.

Among positions with an index between five and 14, all hydrophilic residues were part of the functional epitopes (N107, K151, S180 and S185 for Fab B; R246, T263, S280, T283, T284, T308 and D312 for Fab A). By contrast, most hydrophobic positions with intermediate index values were buried (Fig. 5A), with few exceptions. Y178, L179 and F184 are the three hydrophobic residues exposed on the surface of the Fab B epitope (vs 15 buried residues). Similarly, Y244, L256 and L310 belong to the Fab A functional epitope, while the 18 other hydrophobic residues with

an intermediate index are buried. Altogether, the nature of the amino acids and tolerated substitutions appear to be good predictors of their contribution to the functional epitope, but 3D information was decisive in precisely sorting buried from exposed residues and finely mapping the epitope with unprecedented efficiency in terms of time and resources.

Beyond the determination of the functional epitope of the two antibodies, this study sheds light on the structural determinants of their inter-species cross-reactivity. Indeed, identification of the amino acids that diverge between human LAMP-1 epitopes and their monkey or mouse ortholog provides a better understanding of the differences in affinity. It is particularly interesting to note that a single substitution, such as G187E in the mAb B epitope, appears to be responsible for its lack of cross-reactivity with cynomolgus LAMP-1, with a complete loss of recognition in BLI at the maximum tested concentration (Fig. 1). G187 is located in a loop pointing to the Fab B light chain. According to the crystal structure, a glutamate residue at this position would cause a steric clash with Y32 of the antibody light chain.

Two substitutions present in cynomolgus orthologous form, T283S and I309T, are located in the Fab A epitope. DMS shows a high enrichment score for I309T, suggesting that it substantially impacts Fab A binding to cynoLAMP1. In sharp contrast with I309T, the T283S mutation is tolerated, unlike several other substitutions on position T283, which are found to be deleterious. These finding provide new examples of the fine specificity of antibodies. This is in line with a large body of literature that shows that a small number of substitutions at key epitope locations can significantly or completely abolish antigen-antibody binding. The high specificity of antibodies can even be used to distinguish two isoforms of the same protein in the same species.[37] This is also particularly well documented for many SARS-CoV antibodies, which have seen their binding abolished because of point mutations present in the various emerging variants.[38, 39] The existence of single mutations critical for antibody recognition has also been observed for several other types of antigens,[40-42] including tumor antigens.[43] When a few amino acids differ between antigens, the present methodology is particularly useful in identifying rapidly and precisely those responsible for the difference in affinity. Not surprisingly, the presence of numerous substitutions within the two epitopes in murine LAMP-1 results in a total loss of affinity *in vitro*, making it difficult to precisely evaluate the contribution of each substitution.

This report demonstrates how much combining orthogonal approaches such as DMS and deep learning-based structural modeling strengthens the accurate determination of epitopes. Applying a similar methodology to paratope mapping would provide complementary information and define the interacting domains on the two partners. However, there is an additional challenge, as loop structure prediction by AlphaFold is still inaccurate for loops longer than 10 residues.[44] Accurate prediction of antibody CDR loop structure is the subject of intense research in a fast-paced environment[45-47] and raises much hope in the antibody community. Similarly, despite the amazing progress observed over the past years,[8] most docking models of antibody/antigen complexes still have low success rates and need further development.[48, 49] In the meantime, efficient approaches to experimental determination of the functional epitope/paratope pair augmented by potent protein structure prediction tools will remain the basis for high-throughput antibody engineering.

Materials & Methods

mAb A and mAb B

mAbs A and B are full-length anti-LAMP1 IgG antibodies (mouse and human IgG1, respectively) produced in-house by transient transfection of human HEK293 FreeStyle™ cells (Thermo Fisher).

Affinity measurement by biolayer interferometry

Binding kinetics were determined using an Octet RED96 instrument (Molecular Devices, San Jose, USA). Anti-hIgG Fc Capture (AHC) Biosensors were loaded with mAb A or mAb B IgG molecules (25 nM) for 60 seconds. After baseline determination using kinetic buffer (phosphate-buffered saline (PBS), bovine serum albumin 0.1% (w/v) and Tween 20 0.02% (v/v)), association of human LAMP-1 or cynomolgus LAMP-1 was measured at different concentrations (200 nM to 6.25 nM) for 300 seconds before dissociation in kinetic buffer. Data of the control without antigen were subtracted from all binding curves and binding kinetics were fitted using a global 1:1 Langmuir binding model.

Fab and LAMP-1 protein production

The Fab heavy and light chain sequences were cloned into the AbVec2.0-IGHG1 and AbVec1.1-IGLC plasmids, respectively. [50] In both constructs, the Fc was replaced by a polyhistidine tag. Fab B was also fused to a V5 tag (GKPIPNPLLGLDST) at the C-terminus of the light chain. Human HEK293 FreeStyle™ cells (Thermo Fisher) (2.5×10^6 cells/mL) were transiently co-transfected in 100 mL of FreeStyle™ medium (Thermo Fisher) by adding 150 µg of each plasmid and 1.8 mL of linear polyethylenimine (0.5 mg/mL, Polysciences). Cells were incubated for 7 days at 37°C, 120 rpm, 8% CO_2. The culture supernatant was purified using HisTrap Excel columns (GE Healthcare). Size-exclusion chromatography was performed using Sephacryl-S-200 HR columns (Sigma) with PBS. After purification, Fab A was biotinylated using the EZ-link SulfoNHS-LC-Biotin biotinylation kit (Ref A39257, Thermo Fisher).

Nucleic acid sequences coding for LAMP-1 extracellular domains fused to a polyhistidine tag at its C-terminus were cloned into mammalian expression plasmids under the CMV enhancer/promoter and the SV40 polyA signal. Resulting plasmids were transfected into FreeStyle™ 293-F cells using FreeStyle™ MAX 293 Expression System according to the manufacturer's instructions (Thermo Fisher Scientific; K9000-10). LAMP-1 proteins were purified by immobilized metal affinity chromatography (Chelating Sepharose, 17-0575-01 GE Healthcare) and stored in PBS after concentration and buffer exchange (Sephadex G-25 column, GE Healthcare).

The first luminal domain of the human LAMP-1 (LAMP-1 29-195) sequence fused to a polyhistidine tag and a thrombin recognition site at its N-terminus was cloned in pET-48b(+) vector (Novagen). The resulting plasmid was transformed into SHuffle® T7 Competent *E. coli* cells (New England Biolabs). Protein was purified by immobilized metal affinity chromatography (Chelating Sepharose, 17-0575-01 GE Healthcare) before removal of the thioredoxin domain and polyhistidine tag by thrombin cleavage and stored in PBS after concentration and buffer exchange (Sephadex G-25 column, GE Healthcare).

Libraries

Five human LAMP-1 libraries with single amino acid mutations were constructed using SOE-PCR and NNK codons. Library 1 corresponds to amino acids 29 to 99, library 2: amino acids 100 to 194, library 3: amino acids 195 to 226, library 4: amino acids 227 to 309 and library 5: amino acids 310 to 382. Following the mutagenesis, genes were constructed and amplified by SOE-PCR. Preparation of competent yeast cells EBY100 (ATCC® MYA-4941) and library transformation were performed according to Benatuil *et al*.[51] Libraries were generated by gap repair cloning in yeast cells electroporated with 1 µg of digested vector and a molar ratio of 1.5:1 (library genes/digested vector). Transformation efficiency was determined by plating serial dilutions on selective agar plates. Each library contained at least 10^6 clones. Transformed cells were cultured for two days in SD-CAA medium (6.7 g/L yeast nitrogen base without casamino acids, 20 g/L glucose, 5 g/L casamino acids, 100 mM sodium phosphate, pH 6.0), at 30°C with shaking. After a passage to an OD_{600} of 0.25, cells were grown at 30°C until OD_{600} 0.5–1.0 and re-suspended in 50 mL of SG-CAA for induction and incubated at 20°C.[52]

Flow cytometry sorting

After induction, yeast cells displaying the libraries were incubated in 10 mL of a PBSF solution containing 1 nM of Fab A (biotin-labeled) and 15 nM of Fab B-V5 tag. Cells were incubated with shaking for 2 hours at 20°C. Cells were washed with ice-cold PBSF before incubation with PE-conjugated streptavidin and anti-V5 tag/APC-conjugated antibody in PBSF, for 15 minutes on ice. Cells were washed with 1 mL of ice-cold PBSF and sorted with a BD FACS Aria™ III cytometer (Becton Dickinson, Franklin Lake, USA) using BD FACSdiva™ software. Cells with decreased binding for one of the Fabs while retaining binding for the other one were sorted. Library 3 did not contain such a population and so was not sorted. After sorting, cells were cultured at 30°C for two days in SD-CAA.

NGS sequencing and data analysis

The Zymoprep Yeast Plasmid Miniprep II kit (Zymo Research, Irvine, USA) was used to extract plasmids from the sorted population. Regions of interest were amplified in a first PCR step and then adapters and barcodes needed for Illumina sequencing were added in a second PCR step. NGS was performed with an Illumina MiSeq (Illumina, San Diego, USA) device (2x300bp, v3 kit 600 cycles) with at least one million reads per population. The first steps of data analysis were performed on the Galaxy platform. Reads were first paired using the Pear function. Then the reads with an unexpected length were eliminated using Filter FASTQ function. The following analysis steps were performed using RStudio software and eliminated sequences containing more than one mutation compared to the parental antigen sequence. Reads presenting a quality under 30 were also eliminated. After DNA translation, identical sequences were grouped and counted in order to calculate the mono-mutant enrichment ratio in each sorted population compared to the initial population.

Crystallization

The complex between the first luminal domain of LAMP-1 (LAMP-1 29-195 produced in *E. coli*) and Fab B was concentrated to 12 mg/ml in 10 mM phosphate buffer saline pH 7. Crystallization was done by vapor diffusion using the sitting drop method. Crystals were obtained in 20% (w/v) polyethylene glycol 3350, 200 mM NaF. 25% (v/v) ethylene glycol was included as cryoprotectant prior to freezing. Datasets were collected at beamline ID29 from the synchrotron ESRF (European Synchrotron Radiation Facility) on a Pilatus 6M at wavelength 0.976251 Å. The crystals belong to the space group C2 and diffracted to 2.37 Å. Data were processed using autoproc from GlobalPhasing[53] which relies on the XDS[54] and Aimless[55] programs. Final processing statistics are listed in Supp. table 1.

Structure determination

A model of the constant domain of the Fab was built using the structure 4JG0 as reference. This structure was obtained from the Protein Data Bank (http://www.rcsb.org/pdb/home/home.do). A model of the variable domain was constructed in Maestro (Schrödinger, Inc.: Portland, OR, 2012). Molecular replacement was carried out using Phaser[56] of the CCP4 suite[57] and two complexes LAMP-1/Fab could be constructed in the asymmetric unit. The structure was refined at 2.37 Å by doing multiple cycles of Buster (Buster-TNT 2.11.5, Global Phasing Ltd) followed by

manual corrections in COOT[58] to a final Rfree of 0.261 and Rfactor 0.226. Refinement statistics are available in Supp. table 1.

The Alphafold 2 model of human LAMP-1 is available at: https://alphafold.ebi.ac.uk/entry/P11279.

Acknowledgements

The authors wish to acknowledge Fabienne Soubrier, Cecile Capdevila, Francis Duffieux, Alain Dupuy and Alexey Rakfor the key contributions they made to this work. LAMP-1 proteins and domains as well as Fab B were cloned by FS, produced by CC, and purified by FD (Sanofi, LMR, France). The authors also thank Raphaël Sierocki (Deeptope SAS) for help with the NGS data analysis scripts and useful discussions about the epitope mapping processes.

Disclosure of potential conflicts of interest

TP, MM and EV are Sanofi employees and may hold shares and/or stock options in the company. The authors have no additional conflict of interest.

Funding

This work was financially supported by Sanofi (Collaboration agreement Sanofi/CEA). TP was supported by a CIFRE fellowship (No. 2018/0802) funded in part by ANRT (National Association for Research and Technology) on the behalf of the French Ministry of Education and Research and in part by Sanofi.

Author Contribution Statement

TP, MM and SD contributed to data collection, TP, MM and HN contributed to data analysis, data interpretation was performed by TP, HN, MM, EV and BM. BM, EV and HN contributed to the writing and design of the study.

All authors have approved the final version of this manuscript and agreed both to be personally accountable for their contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even that in

which the author was not personally involved, are appropriately investigated, and resolved and the resolution documented in the literature.

References

1. Stave JW, Lindpaintner K. Antibody and antigen contact residues define epitope and paratope size and structure. J Immunol 2013; 191:1428-35.
2. Nilvebrant J, Rockberg J. An Introduction to Epitope Mapping. Methods Mol Biol 2018; 1785:1-10.
3. Toride King M, Brooks CL. Epitope Mapping of Antibody-Antigen Interactions with X-Ray Crystallography. Methods Mol Biol 2018; 1785:13-27.
4. Qi H, Ma M, Hu C, Xu ZW, Wu FL, Wang N, Lai DY, Li Y, Zhang H, Jiang HW, et al. Antibody Binding Epitope Mapping (AbMap) of Hundred Antibodies in a Single Run. Mol Cell Proteomics 2021; 20:100059.
5. Najar TA, Khare S, Pandey R, Gupta SK, Varadarajan R. Mapping Protein Binding Sites and Conformational Epitopes Using Cysteine Labeling and Yeast Surface Display. Structure 2017; 25:395-406.
6. Infante YC, Pupo A, Rojas G. A combinatorial mutagenesis approach for functional epitope mapping on phage-displayed target antigen: application to antibodies against epidermal growth factor. mAbs 2014; 6:637-48.
7. Puchades C, Kukrer B, Diefenbach O, Sneekes-Vriese E, Juraszek J, Koudstaal W, Apetri A. Epitope mapping of diverse influenza Hemagglutinin drug candidates using HDX-MS. Scientific reports 2019; 9:4735.
8. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, Žídek A, Bates R, Blackwell S, Yim J, et al. Protein complex prediction with AlphaFold-Multimer. bioRxiv 2022:2021.10.04.463034.
9. Bourquard T, Musnier A, Puard V, Tahir S, Ayoub MA, Jullian Y, Boulo T, Gallay N, Watier H, Bruneau G, et al. MAbTope: A Method for Improved Epitope Mapping. J Immunol 2018; 201:3096-105.
10. Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, Wang J, Cong Q, Kinch LN, Schaeffer RD, et al. Accurate prediction of protein structures and interactions using a three-track neural network. Science 2021; 373:871-6.
11. Heyne M, Shirian J, Cohen I, Peleg Y, Radisky ES, Papo N, Shifman JM. Climbing Up and Down Binding Landscapes through Deep Mutational Scanning of Three Homologous Protein-Protein Complexes. J Am Chem Soc 2021; 143:17261-75.
12. Dunham AS, Beltrao P. Exploring amino acid functions in a deep mutational landscape. Mol Syst Biol 2021; 17:e10305.
13. Doolan KM, Colby DW. Conformation-dependent epitopes recognized by prion protein antibodies probed using mutational scanning and deep sequencing. J Mol Biol 2015; 427:328-40.
14. Van Blarcom T, Rossi A, Foletti D, Sundar P, Pitts S, Bee C, Witt JM, Melton Z, Hasa-Moreno A, Shaughnessy LJJomb. Precise and efficient antibody epitope determination through library design, yeast display and next-generation sequencing. J Mol Biol 2015; 427:1513-34.
15. Medina-Cucurella AV, Zhu Y, Bowen SJ, Bergeron LM, Whitehead TA. Pro region engineering of nerve growth factor by deep mutational scanning enables a yeast platform for conformational epitope mapping of anti-NGF monoclonal antibodies. Biotechnol Bioeng 2018; 115:1925-37.
16. Sierocki R, Jneid B, Orsini Delgado ML, Plaisance M, Maillere B, Nozach H, Simon S. An antibody targeting type III secretion system induces broad protection against Salmonella and Shigella infections. PLoS Negl Trop Dis 2021; 15:e0009231.
17. Eguia RT, Crawford KHD, Stevens-Ayers T, Kelnhofer-Millevolte L, Greninger AL, Englund JA, Boeckh MJ, Bloom JD. A human coronavirus evolves antigenically to escape antibody immunity. PLoS pathogens 2021; 17:e1009453.
18. Sourisseau M, Lawrence DJP, Schwarz MC, Storrs CH, Veit EC, Bloom JD, Evans MJ. Deep Mutational Scanning Comprehensively Maps How Zika Envelope Protein Mutations Affect Viral Growth and Antibody Escape. J Virol 2019; 93:e01291-19.
19. Dingens AS, Haddox HK, Overbaugh J, Bloom JD. Comprehensive Mapping of HIV-1 Escape from a Broadly Neutralizing Antibody. Cell Host Microbe 2017; 21:777-87 e4.
20. Wu NC, Xie J, Zheng T, Nycholat CM, Grande G, Paulson JC, Lerner RA, Wilson IAJCh, microbe. Diversity of functionally permissive sequences in the receptor-binding site of influenza hemagglutinin. Cell Host & Microbe 2017; 21:742-53. e8.
21. Lee JM, Huddleston J, Doud MB, Hooper KA, Wu NC, Bedford T, Bloom JD. Deep mutational scanning of hemagglutinin helps predict evolutionary fates of human H3N2 influenza variants. Proc Natl Acad Sci U S A 2018; 115:E8276-E85.
22. Starr TN, Czudnochowski N, Liu ZM, Zatta F, Park YJ, Addetia A, Pinto D, Beltramello M, Hernandez P, Greaney AJ, et al. SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape. Nature 2021; 597:97-+.

23. Terasawa K, Tomabechi Y, Ikeda M, Ehara H, Kukimoto-Niino M, Wakiyama M, Podyma-Inoue KA, Rajapakshe AR, Watabe T, Shirouzu M, et al. Lysosome-associated membrane proteins-1 and -2 (LAMP-1 and LAMP-2) assemble via distinct modes. Biochem Biophys Res Commun 2016; 479:489-95.
24. Cameron B, Dabdoubi T, Berthou-Soulie L, Gagnaire M, Arnould I, Severac A, Soubrier F, Morales J, Leighton PA, Harriman W, et al. Complementary epitopes and favorable developability of monoclonal anti-LAMP1 antibodies generated using two transgenic animal platforms. PLoS One 2020; 15:e0235815.
25. Agarwal AK, Gude RP, Kalraiya RD. Regulation of melanoma metastasis to lungs by cell surface Lysosome Associated Membrane Protein-1 (LAMP1) via galectin-3. Biochem Biophys Res Commun 2014; 449:332-7.
26. Alessandrini F, Pezze L, Ciribilli Y. LAMPs: Shedding light on cancer biology. Semin Oncol 2017; 44:239-53.
27. Wilke S, Krausze J, Bussow K. Crystal structure of the conserved domain of the DC lysosomal associated membrane protein: implications for the lysosomal glycocalyx. BMC Biol 2012; 10:62.
28. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Zidek A, Potapenko A, et al. Highly accurate protein structure prediction with AlphaFold. Nature 2021; 596:583-9.
29. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. Nat Meth 2022; 19:679-82.
30. Deng X, Storz U, Doranz BJ. Enhancing antibody patent protection using epitope mapping information. mAbs 2018; 10:204-9.
31. Kowalsky CA, Faber MS, Nath A, Dann HE, Kelly VW, Liu L, Shanker P, Wagner EK, Maynard JA, Chan C, et al. Rapid Fine Conformational Epitope Mapping Using Comprehensive Mutagenesis and Deep Sequencing. J Biol Chem 2015; 290:26457-U150.
32. Tran MH, Schoeder CT, Schey KL, Meiler J. Computational Structure Prediction for Antibody-Antigen Complexes From Hydrogen-Deuterium Exchange Mass Spectrometry: Challenges and Outlook. Frontiers in immunology 2022; 13:859964.
33. Wittrup KD. Protein engineering by cell-surface display. Curr Opin Biotechnol 2001; 12:395-9.
34. Hamilton SR, Bobrowicz P, Bobrowicz B, Davidson RC, Li H, Mitchell T, Nett JH, Rausch S, Stadheim TA, Wischnewski H, et al. Production of complex human glycoproteins in yeast. Science 2003; 301:1244-6.
35. Fukuda M. Lysosomal membrane glycoproteins. Structure, biosynthesis, and intracellular trafficking. J Biol Chem 1991; 266:21327-30.
36. Van Regenmortel MH. What is a B-cell epitope? Methods Mol Biol 2009; 524:3-20.
37. Orsini Delgado ML, Avril A, Prigent J, Dano J, Rouaix A, Worbs S, Dorner BG, Rougeaux C, Becher F, Fenaille F, et al. Ricin Antibodies' Neutralizing Capacity against Different Ricin Isoforms and Cultivars. Toxins 2021; 13:100.
38. Starr TN, Greaney AJ, Addetia A, Hannon WW, Choudhary MC, Dingens AS, Li JZ, Bloom JD. Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. Science 2021; 371:850-4.
39. Focosi D, McConnell S, Casadevall A, Cappello E, Valdiserra G, Tuccori M. Monoclonal antibody therapies against SARS-CoV-2. Lancet Infect Dis 2022.
40. Doud MB, Lee JM, Bloom JD. How single mutations affect viral escape from broad and narrow antibodies to H1 influenza hemagglutinin. Nature communications 2018; 9:1386.
41. Keeffe JR, Van Rompay KKA, Olsen PC, Wang Q, Gazumyan A, Azzopardi SA, Schaefer-Babajew D, Lee YE, Stuart JB, Singapuri A, et al. A Combination of Two Human Monoclonal Antibodies Prevents Zika Virus Escape Mutations in Non-human Primates. Cell Rep 2018; 25:1385-94 e7.
42. Dao T, Mun S, Korontsvit T, Khan AG, Pohl MA, White T, Klatt MG, Andrew D, Lorenz IC, Scheinberg DA. A TCR mimic monoclonal antibody for the HPV-16 E7-epitope p11-19/HLA-A*02:01 complex. PLoS One 2022; 17:e0265534.
43. Arena S, Bellosillo B, Siravegna G, Martinez A, Canadas I, Lazzari L, Ferruz N, Russo M, Misale S, Gonzalez I, et al. Emergence of Multiple EGFR Extracellular Mutations during Cetuximab Treatment in Colorectal Cancer. Clinical cancer research : an official journal of the American Association for Cancer Research 2015; 21:2157-66.
44. Stevens AO, He Y. Benchmarking the Accuracy of AlphaFold 2 in Loop Structure Prediction. Biomolecules 2022; 12.
45. Ruffolo JA, Chu L-S, Mahajan SP, Gray JJ. Fast, accurate antibody structure prediction from deep learning on massive set of natural antibodies. bioRxiv 2022:2022.04.20.488972.
46. Liang T, Jiang C, Yuan J, Othman Y, Xie X-Q, Feng Z. Differential performance of RoseTTAFold in antibody modeling. Brief Bioinform 2022:bbac152.
47. Abanades B, Georges G, Bujotzek A, Deane CM. ABlooper: Fast accurate antibody CDR loop structure prediction with accuracy estimation. Bioinformatics 2022.

48. Yin R, Feng BY, Varshney A, Pierce BG. Benchmarking AlphaFold for protein complex modeling reveals accuracy determinants. Protein Sci 2022; 31:e4379.
49. Xu Z, Davila A, Wilamowski J, Teraguchi S, Standley DM. Improved Antibody-Specific Epitope Prediction Using AlphaFold and AbAdapt**. ChemBioChem 2022; n/a:e202200303.
50. Smith K, Garman L, Wrammert J, Zheng NY, Capra JD, Ahmed R, Wilson PC. Rapid generation of fully human monoclonal antibodies specific to a vaccinating antigen. Nat Protoc 2009; 4:372-84.
51. Benatuil L, Perez JM, Belk J, Hsieh CM. An improved yeast transformation method for the generation of very large human antibody libraries. Protein Eng Des Sel 2010; 23:155-9.
52. Laroche A, Orsini Delgado ML, Chalopin B, Cuniasse P, Dubois S, Sierocki R, Gallais F, Debroas S, Bellanger L, Simon S, et al. Deep mutational engineering of broadly-neutralizing nanobodies accommodating SARS-CoV-1 and 2 antigenic drift. mAbs 2022; 14:2076775.
53. Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, Womack T, Bricogne G. Data processing and analysis with the autoPROC toolbox. Acta Crystallogr D Biol Crystallogr 2011; 67:293-302.
54. Kabsch W. Xds. Acta Crystallogr D Biol Crystallogr 2010; 66:125-32.
55. Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr 2013; 69:1204-14.
56. Storoni LC, McCoy AJ, Read RJ. Likelihood-enhanced fast rotation functions. Acta Crystallogr D Biol Crystallogr 2004; 60:432-8.
57. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr 2011; 67:235-42.
58. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 2010; 66:486-501.

Figure 1: Cross-reactivity of LAMP-1 antibodies determined by bio-layer interferometry

BLI sensorgrams showing the binding of human LAMP-1 (top panel) and cynomolgus LAMP-1 (bottom panel) to mAbs A and B immobilized on AHC biosensor tips. Data are shown as colored lines at different concentrations of human or cynomolgus LAMP-1 (From 200 nM to 6,25 nM using two-fold serial dilutions). Red lines are the best fit of the data.

Figure 2: Deep Mutational Scanning of Fab A and Fab B binding to the extracellular domain of hLAMP-1

(A) General principle of functional screening by yeast surface display. Five DNA libraries of hLAMP-1 harboring a single mutation (each corresponding to one of the five regions encompassing the sequence of the extracellular domain of hLAMP-1) **(B)** were transformed into yeast using gap repair recombination. **(C)** Bivariate flow cytometry analysis of libraries. Cells were simultaneously incubated with Fab A and Fab B and labeled with secondary reporters before FACS analysis. Selected cells (red gates) were sorted and sequenced with Illumina Deep Sequencing.

Figure 3: Deep Mutational Scanning epitope maps of Fabs A and B

NGS-based heatmaps represent the enrichment scores of hLAMP-1 single mutants after functional sorting in FACS using Fab B **(A)** and Fab A **(B)** as bait. Enrichment score is a log2 function of the frequency fold-change between sorted and unsorted hLAMP-1 yeast populations for a given amino acid substitution. The corresponding table is colored in red for enriched mutations. The index is set as the number of substitutions with an enrichment score higher than 2.

Figure 4: Comparison of functional and structural epitopes of Fab B

(A) Co-crystal structure of the first luminal domain of hLAMP-1 in complex with Fab B. Ribbon diagram illustrating the heavy chain (green) and light chain (blue) of the Fab. Amino acids within a 4.5 Å range from Fab B are colored in yellow on the surface representation of the first luminal domain of hLAMP-1 (grey). **(B)** Graphical view of co-crystal structure of the first luminal domain of hLAMP-1 in complex with Fab B colored with DMS Epitope Mapping data. Amino acids with DMS scores above 10 are marked in red and amino acids with DMS scores between 5 and 9 are shown in yellow (with a surface representation or with spheres on the ribbon representation). **(C)** Representation of the AlphaFold model of the first luminal domain of hLAMP-1. Residues included in the DMS epitope are colored pink.

Figure 5: Functional Epitopes of mAb A and mAb B in human LAMP1 and orthologs

(A) & (B) Functional epitopes of mAb A and mAb B on the AlphaFold model of the full extracellular domain of human LAMP-1 are represented in blue and pink, respectively. **(C)** Molecular surface representation of epitope conservation in the cynomolgus and mouse sequences. The surface area is colored blue or pink if the residue is conserved between the 2 species for Fab B and A, respectively, and orange if the epitope has different residues in the two antigens.

Figure 6: Summary of the steps of the DMS and functional epitope visualization approach

For each library, cells with a lower binding of the considered Fab are sorted by FACS and the sequence of the corresponding clones determine by NGS. A first step of analysis determines the enrichments of the mutations abolishing or reducing the recognition of the Fab. Based on the models established by Alphafold2, the buried amino acids are excluded and a three-dimensional representation of the functional epitope is established.

Supplementary Figure 1: Complete Fab B DMS heatmaps for libraries 1 and 2

NGS-based heatmap representing enrichment values of hLAMP-1 single mutant after functional sorting in FACS. The index is set as the number of substitutions with an enrichment score higher than 2. Amino acids differing between hLAMP-1 and cLAMP-1 are indicated in the bottom line. Functional epitope is highlighted in red in black frames on the top line.

Supplementary Figure 2: Complete Fab A DMS heatmaps for libraries 3 and 4

NGS-based heatmap representing enrichment values of hLAMP-1 single mutant after functional sorting in FACS. The index is set as the number of substitutions with an enrichment score higher than 2. Amino acids differing between hLAMP-1 and cLAMP-1 are indicated in the bottom line. Functional epitope is highlighted in blue in black frames on the top line.

Supplementary Figure 3: Predicted structure of hLAMP-1 and library 3 labelling in FACS

(A) Ribbon representation of the AlphaFold model of human LAMP-1. pLDDT confidence scores of the model are color coded from red to blue (low to high confidence). **(B)** Predicted Aligned Error (PAE) plDBot of the AlphaFold model of hLAMP-1. The color at position (x, y) indicates AlphaFold's expected position error. pLDDT ranges from 73 to 86 for a.a. 82-86. **(C)** Bivariate flow cytometry analysis of library 3

Supplementary Figure 4: Localization of N-glycosylation sites on human LAMP1 in the vicinity of functional epitopes of Fabs A and B

(A) Functional epitopes of Fabs A and B are represented in blue and pink on the AlphaFold model of full extracellular domain of human LAMP-1. Potential N-glycosylation sites are colored in green. **(B)** The 4 potential N-glycosylation sites at N37, N107, N181 and N84 (blue sticks) lying at the periphery of the epitope are highlighted on the purple ribbon of the crystal structure of the first luminal domain of hLAMP-1 in complex with Fab B (heavy chain in cyan and light chain in green). The orientation of putative N-glycans is marked with black arrows.

Supplementary Table 1 Data collection and refinement statistics for the structure of the first luminal domain of LAMP-1 in complex with Fab B (PDB code 8ATH)

Data collection	
Space group	C2
Cell dimensions	
a, *b*, *c* (Å)	149.93, 93.68, 108.01
α, β, γ (°)	90, 115.93, 90
Resolution (Å) *	72.08 – 2.37 (2.65 – 2.37)
R_{merge}	0.047 (0.413)
Mean (I / σI)	14.4 (2.8)
Completeness (%)	99.1 (99.3)
Redundancy	3.3 (3.3)
Refinement	
Resolution (Å)	2.37
No. reflections	54389
R_{work} / R_{free}	0.251 / 0.290
No. atoms	
Protein	8703
Water	269
R.m.s. deviations	
Bond lengths (Å)	0.01
Bond angles (°)	1.04

*Highest resolution shell is shown in parentheses.

Alt text for equal opportunity to all readers, including those with visual or print impairments.

Figure 1: Cross-reactivity of LAMP-1 antibodies determined by bio-layer interferometry

Scheme representing the BLI sensorgrams with association and dissociation of the antigen

Figure 2: Deep Mutational Scanning of Fab A and Fab B binding to the extracellular domain of hLAMP-1

(A) Scheme illustrating the generation of DMS libraries, their cloning in plasmid and transfection of yeast cells. Scheme representing yeast cells each expressing hLAMP1 mutants in interaction with Fab A and Fab B.**(B)** Scheme representing luminal domains of hLAMP1 and localization of the five DMS libraries **(C)** Cytometry dot-plot figures obtained from the analysis of yeast expressing DMS libraries of hLAMP1.

Figure 3: Deep Mutational Scanning epitope maps of Fabs A and B

Table indicating for each amino acid position (in column) and each substitution (in rows) the enrichment of the corresponding mutation in the DMS setup. Corresponding table is coloured in red for enriched mutations.

Figure 4: Comparison of functional and structural epitopes of Fab B

(A) Graphical representation of the co-crystal structure of the first luminal domain of hLAMP-1 in complex with Fab B. **(B)** Graphical view of co-crystal structure of the first luminal domain of hLAMP-1 in complex with Fab B colored with DMS Epitope Mapping data. **(C)** Representation of the AlphaFold model of the first luminal domain of hLAMP-1. Residues included in the DMS epitope are colored pink.

Figure 5: Functional Epitopes of mAb A and mAb B in human LAMP1 and orthologs

(A) Table indicating the localization of residues involved in the functional epitope. The table also indicate the buried positions and the amino acid differences between human, Macaca and Mouse antigens. **(B)** Graphical view of the hLAMP-1 AlphaFold model with DMS epitopes of mAb A and mAb B (in blue and pink, respectively) **(C)** Graphical view of the three species LAMP-1 AlphaFold models with DMS epitopes of mAb A and mAb B (in blue and pink, respectively). Residues diverging from the human epitope are indicated in yellow.

Figure 6: Summary of the steps of the DMS and functional epitope visualization approach

Graphic summary of the steps used to define the functional epitope by DMS. After selection of cells in a dot-plot FACS analysis, the cells are sorted and the corresponding LAMP-1 cDNA sequenced by NGS. Alphafold 2 model is then used to filter out buried residues and to draw a representation of the epitope on the antigen surface.

Supplementary Figure 1: Complete Fab B DMS heatmaps for libraries 1 and 2

Table indicating for each amino acid position (in column) and each substitution (in rows) the enrichment of the corresponding mutation in the DMS setup. Corresponding table is coloured in red for enriched mutations. Functional epitope is highlighted in red in black frames on the top line.

Supplementary Figure 2: Complete Fab A DMS heatmaps for libraries 3 and 4

Table indicating for each amino acid position (in column) and each substitution (in rows) the enrichment of the corresponding mutation in the DMS setup. Corresponding table is coloured in red for enriched mutations. Functional epitope is highlighted in blue in black frames on the top line.

Supplementary Figure 3: Predicted structure of hLAMP-1

(A) Graphical representation of the hLAMP-1 AlphaFold model colored in function of the model confidence. pLDDT confidence scores of the model are color coded from red to blue (low to high confidence). **(B)** Matrix representing the expected position error in angstrom for two given positions.

Supplementary Figure 4: Localization of N-glycosylation sites on human LAMP1 in the vicinity of functional epitopes of Fabs A and B

(A) Graphical view of the hLAMP-1 AlphaFold model with DMS epitopes of mAb A and mAb B (in blue and pink, respectively) and potential N-glycosylation sites. **(B)** Graphical view of co-crystal structure of the first luminal domain of hLAMP-1 in complex with Fab B and potential N-glycosylation sites (in black).