



HAL
open science

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

Tiphanie Pruvost, Magali Mathieu, Steven Dubois, Bernard Maillère,
Emmanuelle Vigne, Hervé Nozach

► **To cite this version:**

Tiphanie Pruvost, Magali Mathieu, Steven Dubois, Bernard Maillère, Emmanuelle Vigne, et al.. Deciphering cross-species reactivity of LAMP-1 antibodies using deep mutational epitope mapping and AlphaFold. *mAbs*, 2023, 15 (1), pp.2175311. 10.1080/19420862.2023.2175311 . cea-03993824

HAL Id: cea-03993824

<https://cea.hal.science/cea-03993824v1>

Submitted on 17 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

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

33

34

35 **Abstract**

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

51

52 Introduction

53
54 Antibodies bind to antigens in a multitude of ways, defining a wide range of possible interacting antigenic surfaces
55 called epitopes. Generally, epitopes cover a surface of 600-900 square Å and involve one to several dozen amino
56 acids of the antigen protein.¹ These interacting amino acids are grouped in a continuous three-dimensional (3D)
57 surface and can be carried by a stretch of linear sequence or, on the contrary, scattered over the primary sequence
58 of the protein. Delineating epitopes can help to understand antibody functions or to facilitate the selection of
59 antibodies that target specific regions of the antigen.

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

72 In recent years, deep mutational scanning (DMS) approaches have considerably accelerated the pace of mutational
73 studies, which can now explore every possible single amino acid substitution in a selected protein.^{11, 12} By combining
74 high-throughput screening methods such as display techniques (e.g., phage display, yeast surface display) with deep
75 sequencing, an increasing number of studies have analyzed the mutational landscape to understand the modalities
76 of interaction between protein partners.¹² More specifically, several studies have allowed the identification of the
77 epitope of monoclonal antibodies (mAbs) against prion protein,¹³ *S. aureus* alpha toxin,¹⁴ nerve growth factor¹⁵ and

78 Salmonella antigens.¹⁶ DMS identifies the functional epitope as the key interacting amino acids that cannot be
79 replaced without causing a major loss in binding activity. By extension, DMS has recently proved useful in predicting
80 antigen mutations that allow escape from the action of therapeutic mAbs. This is a known mechanism of resistance
81 against natural or therapeutic antibodies targeting viral antigens that are subject to high selective pressure and for
82 which substitutions can reduce antibody-mediated neutralization.¹⁷ Comprehensive escape maps were notably
83 established for selected viral antigens of Zika,¹⁸ HIV,¹⁹ influenza^{20,21} and SARS-CoV-2²².

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

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

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

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

107 **Results**

108 **Identification of mutations affecting mAb binding to human LAMP-1**

109 We first determined the affinity constants of the two LAMP-1 mAbs for human LAMP-1 and their non-human
110 primate and mouse counterparts. Affinity measurements using biolayer interferometry (BLI) demonstrated the high
111 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.
112 1). The affinity of mAb A for the cynomolgus antigen was approximately 20-fold lower at 18 nM, while no binding
113 signal was observed with mAb B at the concentration of 200 nM, revealing the lack of cross-reactivity of this
114 antibody (Fig. 1). Furthermore, both antibodies failed to bind the murine antigen at the maximum tested
115 concentration (200 nM). Fluorescence-activated cell sorting (FACS) experiments confirmed the binding profile of the
116 two mAbs to human and cynoLAMP1 presented on the surface of engineered cell lines (data not shown). To link
117 these biochemical data to sequence information, we performed DMS of the human LAMP-1 antigen using the yeast
118 surface display (YSD) technique, thereby expressing mutants of the extracellular domain of human LAMP-1 on the
119 surface of yeast cells. The two luminal domains of human LAMP-1 linked together by its hinge region were anchored
120 in the yeast cell wall through a C-terminal fusion with the Aga2p protein, itself attached to Aga1p by two disulfide
121 bonds (Fig. 2A). We first demonstrated that wild-type human LAMP-1 can be expressed on the yeast surface and
122 retains binding to each antigen-binding fragment (Fab). We further showed that the two mAbs bound
123 simultaneously to LAMP-1, and therefore target independent epitopes. We then generated single mutant libraries of
124 the human LAMP-1 antigen by SOE-PCR using libraries of primers each carrying a single degenerate codon (Fig. 2A).
125 The resulting linear DNA fragment libraries were then transformed into the yeast *S. cerevisiae*. Given the large size of
126 the luminal domain of human LAMP-1 and to facilitate next-generation sequencing (NGS) procedures, we generated
127 5 sub-libraries encompassing each luminal domain and the hinge region (Fig. 2B).

128 The five yeast libraries were then simultaneously labeled with Fabs A and B for FACS sorting. Both antibodies were
129 found to bind without affecting the binding of the other molecule, indicating the independence of their two
130 epitopes. We preferred Fab to IgG to avoid experimental bias related to avidity phenomena. Fabs were used at
131 concentrations close to their K_D affinity constants, to allow the most sensitive discrimination between mutants and
132 isolate those for which a loss of recognition by either of the two Fabs is observed. Flow cytometry showed that most
133 LAMP-1 variants displayed strong fluorescence signals with both Fabs (Fig. 2C), demonstrating that corresponding
134 mutations in LAMP-1 had no effect on Fab binding. More interestingly, some subpopulations within libraries 1 and 2
135 lost binding to Fab B, but not to Fab A (red gates, upper panel of Fig. 2C). Symmetrically, some cells in libraries 4 and
136 5 expressed LAMP-1 mutants that were no longer recognized by Fab A, but still by Fab B (red gates, lower panel of
137 Fig. 2C). Finally, mutations in the hinge had no effect on the binding of either Fab (Library 3, data not shown). These
138 results not only confirmed that both Fabs bind LAMP-1 at independent epitopes, but also demonstrated that Fab A
139 binds the second luminal domain (libraries 4 and 5) and Fab B binds the first luminal domain (libraries 1 and 2). To
140 identify the amino acid substitutions responsible for the loss of recognition by either of the two Fabs, corresponding
141 cells were sorted before bulk sequencing of their human LAMP-1 mutant sequence.

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

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

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

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

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

165 **3D modeling to guide the fine determination of the functional epitope**

166

167 We decided to generate structural and 3D modeling data to distinguish positions directly involved in the epitope
168 from those affecting the overall conformation of the antigen and its folding, and ultimately refine the epitopes. We
169 first solved the crystallographic structure of the complex between Fab B and an aglycosylated form of the first
170 luminal domain of human LAMP-1 (Fig. 4A). This domain adopts the same overall β -prism fold as murine LAMP-1²³
171 and DC-Lamp3.²⁷ Most of the interaction between Fab B and LAMP-1 is mediated by amino acids in the heavy chain
172 complementarity-determining regions (CDRs). Briefly, loop 82-86 of LAMP-1 interacts with CDRH1 and the FR3 loop
173 from the Fab heavy chain. Loop 106-109 interacts with all heavy chain CDRs and loop 149-151 is in contact with
174 CDRH3 and CDRL1. Lastly, loop 178-187 contacts both CDRH1 and CDRH3, along with CDRL1 and CDRL2. All LAMP-1
175 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
176 constitute what might be termed the 'structural epitope'.

177 Next, we examined the localization of the amino acids identified by DMS for the Fab B within the crystallographic
178 structure. We colored in red the 15 positions for which at least 10 substitutions were deleterious to Fab B binding
179 and in yellow the 16 positions for which 5 to 9 substitutions were not tolerated (Fig. 4B). We observed that 11 of the
180 15 positions with an index higher than 10 were accessible to the solvent, from which 10 positions were in direct
181 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

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

186 Given the difficulty of obtaining structures of antigen-antibody complexes, we also sought to use structure models to
187 refine the DMS data. Considering the unparalleled accuracy recently demonstrated by the AlphaFold 2 algorithms,
188 we retrieved the model of the human LAMP-1 first luminal domain from AlphaFold DB (Fig. 4C). The data overlay
189 showed a very good alignment of the AlphaFold model with the crystallographic structure of the LAMP-1 domain,
190 with an root-mean-square deviation (RMSD) of atomic positions of 1.27 Å for all main-chain atoms, demonstrating
191 the good quality of the model. The only noticeable difference between the model and structure lies in the LAMP-1
192 loop 82-86, which is part of the structural epitope. However, the loop 82-86 defined by AlphaFold was found to form
193 a steric clash with the CDRH1 of Fab B when superimposed with the crystallographic structure (circled in red, Fig.
194 4B). It should be noted that this loop appears to have limited structural constraints and the model confidence scores
195 for this part of the protein are not very high (Supp. Fig. 3A). Importantly, the introduction of mutations in the 82-86
196 loop of LAMP-1 was not identified as important for Fab B recognition. Altogether, this showed that the AlphaFold
197 model of the first LAMP-1 luminal domain can be used to refine the functional epitope of Fab B and, more broadly,
198 gives confidence to use of the algorithm for predicting the 3D structure of the second LAMP-1 luminal domain for
199 which no structural data is available. On this basis, we filtered out the positions identified as buried in the second
200 luminal domain of human LAMP-1 using the AlphaFold model and mapped the functional epitope of Fab A (lower
201 table in Fig. 5A). We have summarized the different steps of the method, from cell sorting to the different data
202 processing steps in Fig. 6.

203

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

205

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

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

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

218

219 **Discussion**

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

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

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

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

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

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

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

254 Yeast cells are known to be capable of expressing a large variety of proteins on their surface.³³ It is remarkable that,
255 despite the presence of several disulfide bridges and numerous glycosylation sites, the cellular machinery of *S.*
256 *cerevisiae* allows surface expression of the full extracellular domain of the human LAMP-1 protein and its proper
257 recognition by the two studied mAbs. The probable presence of mannose-rich glycans typical of yeast glycosylation
258 machinery³⁴ in place of mammalian glycosylation patterns did not affect antibody recognition. This is consistent with
259 the successful complex formation between Fab B and the aglycosylated form of the first luminal domain of human
260 LAMP-1 used for the crystallography study. The functional and structural epitopes uncovered in this study ultimately

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

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

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

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

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

294 A close examination of the nature of the affected positions in LAMP-1 and tolerated substitutions led to definition of
295 DMS patterns governing the outlines of the epitope. We first chose to focus on positions with an index greater than
296 15, i.e., those for which a very small number of substitutions were tolerated. The threshold of 15 is quite stringent,
297 but allows the selection of positions for which conservative mutations are sometimes tolerated. All five Fab A
298 positions and six of the eight Fab B positions with such a high index are exposed on the LAMP-1 surface and belong
299 to the functional epitopes. However, two residues did not follow this rule in the Fab B DMS map. C155 displayed an
300 index of 17, mirroring the index of 11 for C191, showing that the abolition of the disulfide bridge between these two
301 cysteines of the first luminal domain was very unfavorable for the recognition of mAb B. Each LAMP-1 luminal
302 domain has four cysteine residues that form two disulfide bonds and are conserved among the family of lysosome-
303 associated membrane proteins LAMP-1, 2 and 3 and across species.²³ These disulfide bridges likely play a critical role
304 in the overall assembly and stability of these proteins, which explains why disruption of the C155-C191 bond alters
305 the epitope while being buried in the protein core. Remarkably, Q176 is the only hydrophilic amino acid highly
306 intolerant to substitution and being buried. With A177, it is located at the base of the Y₁₇₈L₁₇₉S₁₈₀ triplet in the
307 epitope and can be considered as scaffolding residues having no direct interaction with the antigen. While the
308 C155/C191 cysteine pairs could be a priori discarded from the Fab B epitope, Q176 was ruled out due to the antigen
309 structure, whether experimental or predicted.

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

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

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

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

341 This report demonstrates how much combining orthogonal approaches such as DMS and deep learning-based
342 structural modeling strengthens the accurate determination of epitopes. Applying a similar methodology to
343 paratope mapping would provide complementary information and define the interacting domains on the two
344 partners. However, there is an additional challenge, as loop structure prediction by AlphaFold is still inaccurate for
345 loops longer than 10 residues.⁴⁴ Accurate prediction of antibody CDR loop structure is the subject of intense research
346 in a fast-paced environment⁴⁵⁻⁴⁷ and raises much hope in the antibody community. Similarly, despite the amazing
347 progress observed over the past years,⁸ most docking models of antibody/antigen complexes still have low success
348 rates and need further development.^{48, 49} In the meantime, efficient approaches to experimental determination of
349 the functional epitope/paratope pair augmented by potent protein structure prediction tools will remain the basis
350 for high-throughput antibody engineering.

351

352

353 **Materials & Methods**

354 **mAb A and mAb B**

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

357 **Affinity measurement by biolayer interferometry**

358 Binding kinetics were determined using an Octet RED96 instrument (Molecular Devices, San Jose, USA). Anti-hIgG Fc
359 Capture (AHC) Biosensors were loaded with mAb A or mAb B IgG molecules (25 nM) for 60 seconds. After baseline
360 determination using kinetic buffer (phosphate-buffered saline (PBS), bovine serum albumin 0.1% (w/v) and Tween
361 20 0.02% (v/v)), association of human LAMP-1 or cynomolgus LAMP-1 was measured at different concentrations
362 (200 nM to 6.25 nM) for 300 seconds before dissociation in kinetic buffer. Data of the control without antigen were
363 subtracted from all binding curves and binding kinetics were fitted using a global 1:1 Langmuir binding model.

364 **Fab and LAMP-1 protein production**

365 The Fab heavy and light chain sequences were cloned into the AbVec2.0-IGHG1 and AbVec1.1-IGLC plasmids,
366 respectively.⁵⁰ In both constructs, the Fc was replaced by a polyhistidine tag. Fab B was also fused to a V5 tag
367 (GKPIPPLLGLDST) at the C-terminus of the light chain. Human HEK293 FreeStyle™ cells (Thermo Fisher) (2.5×10^6
368 cells/mL) were transiently co-transfected in 100 mL of FreeStyle™ medium (Thermo Fisher) by adding 150 µg of each
369 plasmid and 1.8 mL of linear polyethylenimine (0.5 mg/mL, Polysciences). Cells were incubated for 7 days at 37°C,
370 120 rpm, 8% CO₂. The culture supernatant was purified using HisTrap Excel columns (GE Healthcare). Size-exclusion
371 chromatography was performed using Sephacryl-S-200 HR columns (Sigma) with PBS. After purification, Fab A was
372 biotinylated using the EZ-link SulfoNHS-LC-Biotin biotinylation kit (Ref A39257, Thermo Fisher).

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

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

385 **Libraries**

386 Five human LAMP-1 libraries with single amino acid mutations were constructed using SOE-PCR and NNK codons.
387 Library 1 corresponds to amino acids 29 to 99, library 2: amino acids 100 to 194, library 3: amino acids 195 to 226,
388 library 4: amino acids 227 to 309 and library 5: amino acids 310 to 382. Following the mutagenesis, genes were
389 constructed and amplified by SOE-PCR. Preparation of competent yeast cells EBY100 (ATCC® MYA-4941) and library
390 transformation were performed according to Benatuil *et al.*⁵¹ Libraries were generated by gap repair cloning in yeast
391 cells electroporated with 1 µg of digested vector and a molar ratio of 1.5:1 (library genes/digested vector).
392 Transformation efficiency was determined by plating serial dilutions on selective agar plates. Each library contained
393 at least 10⁶ clones. Transformed cells were cultured for two days in SD-CAA medium (6.7 g/L yeast nitrogen base
394 without casamino acids, 20 g/L glucose, 5 g/L casamino acids, 100 mM sodium phosphate, pH 6.0), at 30°C with
395 shaking. After a passage to an OD₆₀₀ of 0.25, cells were grown at 30°C until OD₆₀₀ 0.5–1.0 and re-suspended in 50 mL
396 of SG-CAA for induction and incubated at 20°C.⁵²

397 **Flow cytometry sorting**

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

405 **NGS sequencing and data analysis**

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

416 **Crystallization**

417 The complex between the first luminal domain of LAMP-1 (LAMP-1 29-195 produced in *E. coli*) and Fab B was
418 concentrated to 12 mg/ml in 10 mM phosphate buffer saline pH 7. Crystallization was done by vapor diffusion using
419 the sitting drop method. Crystals were obtained in 20% (w/v) polyethylene glycol 3350, 200 mM NaF. 25% (v/v)
420 ethylene glycol was included as cryoprotectant prior to freezing. Datasets were collected at beamline ID29 from the
421 synchrotron ESRF (European Synchrotron Radiation Facility) on a Pilatus 6M at wavelength 0.976251 Å. The crystals
422 belong to the space group C2 and diffracted to 2.37 Å. Data were processed using autoproc from GlobalPhasing⁵³
423 which relies on the XDS⁵⁴ and Aimless⁵⁵ programs. Final processing statistics are listed in Supp. table 1.

424 **Structure determination**

425 A model of the constant domain of the Fab was built using the structure 4JG0 as reference. This structure was
426 obtained from the Protein Data Bank (<http://www.rcsb.org/pdb/home/home.do>). A model of the variable domain
427 was constructed in Maestro (Schrödinger, Inc.: Portland, OR, 2012). Molecular replacement was carried out using
428 Phaser⁵⁶ of the CCP4 suite⁵⁷ and two complexes LAMP-1/Fab could be constructed in the asymmetric unit. The
429 structure was refined at 2.37 Å by doing multiple cycles of Buster (Buster-TNT 2.11.5, Global Phasing Ltd) followed by

430 manual corrections in COOT⁵⁸ to a final Rfree of 0.261 and Rfactor 0.226. Refinement statistics are available in Supp.
431 table 1.

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

433

434

435 **Acknowledgements**

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

440

441 **Disclosure of potential conflicts of interest**

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

444

445 **Funding**

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

449 **Author Contribution Statement**

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

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

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

456

457 References

- 458 1. Stave JW, Lindpaintner K. Antibody and antigen contact residues define epitope and paratope size and
459 structure. *J Immunol* 2013; 191:1428-35.
- 460 2. Nilvebrant J, Rockberg J. An Introduction to Epitope Mapping. *Methods Mol Biol* 2018; 1785:1-10.
- 461 3. Toride King M, Brooks CL. Epitope Mapping of Antibody-Antigen Interactions with X-Ray Crystallography.
462 *Methods Mol Biol* 2018; 1785:13-27.
- 463 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
464 Mapping (AbMap) of Hundred Antibodies in a Single Run. *Mol Cell Proteomics* 2021; 20:100059.
- 465 5. Najar TA, Khare S, Pandey R, Gupta SK, Varadarajan R. Mapping Protein Binding Sites and Conformational
466 Epitopes Using Cysteine Labeling and Yeast Surface Display. *Structure* 2017; 25:395-406.
- 467 6. Infante YC, Pupo A, Rojas G. A combinatorial mutagenesis approach for functional epitope mapping on
468 phage-displayed target antigen: application to antibodies against epidermal growth factor. *mAbs* 2014; 6:637-48.
- 469 7. Puchades C, Kukrer B, Diefenbach O, Sneekes-Vriese E, Juraszek J, Koudstaal W, Apetri A. Epitope mapping of
470 diverse influenza Hemagglutinin drug candidates using HDX-MS. *Scientific reports* 2019; 9:4735.
- 471 8. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, Židek A, Bates R, Blackwell S, Yim J, et al.
472 Protein complex prediction with AlphaFold-Multimer. *bioRxiv* 2022:2021.10.04.463034.
- 473 9. Bourquard T, Musnier A, Puard V, Tahir S, Ayoub MA, Jullian Y, Boulo T, Gallay N, Watier H, Bruneau G, et al.
474 MAbTope: A Method for Improved Epitope Mapping. *J Immunol* 2018; 201:3096-105.
- 475 10. Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, Wang J, Cong Q, Kinch LN, Schaeffer
476 RD, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*
477 2021; 373:871-6.
- 478 11. Heyne M, Shirian J, Cohen I, Peleg Y, Radisky ES, Papo N, Shifman JM. Climbing Up and Down Binding
479 Landscapes through Deep Mutational Scanning of Three Homologous Protein-Protein Complexes. *J Am Chem Soc*
480 2021; 143:17261-75.
- 481 12. Dunham AS, Beltrao P. Exploring amino acid functions in a deep mutational landscape. *Mol Syst Biol* 2021;
482 17:e10305.
- 483 13. Doolan KM, Colby DW. Conformation-dependent epitopes recognized by prion protein antibodies probed
484 using mutational scanning and deep sequencing. *J Mol Biol* 2015; 427:328-40.
- 485 14. Van Blarcom T, Rossi A, Foletti D, Sundar P, Pitts S, Bee C, Witt JM, Melton Z, Hasa-Moreno A, Shaughnessy
486 LJomb. Precise and efficient antibody epitope determination through library design, yeast display and next-
487 generation sequencing. *J Mol Biol* 2015; 427:1513-34.
- 488 15. Medina-Cucurella AV, Zhu Y, Bowen SJ, Bergeron LM, Whitehead TA. Pro region engineering of nerve growth
489 factor by deep mutational scanning enables a yeast platform for conformational epitope mapping of anti-NGF
490 monoclonal antibodies. *Biotechnol Bioeng* 2018; 115:1925-37.
- 491 16. Sierocki R, Jneid B, Orsini Delgado ML, Plaisance M, Maillere B, Nozach H, Simon S. An antibody targeting
492 type III secretion system induces broad protection against Salmonella and Shigella infections. *PLoS Negl Trop Dis*
493 2021; 15:e0009231.
- 494 17. Eguia RT, Crawford KHD, Stevens-Ayers T, Kelnhofer-Millevolte L, Greninger AL, Englund JA, Boeckh MJ,
495 Bloom JD. A human coronavirus evolves antigenically to escape antibody immunity. *PLoS pathogens* 2021;
496 17:e1009453.
- 497 18. Sourisseau M, Lawrence DJP, Schwarz MC, Storrs CH, Veit EC, Bloom JD, Evans MJ. Deep Mutational Scanning
498 Comprehensively Maps How Zika Envelope Protein Mutations Affect Viral Growth and Antibody Escape. *J Virol* 2019;
499 93:e01291-19.
- 500 19. Dingens AS, Haddox HK, Overbaugh J, Bloom JD. Comprehensive Mapping of HIV-1 Escape from a Broadly
501 Neutralizing Antibody. *Cell Host Microbe* 2017; 21:777-87 e4.
- 502 20. Wu NC, Xie J, Zheng T, Nycholat CM, Grande G, Paulson JC, Lerner RA, Wilson IA, Ch, microbe. Diversity of
503 functionally permissive sequences in the receptor-binding site of influenza hemagglutinin. *Cell Host & Microbe* 2017;
504 21:742-53. e8.
- 505 21. Lee JM, Huddleston J, Doud MB, Hooper KA, Wu NC, Bedford T, Bloom JD. Deep mutational scanning of
506 hemagglutinin helps predict evolutionary fates of human H3N2 influenza variants. *Proc Natl Acad Sci U S A* 2018;
507 115:E8276-E85.
- 508 22. Starr TN, Czudnochowski N, Liu ZM, Zatta F, Park YJ, Addetia A, Pinto D, Beltramello M, Hernandez P,
509 Greaney AJ, et al. SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape. *Nature* 2021; 597:97-
510 +.

511 23. Terasawa K, Tomabeche Y, Ikeda M, Ehara H, Kukimoto-Niino M, Wakiyama M, Podyma-Inoue KA, Rajapakshe
512 AR, Watabe T, Shirouzu M, et al. Lysosome-associated membrane proteins-1 and -2 (LAMP-1 and LAMP-2) assemble
513 via distinct modes. *Biochem Biophys Res Commun* 2016; 479:489-95.

514 24. Cameron B, Dabdoubi T, Berthou-Soulie L, Gagnaire M, Arnould I, Severac A, Soubrier F, Morales J, Leighton
515 PA, Harriman W, et al. Complementary epitopes and favorable developability of monoclonal anti-LAMP1 antibodies
516 generated using two transgenic animal platforms. *PLoS One* 2020; 15:e0235815.

517 25. Agarwal AK, Gude RP, Kalraiya RD. Regulation of melanoma metastasis to lungs by cell surface Lysosome
518 Associated Membrane Protein-1 (LAMP1) via galectin-3. *Biochem Biophys Res Commun* 2014; 449:332-7.

519 26. Alessandrini F, Pezze L, Ciribilli Y. LAMPs: Shedding light on cancer biology. *Semin Oncol* 2017; 44:239-53.

520 27. Wilke S, Krausze J, Bussow K. Crystal structure of the conserved domain of the DC lysosomal associated
521 membrane protein: implications for the lysosomal glycoalyx. *BMC Biol* 2012; 10:62.

522 28. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Zidek A,
523 Potapenko A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature* 2021; 596:583-9.

524 29. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding
525 accessible to all. *Nat Meth* 2022; 19:679-82.

526 30. Deng X, Storz U, Doranz BJ. Enhancing antibody patent protection using epitope mapping information. *mAbs*
527 2018; 10:204-9.

528 31. Kowalsky CA, Faber MS, Nath A, Dann HE, Kelly VW, Liu L, Shanker P, Wagner EK, Maynard JA, Chan C, et al.
529 Rapid Fine Conformational Epitope Mapping Using Comprehensive Mutagenesis and Deep Sequencing. *J Biol Chem*
530 2015; 290:26457-U150.

531 32. Tran MH, Schoeder CT, Schey KL, Meiler J. Computational Structure Prediction for Antibody-Antigen
532 Complexes From Hydrogen-Deuterium Exchange Mass Spectrometry: Challenges and Outlook. *Frontiers in*
533 *immunology* 2022; 13:859964.

534 33. Wittrup KD. Protein engineering by cell-surface display. *Curr Opin Biotechnol* 2001; 12:395-9.

535 34. Hamilton SR, Bobrowicz P, Bobrowicz B, Davidson RC, Li H, Mitchell T, Nett JH, Rausch S, Stadheim TA,
536 Wischnewski H, et al. Production of complex human glycoproteins in yeast. *Science* 2003; 301:1244-6.

537 35. Fukuda M. Lysosomal membrane glycoproteins. Structure, biosynthesis, and intracellular trafficking. *J Biol*
538 *Chem* 1991; 266:21327-30.

539 36. Van Regenmortel MH. What is a B-cell epitope? *Methods Mol Biol* 2009; 524:3-20.

540 37. Orsini Delgado ML, Avril A, Prigent J, Dano J, Rouaix A, Worbs S, Dorner BG, Rougeaux C, Becher F, Fenaille F,
541 et al. Ricin Antibodies' Neutralizing Capacity against Different Ricin Isoforms and Cultivars. *Toxins* 2021; 13:100.

542 38. Starr TN, Greaney AJ, Addetia A, Hannon WW, Choudhary MC, Dingens AS, Li JZ, Bloom JD. Prospective
543 mapping of viral mutations that escape antibodies used to treat COVID-19. *Science* 2021; 371:850-4.

544 39. Focosi D, McConnell S, Casadevall A, Cappello E, Valdiserra G, Tuccori M. Monoclonal antibody therapies
545 against SARS-CoV-2. *Lancet Infect Dis* 2022.

546 40. Doud MB, Lee JM, Bloom JD. How single mutations affect viral escape from broad and narrow antibodies to
547 H1 influenza hemagglutinin. *Nature communications* 2018; 9:1386.

548 41. Keefe JR, Van Rompay KKA, Olsen PC, Wang Q, Gazumyan A, Azzopardi SA, Schaefer-Babajew D, Lee YE,
549 Stuart JB, Singapuri A, et al. A Combination of Two Human Monoclonal Antibodies Prevents Zika Virus Escape
550 Mutations in Non-human Primates. *Cell Rep* 2018; 25:1385-94 e7.

551 42. Dao T, Mun S, Korontsvit T, Khan AG, Pohl MA, White T, Klatt MG, Andrew D, Lorenz IC, Scheinberg DA. A
552 TCR mimic monoclonal antibody for the HPV-16 E7-epitope p11-19/HLA-A*02:01 complex. *PLoS One* 2022;
553 17:e0265534.

554 43. Arena S, Bellosillo B, Siravegna G, Martinez A, Canadas I, Lazzari L, Ferruz N, Russo M, Misale S, Gonzalez I, et
555 al. Emergence of Multiple EGFR Extracellular Mutations during Cetuximab Treatment in Colorectal Cancer. *Clinical*
556 *cancer research : an official journal of the American Association for Cancer Research* 2015; 21:2157-66.

557 44. Stevens AO, He Y. Benchmarking the Accuracy of AlphaFold 2 in Loop Structure Prediction. *Biomolecules*
558 2022; 12.

559 45. Ruffolo JA, Chu L-S, Mahajan SP, Gray JJ. Fast, accurate antibody structure prediction from deep learning on
560 massive set of natural antibodies. *bioRxiv* 2022:2022.04.20.488972.

561 46. Liang T, Jiang C, Yuan J, Othman Y, Xie X-Q, Feng Z. Differential performance of RoseTTAFold in antibody
562 modeling. *Brief Bioinform* 2022:bbac152.

563 47. Abanades B, Georges G, Bujotzek A, Deane CM. ABlooper: Fast accurate antibody CDR loop structure
564 prediction with accuracy estimation. *Bioinformatics* 2022.

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

587

588

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

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

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

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

600

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

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

607

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

609 **(A)** Co-crystal structure of the first luminal domain of hLAMP-1 in complex with Fab B. Ribbon diagram illustrating
610 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
611 yellow on the surface representation of the first luminal domain of hLAMP-1 (grey). **(B)** Graphical view of co-crystal
612 structure of the first luminal domain of hLAMP-1 in complex with Fab B colored with DMS Epitope Mapping data.
613 Amino acids with DMS scores above 10 are marked in red and amino acids with DMS scores between 5 and 9 are
614 shown in yellow (with a surface representation or with spheres on the ribbon representation). **(C)** Representation of
615 the AlphaFold model of the first luminal domain of hLAMP-1. Residues included in the DMS epitope are colored pink.

616

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

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

623

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

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

629

630

631

632

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

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

638

639

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

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

645

646

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

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

652

653

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

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

661

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

664

Data collection

Space group	C2
Cell dimensions	
<i>a, b, c</i> (Å)	149.93, 93.68, 108.01
α, β, γ (°)	90, 115.93, 90
Resolution (Å) *	72.08 – 2.37 (2.65 – 2.37)
<i>R</i> _{merge}	0.047 (0.413)
Mean (<i>I</i> / σ)	14.4 (2.8)
Completeness (%)	99.1 (99.3)
Redundancy	3.3 (3.3)

Refinement

Resolution (Å)	2.37
No. reflections	54389
<i>R</i> _{work} / <i>R</i> _{free}	0.251 / 0.290
No. atoms	
Protein	8703
Water	269
R.m.s. deviations	
Bond lengths (Å)	0.01
Bond angles (°)	1.04

665 *Highest resolution shell is shown in parentheses.

666

667

668

669

670

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

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

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

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

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

679

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

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

683

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

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

689

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

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

696

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

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

701

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

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

706

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

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

711

712 **Supplementary Figure 3: Predicted structure of hLAMP-1**

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

716

717 **Supplementary Figure 4: Localization of N-glycosylation sites on human LAMP1 in the vicinity of functional** 718 **epitopes of Fabs A and B**

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

722