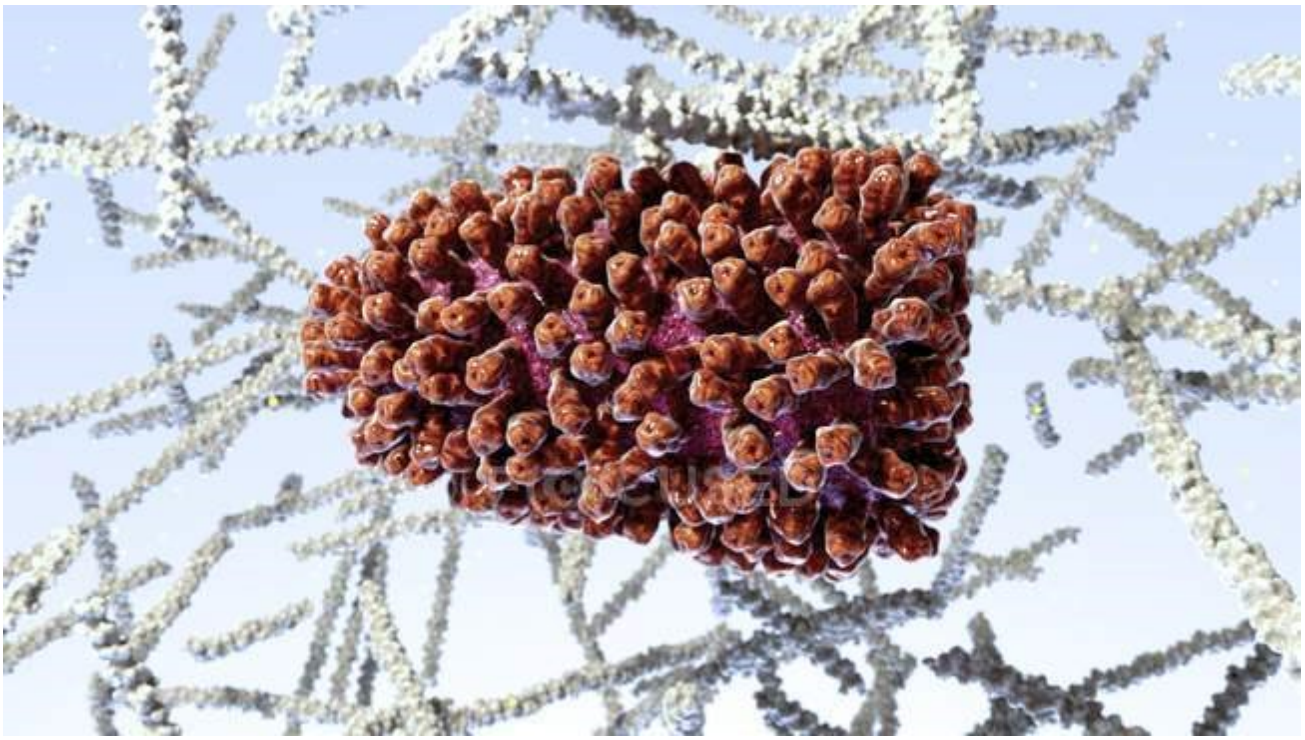
Utrecht University

A structural study on the Rabies Virus Glycoprotein combining cryo-EM and mass spectrometry-based approaches



Bob Coelingh (5995167)

Examined by dr. Joost Snijder of Biomolecular Mass Spectrometry and Proteomics group at Utrecht University

Daily supervision by dr. Matti Pronker of Biomolecular Mass Spectrometry and Proteomics group at Utrecht University

Utrecht, from 22-11-2020 until 29-10-2021

Abstract

Rabies Virus (RABV) is a zoonotic neurotropic virus and the main causative agent for Rabies, a lethal encephalitis. The Rabies Virus Glycoprotein (RABV G) is a homo-trimer present on the viral envelope and is essential in the viral life cycle because of its role in receptor recognition and membrane fusion and is the main immunogenic target. Here we report an updated protocol for the purification of RABV G from cell supernatant, which was used for cryo-EM single particle analysis to elucidate its structure. Additionally we mapped the N-linked glycans of the protein and characterized the oligomerization using mass photometry, native MS and Charge-Detection MS, which hinted at a concentration-dependent equilibrium. We also assessed the binding of two therapeutically used neutralizing antibodies by HDX-MS and can point to a probable epitope on RABV G for both.

Layman's summary

The Rabies Virus (RABV) is a virus transmitted to humans via animals causing 60.000 deaths yearly. A very important protein in the life cycle of the Rabies Virus is the Glycoprotein (RABV G) which is essential for a successful infection. In this research we purified RABV G with the goal of obtaining its structure by using cryogenic Electron Microscopy. Additionally we characterized our purified protein by investigating binding interactions of RABV G with its receptors from host cells as well as with antibodies that neutralize the virus. We showed binding of RABV G to these antibodies of high affinity and most likely located the location where it binds. We also characterized the occupation of sugar groups on the protein. Furthermore we examined the interactions of RABV G itself that form larger assemblies of multiple subunits.

Introduction

Rabies virus (RABV) is a neurotropic virus and the main causative agent of rabies, a fatal encephalitis with 59.000 yearly reported deaths[i]. Approximately 95% of cases occur in Africa or Asia, mostly in low-income countries, and they almost exclusively originate from infected dogs[ii]. Currently, the use of preventative as well as post-exposure vaccines prevents millions of deaths, as the post-exposure vaccines alone are already being used in 15 million cases each year. Several pre-exposure vaccines can elicit long-lasting immunity, but they are expensive and require multiple doses to achieve neutralizing titers[iii]. A single-dose vaccine would be necessary to establish a cost-effective approach to exterminate rabies in humans in lower-income countries[iv]. However, complete eradication of rabies is impossible due to the presence of a large reservoir of lyssaviruses in bats[v], which means an effective approach to terminate rabies in humans remains elusive.

RABV, a member of the genus *Lyssavirus* and the family of *Rhabdoviridae*, is a bullet-shaped enveloped virus. It has a 12 kb non-segmented negative stranded RNA genome encoding five essential viral proteins: nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G) and RNA polymerase (L). N proteins encapsulate the viral RNA and together with P and L proteins form a helical ribonucleoprotein (RNP) that is closely associated with the M protein, which also associates with the lipid bilayer and thereby contributes to the tight packing of the virion (*Fig. 1*).[vi]. The glycoprotein is the only protein exposed on the outside of the virus and therefore the main immunogenic target of RABV[vii]. Currently, strains of RABV overexpressing G protein are already being widely used to vaccinate dogs because they invoke a strong immune response which leads to neutralizing titers of antibodies [viii].

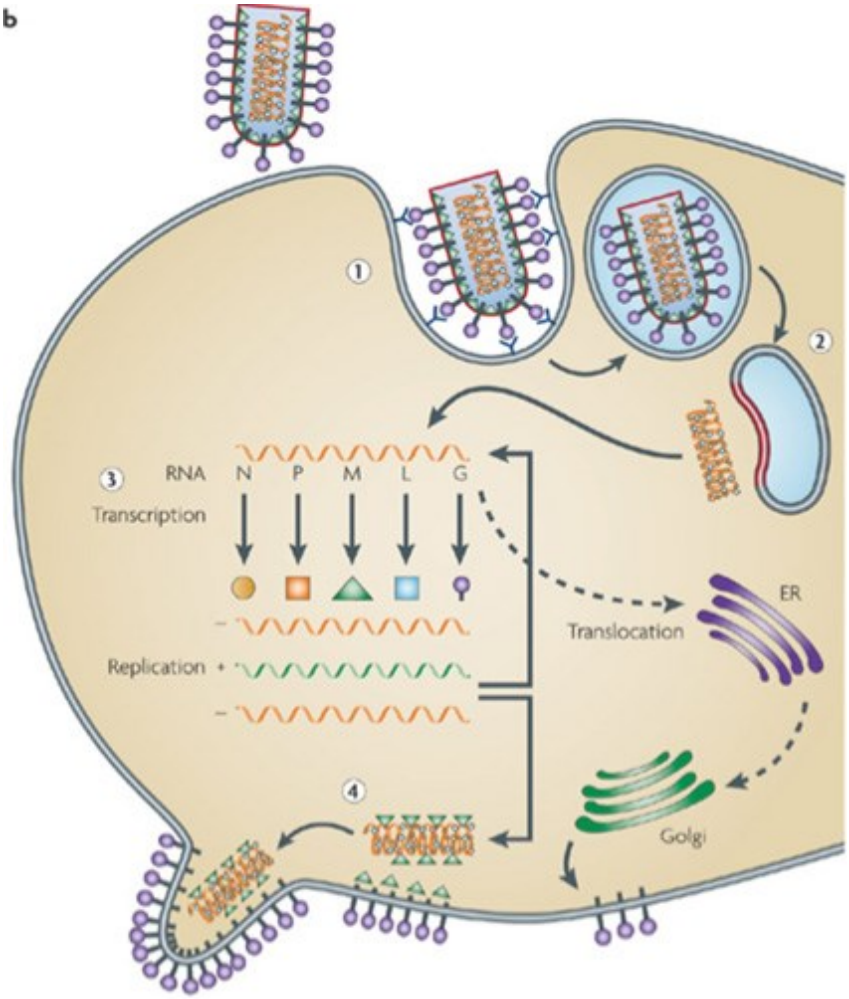


Figure 1 | Rabies Virus cellular life cycle[ix]

G has an ectodomain consisting of approximately 500 amino acids, a transmembrane helix and a small endodomain which has been shown to interact with Matrix proteins in Vesicular Stomatitis Virus (VSV)[x] as well as with host cell proteins, in turn dysregulating neuronal cell homeostasis and contributing to success of viral infection[xi]. Previously, it has been determined that G has three predicted glycosylation sites, of which only one or two have been found to be glycosylated. The glycans have been suggested to solubilize hydrophobic folding intermediates to prevent aggregation during folding[xii], which could explain the large variety in animals susceptible to infection without the differences in glycosylation altering infectivity.

G is essential for viral pathogenicity and has two distinct tasks in the viral life cycle: it enables cell-specific attachment by viral recognition of host cell receptors, as well as membrane fusion of the virion with endocytic vesicles upon internalization. This releases the nucleocapsid into the host cell which is the start of infection. RABV has been suggested to use several internalization methods including the clathrin-mediated endocytosis as well as actin-dependent internalization[vii]. Upon maturation of these endocytic vesicles the pH decreases which triggers a conformational change in the glycoproteins that facilitates membrane fusion. Several host cell receptors have been proposed as receptors for RABV, such as the p75 neutotrophin receptor ($p75^{NTR}$)[xiii], the Neural Cell Adhesion Molecule (NCAM1)[xiv], the metabotropic Glutamate Receptor subtype 2 (mGluR2)[xv] and the nicotinic acetylcholine receptor (nAChR)[xvi]. However, the fact that nAChR is located on the post-synaptic side of the neuro-muscular junction likely indicates the nAChR is not involved in entry into the neurons.

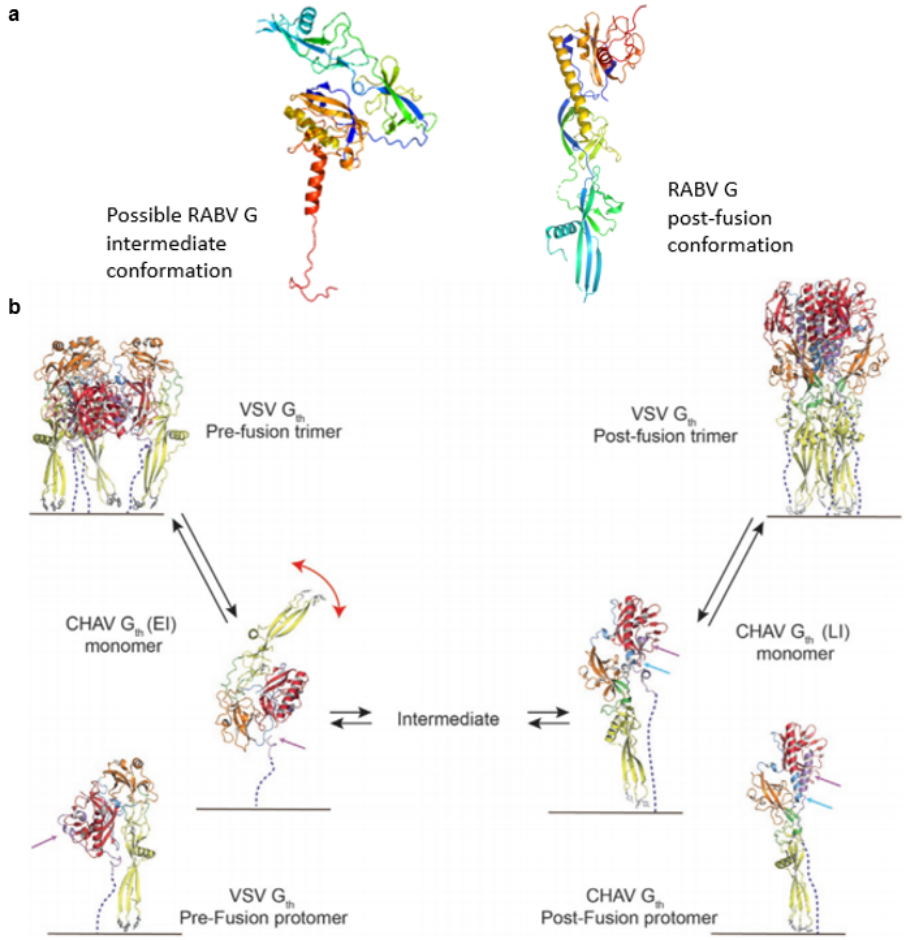


Figure 2 | Rhabdoviridae structures (a) Current RABV G structures (b) Rhabdoviridae glycoprotein structures and possible rearrangement mechanism[xvii]

Many viral fusion proteins are trimeric on their viral envelope, which is also true for RABV G, as it has been found to form homo-trimers on the viral envelope in pre- and post-fusion conformation[xviii]. In fact, formation of this trimer has been shown to be necessary for interactions with host cell receptors[xiii]. For many viruses, the fusion proteins undergo an irreversible conformational change to achieve fusion with the host cell membrane, often also due to proteolytic cleavage by co-factors. However, *Rhabdoviridae* glycoproteins such as those of RABV, VSV and CHAV (Chandipura Virus) are class III viral fusion proteins. All three of these viruses have the very interesting and rare property that the (pH-induced) conformational change the Glycoprotein undergoes to achieve membrane fusion is reversible[xixxx]. Recently, crystal structures of RABV G have been solved in a monomeric state, which were suggested to be monomeric intermediates in the conformational change from pre- to post-fusion trimers[xxi]. Aside from this, much remains unknown on the RABV Glycoprotein, such as the trimeric structures and the domain rearrangements occurring during membrane fusion. This information would greatly help in our understanding of the virus as well as with designing next-generation vaccines and anti-viral therapeutics.

Currently, structures of several Glycoproteins from the *Rhabdoviridae* family have been solved as well as two crystal structures for RABV G[xxi]. These structures are both monomeric, one is suggested to be an intermediate closer to the post-fusion conformation and the other one is suggested to be an earlier intermediate conformation which looks similar to a Chandipura virus monomeric intermediate state (*Fig. 2*). The structure of the pre-fusion conformation has not been reported for Rabies currently. This would be the most interesting structure from a vaccine-based view. Nor have there been trimeric structures elucidated. Aside from solving these structures, the binding of host cell receptors could be evaluated a lot more seeing that for most of the receptors the binding has only been demonstrated once. Demonstrating and characterizing these interactions could give invaluable information for the viral attachment of Rabies. The working mechanism of the conformational change and the membrane fusion is yet to be elucidated. Possible manners to do this would be by investigating the binding of neutralizing antibodies to RABV G using HDX-MS to study their binding epitopes and gain insight into the mechanism of neutralization which would teach us a lot on the mechanism of membrane fusion.

Results

RABV G post-fusion conformation purification and cryo-EM

For many of the applications of structural biology, a pure protein sample of high concentration is needed. The RABV G yields were originally quite low, so a lot of time was spent optimizing the purification protocol. The construct (*Fig. 3a*) was transiently expressed in a HEK293E+ cell line, which was harvested after one week. RABV G was purified from cell supernatant using a tandem affinity approach consisting of IMAC performed on an Äkta Go fast protein liquid chromatography (FPLC) system followed by Strep-tag purification performed in a gravity flow column using streptactin beads. To obtain a sharper elution peak during the first IMAC step, the flow was stopped after equilibration of the column in 500 mM imidazole. Elution was continued after one hour resulting in a sharp peak consisting of RABV-G (*Sup. Fig. 1*). This protocol gives a yield of approximately half a microgram of glycoprotein from 2 L of HEK293E+ supernatant in its post-fusion conformation which means that it is eluted from the Streptactin beads with an elution buffer at a pH of 5,5. The sample is also of good purity (*Fig. 3b*) and compared to doing the IMAC and Strep-tag purification in batch, both yield and purity are greatly improved.

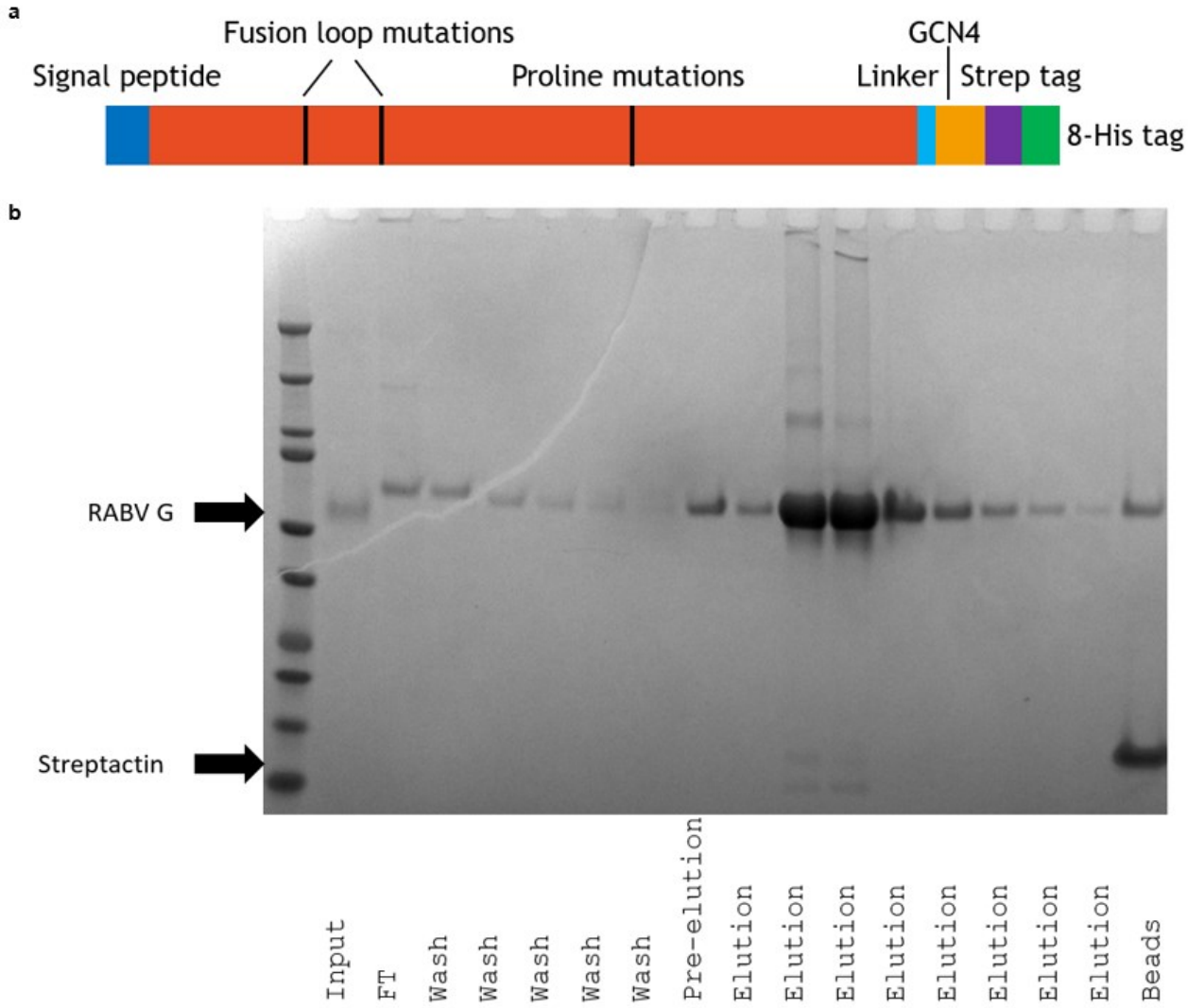


Figure 3 | RABV G construct and purification. (a) RABV G construct with modifications (b) SDS-PAGE of Strep-tag step of RABV G purification

After the purification of RABV G in the post-fusion conformation a cryo-EM dataset was recorded. Upon close inspection of the micrographs, there clearly is a good distribution of particles visible on the grids (*Fig. 4a*). Particles of the expected size are abundant on the grids, although many larger particles are also visible on all micrographs which are likely aggregated RABV G (*Fig. 4a*). Upon image processing, 2D classification in 25 classes show a large amount of particles present in classes that look very similar to VSV G post-fusion (F*ig. 4c and d*). The structure for this virus from the same family has already been solved (*Fig. 4d*)[xxii]. The 2D classes clearly show great similarity in terms of shape, namely being elongated with a sharper side as well as a more dull side. There also seems to be clear density for the glycans which are extended outwards from the trimer. However, there is a large discrepancy in size. The mask size used for the particle picking and 2D classification was 200 Å which is quite large considering previously published results showed the full protein containing transmembrane region to have an estimated size around 132 Å[xviii]. It could be possible that the protein is somehow larger under these conditions or in some sort of enlarged intermediate state or that the mask size is simply chosen too large which biases the image processing towards larger particles consisting of higher oligomers. The data processing was continued, after which 3D classification (*Fig. 4c*) showed a best class much more resembling the structure of the VSV Glycoprotein. The 3D class has an estimated resolution of 7 Å (*Sup. Fig. 2)*, close to side-chain resolution, which would be needed to build a structure but it clearly looks very similar to the VSV G post-fusion. The overall size and shape of the trimer seem to be very similar and it's clearly possible to distinguish the fusion domains from the 3D class.

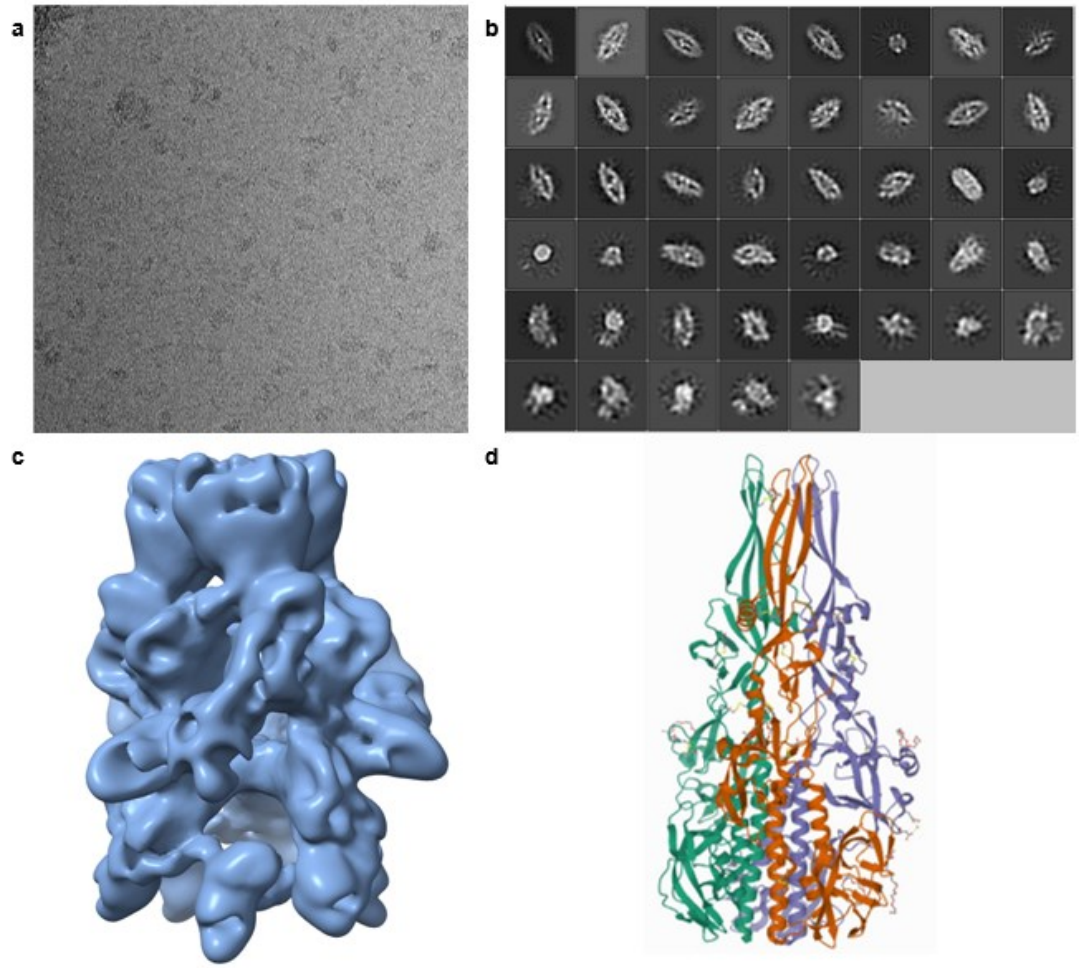


Figure 4| cryo-EM dataset shows RABV G post-fusion is similar in structure to VSV G post-fusion. (a) representative micrograph of cryo-EM dataset collected from post-fusion conformation RABV G (b) 2D class averages of 200 Å mask (c) 3D class (d) structure of VSV G post-fusion

Glycosylation of RABV G

The Pasteur Vaccine strain used for our research has four potential N-linked glycans (N61, N182, N271, N343) (*Fig. 5a*) in the glycoprotein ectodomain. Other strains of Rabies frequently miss an Asparagine at position 182 or 271. The glycan sites all extend from the outside of the trimer interface and are not thought to contribute to host cell recognition or membrane fusion. They have been found to solubilize hydrophobic intermediate states during folding[xii]. Another possible role for these glycans could be in the shielding of antigenic epitopes from the immune system which has been described for many other enveloped virus glycoproteins and famously for HIV[xxiii]. As previously reported the only necessary glycosylation of RABV G for the excretion of soluble Glycoprotein is N343. Additionally a N61 glycan was not found[xx]. We also found no evidence of glycosylation on the N61 and only two glycan isoforms detected at position 182 (N4H5F1S1, N6H3F2) possibly due to the presence of a cysteine in the glycopeptide which forms an intramolecular disulfide bridge. Positions 271 and 343 showed a variety of glycan isoforms (*Fig. 5b and c*) with a relatively high fucosylated glycan being the most prevalent for both positions. The variety for the 343 position was larger than for the 271 position. This is possibly related to the fact that position 343 is necessary for the correct folding of the protein while the glycan at 271 can be removed by mutagenesis without adverse effect to the excretion[xx]. Interestingly there is a large amount of biantennary glycans which is quite unusual but has previously been reported for the specific cell lines and medium used for transfection and expression. The average weight of the glycan isoforms per site was calculated according to their abundance and was found to be 2,0 kDa for each site which brings the total weight of RABV G, which without glycosylation weighs 55 kDa, to 61 kDa.

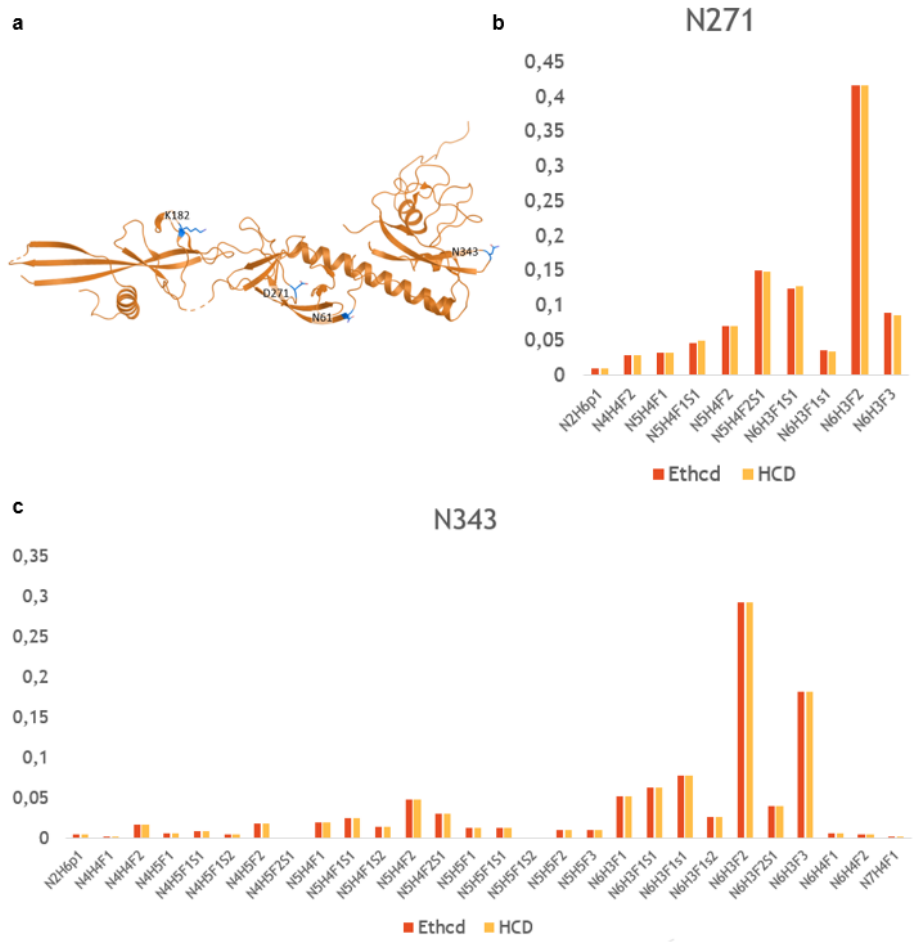


Figure 5 | N-linked Gycosylation sites of RABV G and glycan isoforms. (a) Predicted N-linked glycosylation sites for the Pasteur Vaccine strain shown on the structure of the post-fusion (b and c) glycan isoforms of N271 and N343 site respectively

Oligomerization of RABV G by Mass Photometry, Native MS and Charge Detection MS

To evaluate the oligomerization of the protein after observing significant aggregation on the micrographs of the dataset, we had turned to a quick and easy screening tool we had in our possession which makes use of mass photometry. We measured purified RABV G after every purification and hoped to get a better idea of different conditions on the oligomerization. These measurements clearly show a very large amount of the protein being in a monomeric state with a very small fraction in trimers for both conformations (*Fig. 6 a and b*). For the pre-fusion RABV G there is also a lot of noise in higher mass ranges likely due to unstructured aggregates (not shown). For the post-fusion material there is virtually no noise and only a small peak for what could be a tetramer of trimers when looking at the mass besides peaks for monomer and trimer. It's hard to tell from these measurements what the 763 kDa peak belongs to because of the mass accuracy of the technique but it seems very much like an interesting feature because of the signal compared to the noise and the bell-shape. All measurements for mass photometry have to be done at nanomolar concentration ranges to allow for single event binding to be picked up by the detector. It seems to be that these conditions are not suitable for the formation of trimers. This could indicate a dynamic equilibrium for oligomerization which lies more in the micromolar concentration ranges. Cryo-EM grids were prepared from much more concentrated protein samples so these results do not teach us anything useful in that aspect.

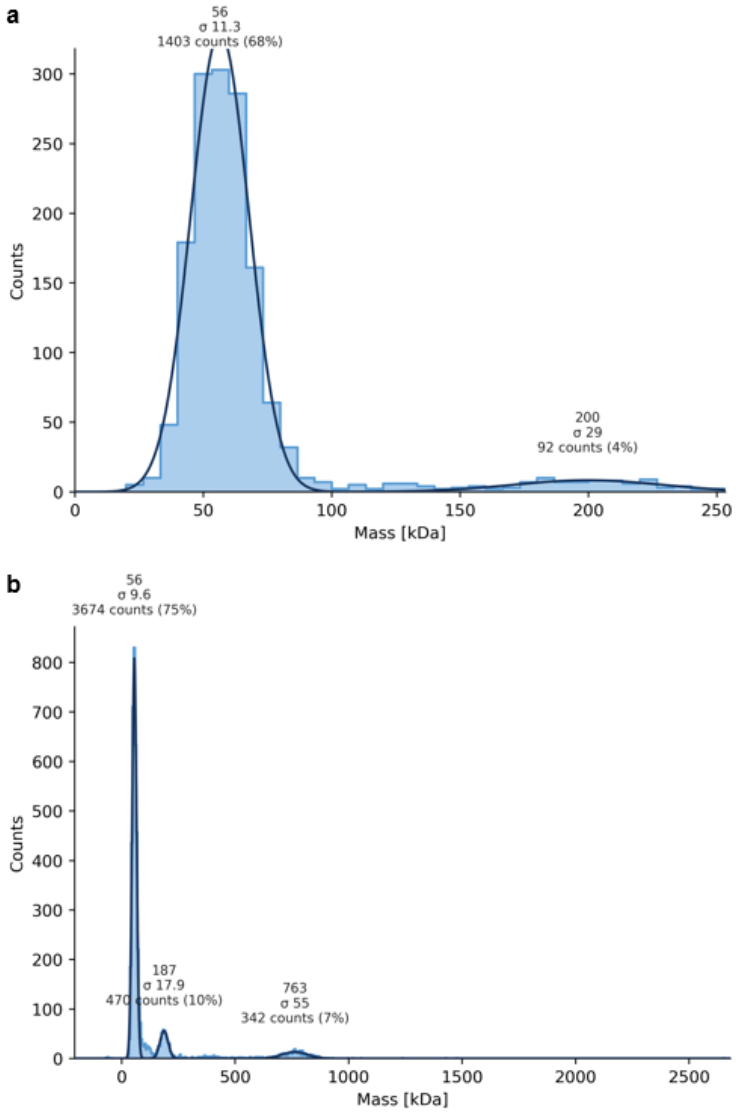


Figure 6 | Mass photometry measurements of RABV G in (a) pre-fusion conformation and (b) post-fusion conformation

We then turned to Native MS and Charge Detection MS (CD-MS) to evaluate the oligomerization of the RABV G post-fusion protein. These techniques, although less native in the buffer environment of the sample do allow for measurements to be done at much higher concentrations and mass accuracies. Purified RABV G samples in post-fusion were buffer exchanged to ammonium acetate buffers of either pH 6 or 8 from the pH 5,5 buffer the protein was in. According to previous studies[xxii], this should force RABV G from the post-fusion conformation back to the pre-fusion conformation, a feature that is almost unique to the *Rhabdoviridae* family glycoproteins. The post-fusion material showed several ordered oligomers starting with monomer, dimer and trimer which would all be expected (*Fig. 7a*). Interestingly, there are also higher oligomers belonging to oligomers of trimers in a very ordered fashion. The masses were determined for the distinct signals and they corresponded to dimers, trimers and tetramers of RABV G trimers. These signals were high compared to the single monomer and trimer signals, taking into account that they were not corrected for the amount of monomers per oligomer. That would mean when corrected for this the signal of the higher oligomers would be the highest by far. The higher oligomers measured here would indicate that the observed problem in the cryo-EM could not be attributed to aggregation caused by unfolded protein, although this could still be a problem which happens on the air-water interface, but more likely the trimers of RABV G have an interface at which they bind to other Glycoproteins trimers. This would be in agreement with previous observations of Rabies virions by EM, in which the trimers appear packed very closely to each other and they have been shown to arrange in regular arrays before . The sample at pH 8 (*Fig. 4b*) has a very similar signal in the lower mass ranges corresponding to monomer, dimer and trimer but in the higher mass ranges the signal looks much less regularly ordered and much more like unstructured aggregates in higher mass ranges than for the post-fusion material. This clearly shows that the oligomerization properties change due to the change in pH, which is most likely due to the conformational change from post-fusion to pre-fusion conformation. The signal does not correspond to a regular oligomer, which could indicate that the conformational change is not complete and some intermediate states aggregate together, or that the pre-fusion conformation somehow does not form these regular arrays but does expose different residues on to the environment which can bind to each other. The tetramer of trimers peak was also picked up by the mass photometry measurements and the presence of noise in the pre-fusion material also corresponds very much here. This is a good comparison of the reversibility of the pre-fusion from native MS with pre-fusion material that was measured directly with mass photometry, which indicates a good efficiency of the reversibility.

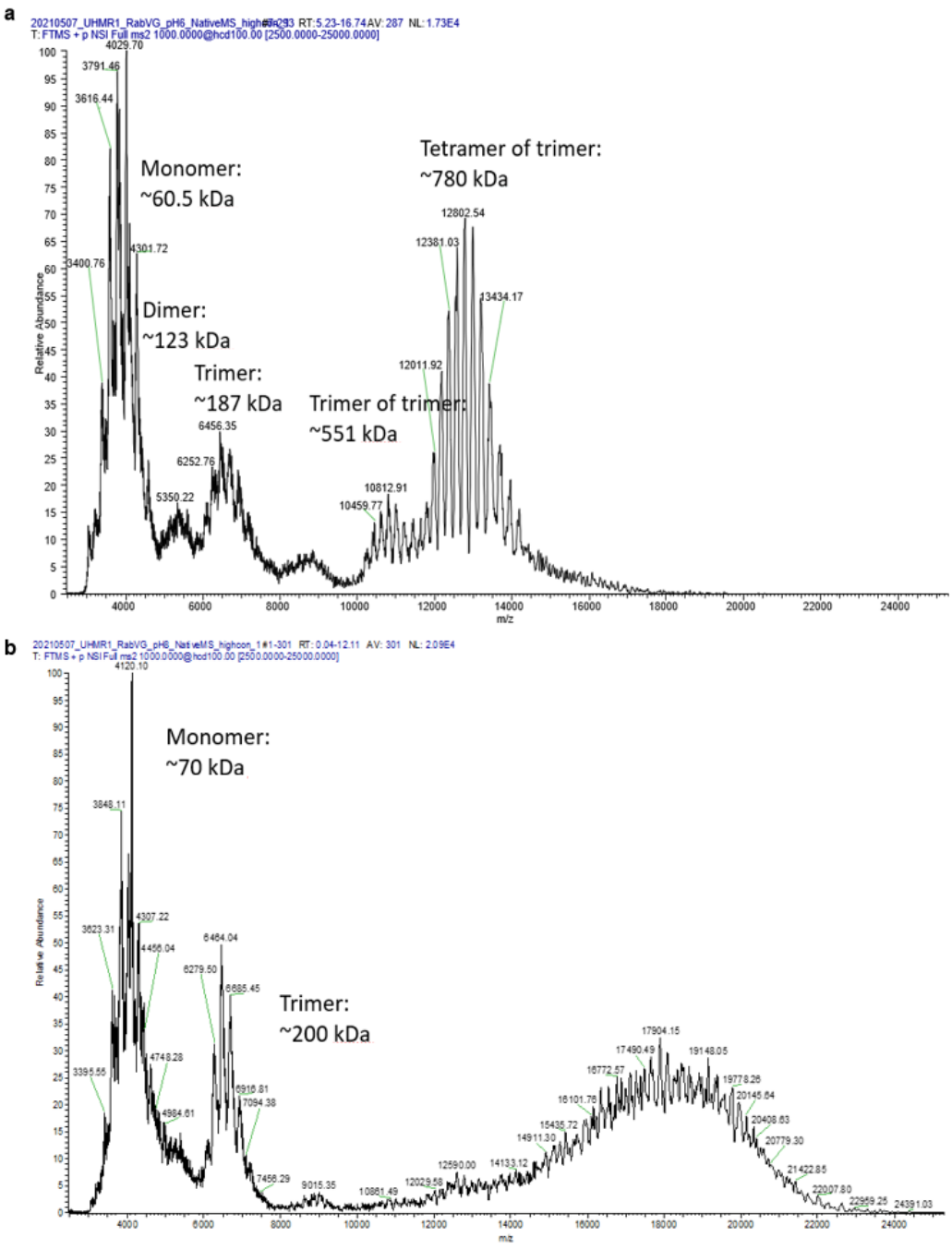


Figure 7 | Native MS measurements of post-fusion RABV G. (a) post-fusion RABV G in pH 6 ammonium acetate buffer (b) post-fusion RABV G in pH 8 ammonium acetate buffer

The CD-MS measurements were performed using the same buffer exchanged samples to pH 6 (*Fig. 8a and c*) and pH 8 (*Fig. 8b and d*). The measurements had to be done in two separate recordings because the total mass range difference between the monomer and the higher oligomeric species was larger than fourfold in m/z which is not desirable for CD-MS experiments. For both samples, the lower mass range species of monomer, dimer and trimer have a very similar signal but the difference in the higher mass ranges is striking again. However, interestingly the CD-MS seems to contradict the native MS because the higher mass ranges of the post-fusion material show less distinctly structured species. It only has distinct species corresponding to trimer of trimer and tetramer of trimer whereas the material at pH 8 has oligomers consisting of two to six trimers. The differences between both mass spectrometry experiments is very interesting and indicates there might be more to the

oligomerization of the material than what we would have thought using only one of these methods. It seems contradictory that the different samples seem to have different oligomerization in both measurements. This could be an artefact introduced by the different conditions for which both measurements were optimized. In conclusion, RABV G aggregation, which seemed to be problematic from the cryo-EM seems to be much more structured than randomly. The larger particles are more likely regular oligomers consisting of multiple trimers. For all techniques the difference between both samples was very obvious, thereby demonstrating the unique property of reversible pH-dependent conformational changes of *Rhabdoviridae* Glycoproteins.

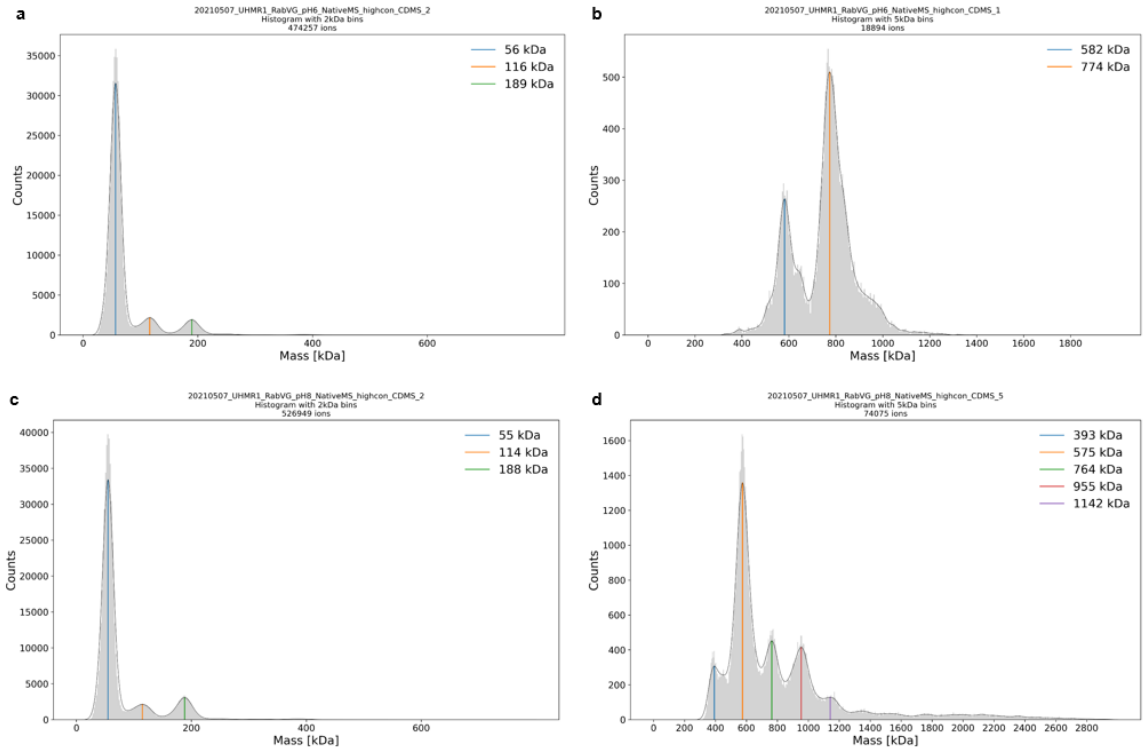


Figure 8 | CD-MS measurements of post-fusion RABV G (a and b) post-fusion RABV G in pH 6 ammonium acetate buffer in lower and higher mass ranges respectively (c and d) post-fusion RABV G in pH 8 ammonium acetate buffer in lower and higher mass ranges respectively

Neutral pH negative stain EM of pre-fusion RABV G

While data processing of the post-fusion cryo-EM dataset was in progress we turned our attention to the biologically more important pre-fusion conformation with regards to the purification. The yield of the purification of this conformation was considerably lower after Strep-tag purification resulting in a concentration too low for use in cryo-EM (*Sup. Fig. 3*). The low yield seemed to arise in the Strep-tag step of the purification since the IMAC step was very reproducible in terms of UV chromatogram. The reduced yield could be due to a decreased stability of the pre-fusion conformation compared to the post-fusion, leading to a higher aggregation rate, possibly reducing the efficiency of the binding of the Strep-tag. To still attempt to evaluate this material we shifted our attention to negative stain EM to assess the quality of the material for electron microscopy purposes. For negative stain screening normally a uranyl-based stain would be used since these have a very fine particle size, resulting in a good resolution, as well as the fact that they can be made in larger batches and stored for long periods of time. However, these uranyl-based stains have a low pH (~4), which is not convenient for a pH-sensitive protein such as RABV G, especially for this sample which was still at a pH of 7,5 and in its pre-fusion conformation. An alternative was to use a phosphotungstic acid stain which could be set to a pH matching the buffer the protein was present in. In this manner, the stain could be a valuable

screening tool for samples such as these where pH could influence the sample drastically. It could even provide a more meaningful context to biological samples that in their native cellular environment don't encounter a pH outside of the physiological range. The sample had a concentration which was well suited for negative stain EM, with a good particle density and distribution (*Fig. 9*). The particles seem to be in the expected size range with shapes resembling the short and stumpy shape of VSV G pre-fusion trimers[xvii]. Overall, the material when purified at a high enough concentration seemed very promising for cryo-EM applications, and the phosphotungstic acid stain a valuable tool for screening of samples at a neutral pH.

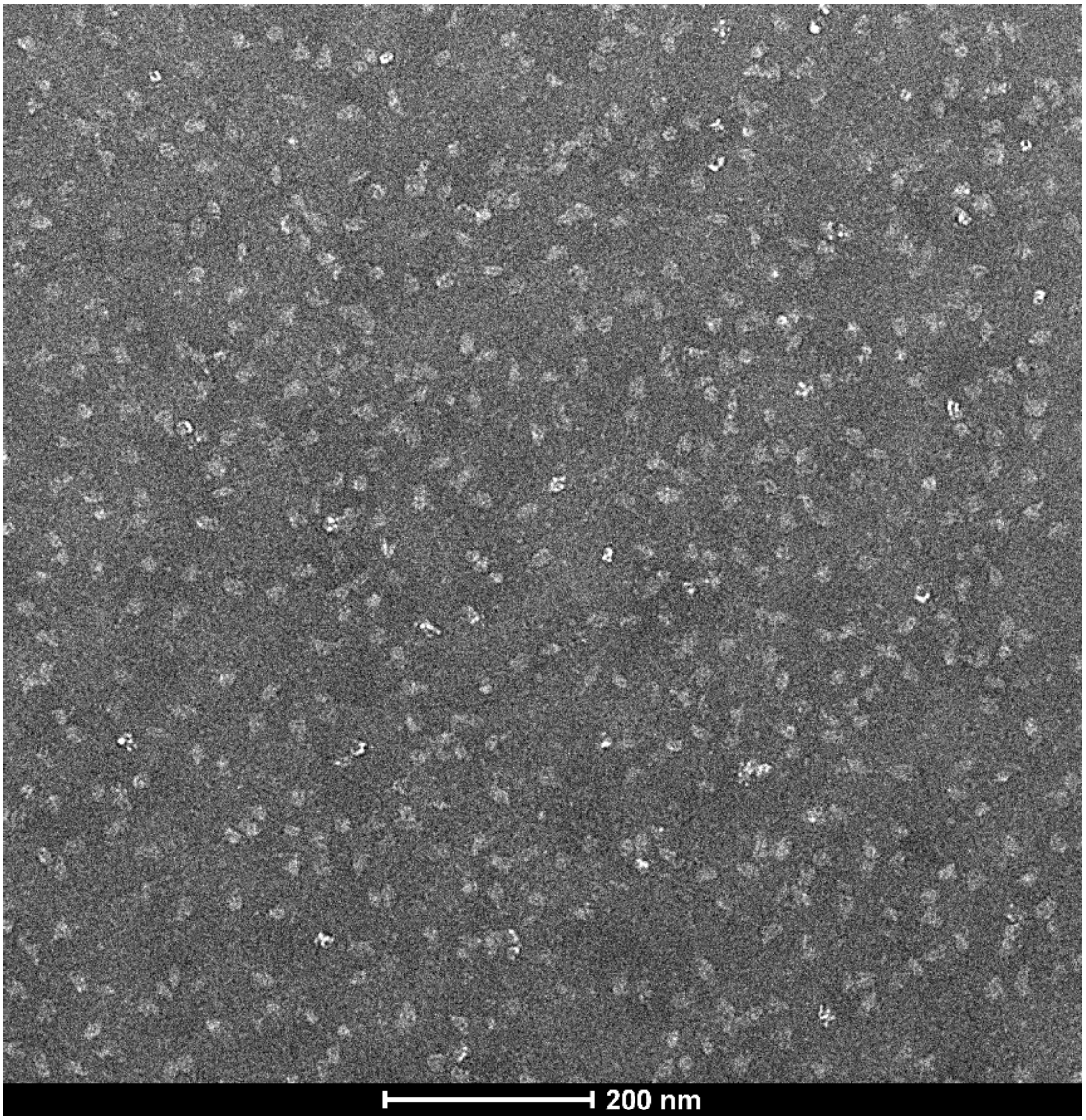


Figure 9 | Negative stain EM micrograph of pre-fusion RABV G stained with phosphotungstic acid.

Co-Immunoprecipitation of RABV G with binding partners

To evaluate the binding of RABV G to host cell receptors we first sought out to demonstrate binding of our construct to some purified proteins which had previously been shown to be a receptor for the Rabies Virus. Previously, the receptor $p75^{NTR}$ and two Fabs made from neutralizing antibodies called 6G11 and 17C7 had been purified. We also purified several different C-terminal truncations of NCAM1 missing either one or multiple domains using IMAC followed by SEC (*Sup. Fig. 4*).

During our standard purification protocol, Streptactin beads bound to RABV G purified from HEK293 cell lines containing a N-acetyl glucosamine (GlcNAc) transferase I (GnTI) knockout (ES) and cell lines without GnTI knockout (E+) in pre-fusion conformation were used to pull down the purified host cell receptors. After washing away unbound protein, the samples were eluted from the beads and run on SDS-PAGE before western blotting and probing with mouse anti-His (Qiagen "penta-his") and secondary anti-mouse fused to horse radish peroxidase (HRP). All proteins used have a C-terminal His_8-tag. The GnTI-/- cell line had an undetectably low yield on western blot of RABV G. The Fabs clearly bound to RABV G, because bands were visible in both the shorter and longer exposure of the blot (*Fig. 10a and b*). The putative host receptor $p75^{NTR}$ appears to bind the Streptactin beads with higher affinity than RABV G, so no clear binding was demonstrated for this protein. The truncated NCAM1 construct consisting of IG1-5 and Fn1 (but lacking the second Fn domain) seemed to bind, because a very small band was observed slightly above the large band for RABV G in the long exposure of the blot, whereas nothing was seen on the short exposure (*Fig. 10c and d*). We demonstrated binding of RABV G to Fabs, that are used as neutralizing antibodies in the clinic, with an affinity that seems to be quite high. The possible binding of NCAM1 IG1-5 Fn1 was unexpected, seeing that the same construct with an extra Fn domain did not show any binding. This longer truncation seemed to be prone to aggregation during its purification and could possibly have aggregated during the Co-Immunoprecipation. This could have resulted in the protein being pulled down with RABV G being undetectably low on the blot. The affinity for NCAM1 to RABV G seems to be quite low which would not be strange in a physiological context where a virion richly decorated with Glycoproteins travels between neurons in synapses densely populated with NCAM1 molecules.

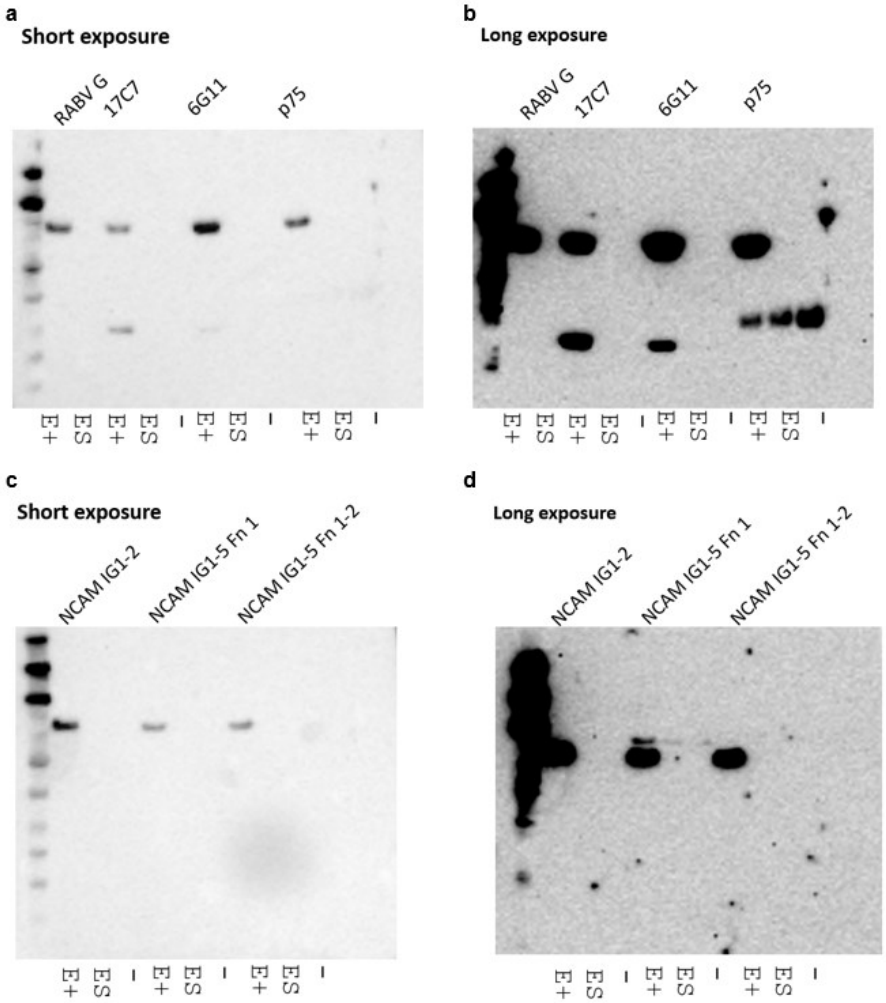


Figure 10 | Co-Immunoprecipitation of RABV G pre-fusion with binding partners.

Purification of pre-fusion stabilized RABV G and HDX-MS with Fabs

Because of disappointing yields for the purification of RABV G in the pre-fusion conformation, we tried another approach to stabilize the pre-fusion conformation. Using scanning mutagenesis, we made mutations to proline residues in a part of the alpha helix which is expected to form the central coiled coil in the post-fusion state, but not in the pre-fusion state of the trimer. Twelve individual point mutations to proline were made and tested for expression, after which combinations between the best-expressing individual mutations were made and tested for expression in small scale. We selected a 2P mutant (E288P, L290P) with relatively high expression in small-scale and purified it in large-scale (*Sup. Fig. 5*). Yield and purity were comparable to a previous post-fusion conformation purification. In parallel to trying to obtain a cryo-EM dataset of this sample, we set out to investigate the binding of the Fabs to RABV G further. We had demonstrated binding of high affinity in the co-IP but the elucidation of the epitopes by using HDX-MS would give valuable information on the possible mechanism of neutralization, as well as possibly identifying important residues for the conformational changes needed to achieve membrane fusion.

Before the experiments pilots were done using only RABV G 2P peptides were identified from initial measurements. The coverage obtained for RABV G during the experiment reached almost 75% (*Fig. 11*) with a good peptide coverage for the more C-terminal parts of the protein. There were some areas of the protein where there was no coverage, but these can in large part be explained by the fact they are part of the signal peptide, contain a Cysteine forming a disulfide bridge or contain a N-linked glycan.

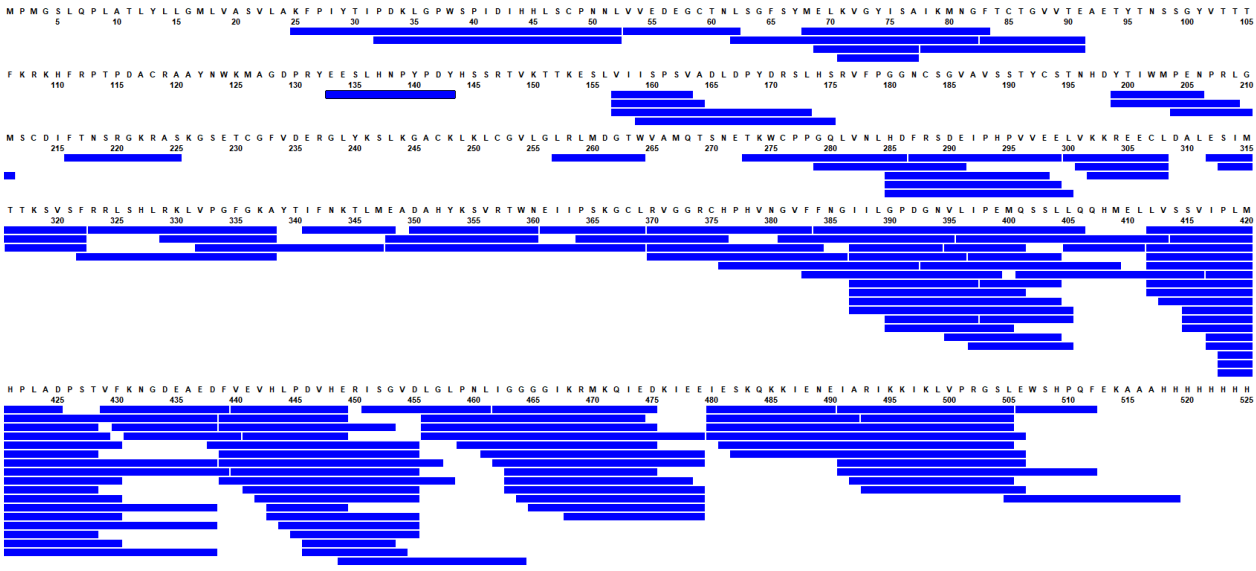


Figure 11 | Coverage map of RABV G 2P-mutant

The experiment was set up as follows: incubation of RABV G 2P-mutant was done in deuterium with excess of either Fab, followed by low pH quench, pepsin column digestion, LC-separation and tracking of peptide m/z values on a TOF-MS that were previously identified in the pilot experiment. Several peptides showed clearly significant mass shifts compared to the control (*Sup. Fig. 6*). These differences in relative uptake of deuterons correlate with the addition of the Fabs. Therefore, they most likely stem from the binding of the Fab protecting the peptides from the deuterium solution and thus altering the rate of exchange. After statistical analysis by T-test, the peptides containing a significant (p value < 0.01) difference in relative uptake compared to the control were selected and the relative uptakes plotted on a homology model made from RABV G pre-fusion conformation threaded on the trimeric VSV G pre-fusion structure (5OYL)[xvii]. The peptides that seem to change the most drastically upon addition of either Fab appear to cluster on a specific region (*Fig. 12a and c*). This area is located on the

outside of the trimer and points outwards from the trimer interface on the beta-sheet-rich lateral domain of the protein.

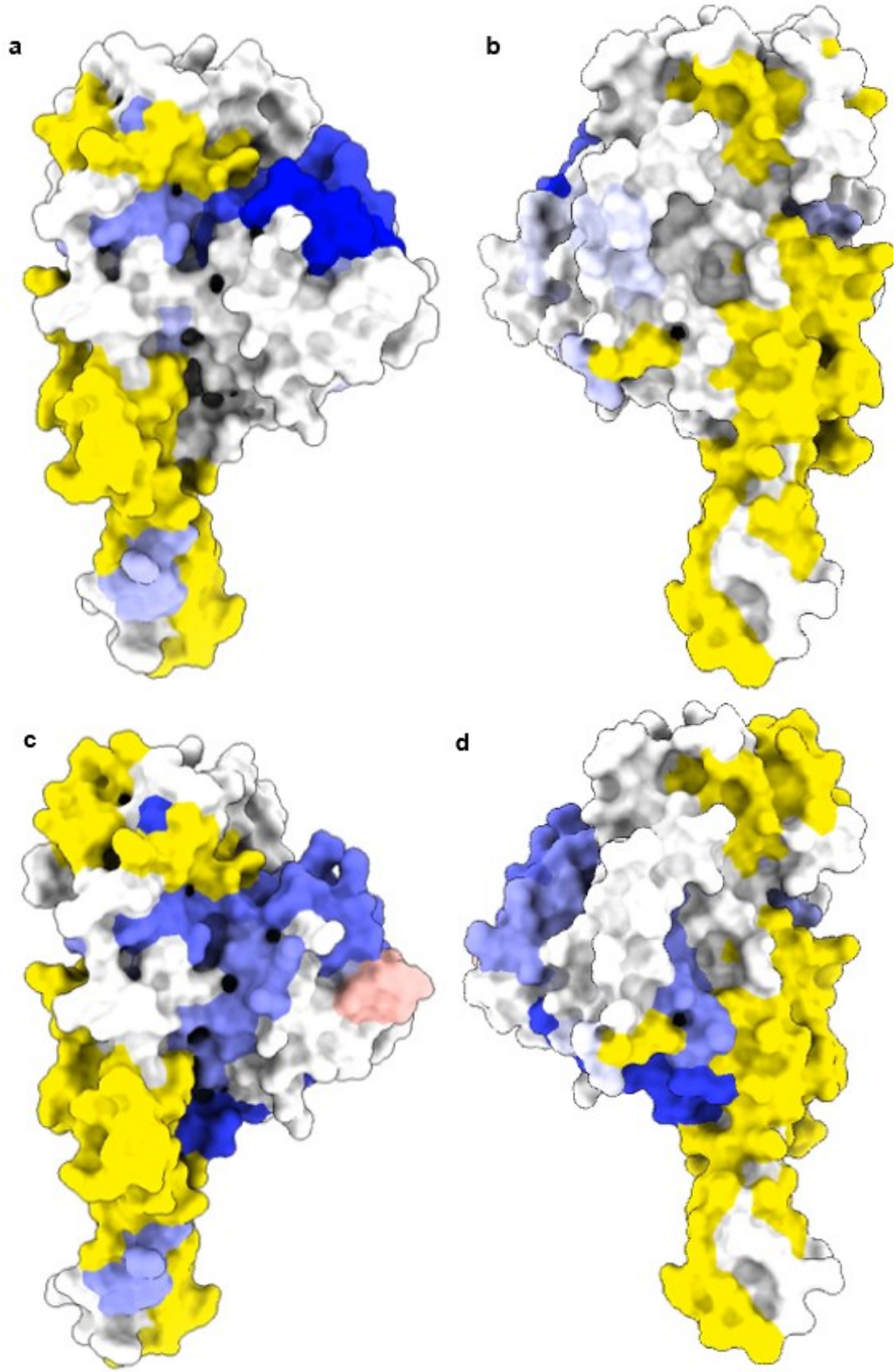


Figure 12 | Relative uptake of deuterium by RABV G peptides in addition of Fabs. (a and b) Addition of Fab 17C7 (c and d) addition of Fab 6G11. (a and c) view of non-trimer interface extending outwards of trimer (b and d) view of trimer interface. Scale: blue to red, protection to deprotection; yellow, no coverage.

For both Fabs there is also an effect of protection visible on the trimer interface (*Fig. 12b and d*) but this effect seems less pronounced. It appears these two neutralizing antibody-derived Fabs share the same binding epitope. The difference at the trimerization interface could be explained by Fab binding at the outer epitope allosterically stabilizing the trimer, thereby causing more protection for peptides present at the trimer interface. The epitope could sterically be bound by antibodies because the virus envelope is located oppositely from the epitope on the virion (*Fig. 13*). An antibody could conceivably bind two monomers in different trimers seeing as they are packed so tightly together. In this manner the antibody could prevent adherence of the virus to host cells. Other possible manners of neutralization could involve immune response activation or preventing the conformational change from pre-fusion to post-fusion which is necessary to achieve membrane fusion and thus for the viral life cycle because the virus has to escape the endocytic vesicle to start replication before it is degraded in a lysosome.

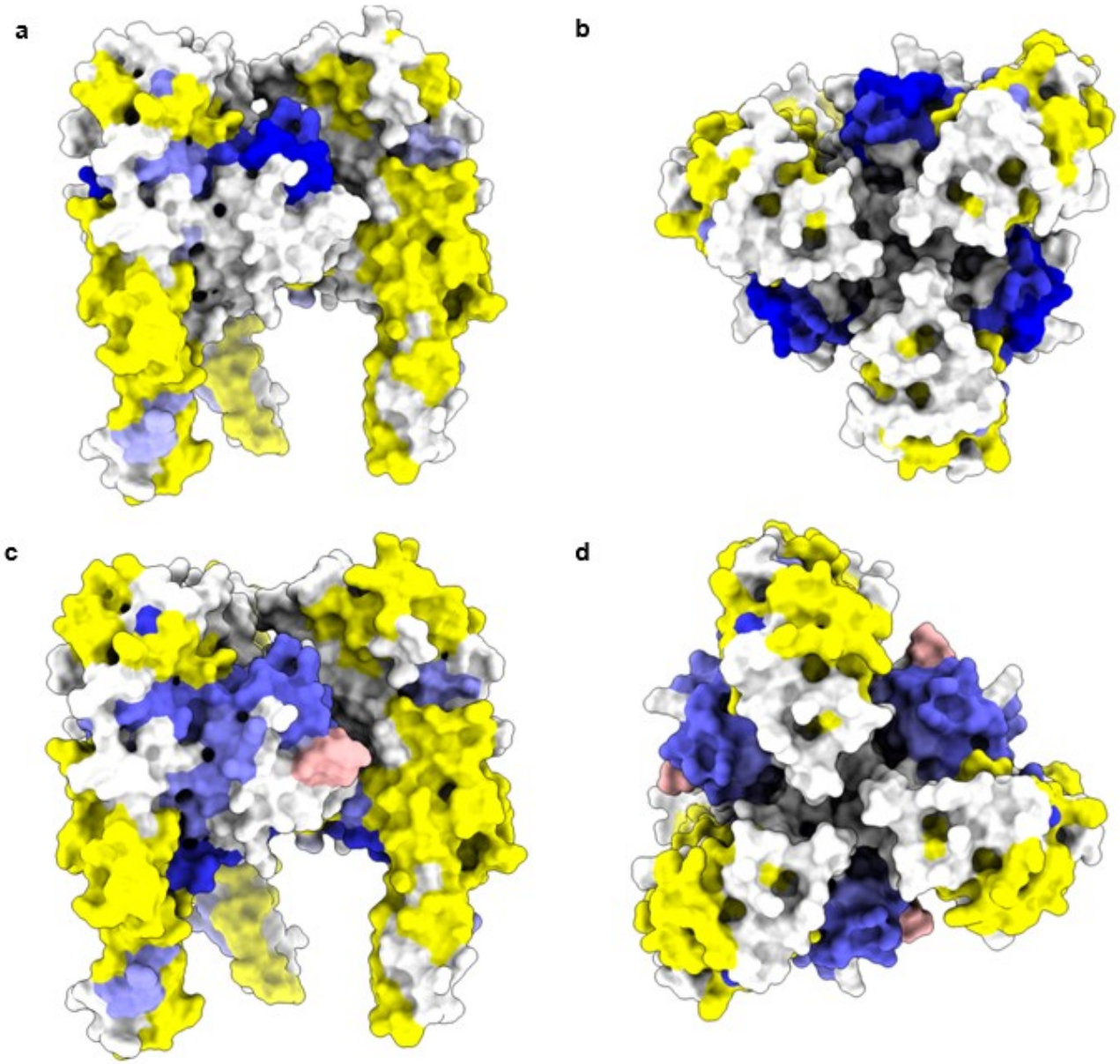


Figure 13 Trimeric view of relative uptake of deuterium by RABV G peptides in addition of Fabs. (a and b) side- and top-view of trimer in addition of 17C7 respectively (c and d) side- and top-view of trimer in addition of 6G11 respectively

Conclusion

The updated purification protocol for RABV G gives a much improved yield and the material is of good purity and quality to be used for experiments to study its structure and binding to host cell receptors. The N-linked glycans of our Pasteur Vaccine strain construct were mapped and glycans were found on three of the four predicted sites. The oligomerization of RABV G is still a serious hinderance to achieving a high-resolution structure using cryo-EM because a large part of the particles is much too large to be a trimer. This limits the amount of useful particles that can be extracted from micrographs. The Mass Photometry, native MS and CD-MS shine a light on this oligomerization, as the data show structured higher order oligomers. Both MS-based methods show contradictory evidence for which conformation would be most prone to form these higher oligomers. What it taught us however is that the formation of these higher order oligomers seems to be concentration-dependent. In the nanomolar range in solution the main species to be seen by far was the monomer, whereas in the micromolar range the main species seems to be shifted more to trimer and higher oligomers. It seems as if there is a very small range of concentrations in which there is a high proportion of trimers. When the trimer concentration becomes saturated, the formation of higher oligomers commences. This could mean that a creative approach has to be tried to solve the RABV G structure with a high resolution. It might be worth investigating these trimer-trimer interfaces further to possibly mutate some residues on this interface or try to use additives which would disrupt this interaction. None of the additives tried have worked so far. The use of vesicles decorated with RABV G could be one such manner. They could be purified from cells transfected with the full length RABV G including its transmembrane region. The addition of a membrane would orient the trimers all in the same direction as well as stabilizing the fusion loops on the membrane and allowing for very regular trimer-trimer contacts. A single particle analysis wouldn't work anymore so a tomography approach or a hybrid approach would have to be used instead.

The binding of the Fabs is of high affinity and we have probably found their binding epitopes. Follow-up experiments are necessary to structurally assess the binding mode, such as using cryo-EM or negative stain EM of G:Fab complexes to distinguish between the two protected regions on the Glycoprotein. Additionally, SPR experiments could be used to test whether the epitopes of both Fabs overlap in a competition experiment and both binding affinities could be quantified as well. If both Fabs bind the same epitope this could be a clue that that site is a good target for neutralization and thus most likely important for the working mechanism of the Glycoprotein. SPR could then also be used for the study of the interactions between RABV G and its host receptors. Quantifying these affinities would be interesting to estimate the amount of binding receptors needed for endocytosis and infection. Many questions remain on these interactions. The binding affinities, possibly paired with epitopes, would give us a much greater insight in the host cell attachment of Rabies and could possibly help in developing novel vaccines or anti-viral therapeutics.

Methods

Constructs

Pasteur vaccine strain RABV-G ectodomain (res. 1-446) preceded by signal peptide and with a C-terminal LIGGGGIKR linker followed by a homotrimeric coiled coil trimerization tag (MKQIEDKIEEIESKQKKIENEIARIKK), a thrombin cleavage site, a Strep-II tag and a His8 tag with mutations in the fusion loops (F93S, V94S, W140S, L141S) and for the 2P mutant additionally (E288P, L290P) was subcloned in a pUPE mammalian expression vector using BamHI/NotI sites.

RABV G expression

Mammalian expression vectors containing the constructs were used for transient transfection using polyethylenimine of in suspension culture growing GntI-/- Epstein-Barr virus nuclear antigen (EBNA) I-expressing (HEK293ES) or only EBNA I-expressing (HEK293E+) cells (U-Protein Express) in Freestyle™ medium. After one week, cell supernatant was collected by centrifugation at 1000 g for 15 min and filtering through 0.22 µm Steritop™ Filter Units (Millipore).

RABV G purification

Protein was purified by performing IMAC. Filtered HEK293 supernatant was loaded onto a prepacked Histrap™ Excel 5 mL column with peristaltic pump, then attached to an Äkta Go system (GE Healthcare) and washed with IMAC A (500 mM NaCl, 20 mM HEPES, 2 mM $CaCl_2$ and pH 7,5) with 50 mM imidazole for 40 column volumes. Elution was performed using IMAC A containing 500 mM imidazole. The flow was paused for one hour after equilibration with elution buffer. Following pause, elution was finished. Elution was diluted to a concentration of 100 mM imidazole with IMAC A buffer, then loaded onto a gravity flow column packed with washed Streptactin Sepharose® beads. The column was washed with IMAC A for 10 column volumes, and then equilibrated with SEC buffer (150 mM NaCl, 20 m MES, 2 mM $CaCl_2$ and pH 5,8) for 3 column volumes. Protein was eluted with SEC buffer containing 10 mM desthiobiotin, elution was paused for 15 min before completion. Yield and purity were evaluated using SDS-PAGE followed by Imperial Stain.

Cryo-EM and image processing

Grids were prepared by applying 3 µL of sample to glow-discharged holey carbon grids (C-Flat thick) three consecutive times with blotting of the excess using filter paper. Grids were plunged in liquid ethane cooled with liquid nitrogen using a Vitrobot (FEI). Micrographs were collected using a Titan Krios (Thermo Fisher) at an acceleration voltage of 300 kV equipped with a Falcon IV direct electron detector in super-resolution mode (EER format) and a Selectris energy filter with a slidwith of 10 eV. Defocus was set -2, -1.7, -1.4, -1.1 and -0.8 µm, with a beamsize of 0.7 µm, a total exposure of 7.7 seconds, a total electron dose of ~50 electrons/$Å^2$, a dose rate of 6.6 e/pixel/second and a pixel size corresponding to 0.96 Å. Motion correction was performed using MotionCor2, ctf correction by Gctf and particles were picked using Cryolo. Further processing was performed using Relion3.1.

Glycoproteomics analysis

Sample was prepared from SDS-PAGE by in-gel digestion from Strep-tag RABV G purification. Bands were excised and washed in 25 mM NH_4CO_3 and 100% acetonitrile consecutively, followed by a 1 hour reduction step with 10 mM DTT in 25 mM NH_4CO_3 at 56°C. Washing with 100% acetonitrile preceded a 20 minute alkylation step with 55 mM iodoacetic acid (IAA) in 25 mM NH_4CO_3 at room temperature in complete darkness. Samples were washed again with acetonitrile, then digestion was performed by

using 100 ng of trypsin (R/K) and Glu-C (D/E) sequentially. The supernatant was collected by elution with 100% acetonitrile after which fractions were pooled and vacuum dried. Peptides were reconstituted in 15 µl 2% formic acid solution for mass spectrometry analysis.
Six µL of the resuspended peptides were run on an Orbitrap Fusion Tribrid mass spectrometer coupled to nanospray UHPLC system Agilent 1290. Peptides were separated in a 90 minute LC gradient with increasing percentage of acetonitrile from zero to 35% at a flow rate of 300 nl/min. An analytical Poroshell 120 EC-C18 column and a ReproSil-Pur C18 trap column were used to separate the peptides. The measurements were done in data-dependent mode using Orbitrap Fusion parameters of: AGC target 4 x 10^5 at 60.000 resolution, a maximum injection time of 50 ms and over a 350 to 2000 m/z mass range. The most intense ions were fragmented using EThcD and HCD. MS2 spectra were acquired with AGC target 5 x 10^5 at resolution 30.000, maximum injection time of 250 ms and over a 120 to 4000 m/z mass range with exclusion of peptides 16 seconds after selection. Data were analyzed using Byonic searching against a 485 glycan database excluding glycopeptides with a Byonic score below 200 and LogProb below 2 and a length below 6 peptides.

Mass photometry of RABV G

Glass microscope cover slips were cleaned in a sonic bath with isopropanol for 5 minutes and then dried with nitrogen gas. Purified RABV G was measured at a concentration of 33 nM on Refeyn for 60 seconds. Histograms were made and gaussians fitted according to resolved peaks.

Native MS

35 µL of ~20 µM RABV G in SEC buffer of pH 5,8 was buffer exchanged into 170 mM ammonium acetate of either pH 6 or 8 with a 10 kDa Molecular Weight Cut-Off filter. The final volume after six rounds of buffer exchanging was approximately similar to the starting volume. 3 µL sample was loaded into in-house gold-coated borosilicate capillaries and infused into a Q exactive UHMR Hybrid Quadrupole-Orbitrap (Thermo Fisher) mass spectrometer. The gas in the collision cell was nitrogen. Settings optimized for this sample were as applied to the mass spectrometry measurements were as follows: 3125 resolution at 200 m/z with 16 ms transient, HCD energy of 100 V, injection of 100 ms, in-source trapping of -75 V, read-out pressure of $1{,}5*10^{-10}$ mBar, capillary temperature of 250°C and spray voltage of 1,4 kV.

Orbitrap-based Charge Detection MS

For the CD-MS the same buffer-exchanged samples were used as for the native MS and measured on a Q Exactive QE-UHMR mass spectrometer using the same gold-coated capillaries. Settings that were optimized for this sample are as follows: capillary voltage of 1,5 kV, capillary temperature of 250°C, injection time of 10 ms, resolution of 280.000 at 200 m/z with a transient time of 1 sec, in-source trapping voltage of -75 V, HCD energy of 100 V and read-out pressure of $6*10^{-11}$ mBar with nitrogen gas. Calibration of the mass spectrometer was performed according to a previously established protocol (*Wörner et al.*, Nature Methods 2020)

Negative Stain EM

Commercial grids with continuous carbon were glow-discharged and 3 µL of sample was applied for 1 minute, excess was removed with filter paper. Grids were washed thrice with Milli-Q water, staining was performed using either Uranyl Acetate or a 2% Phosphotungstic Acid solution at pH 7,5 for a 1 minute incubation. PTA stain was made by dissolving PTA in milli-Q water by constant shaking in hand for approximately 30 minutes followed by setting of the pH. Micrographs were taken on a 200 kV TEM with LaB6 filament Tecnai T20 (FEI) at 45,000x magnification.

NCAM1 purification

Mammalian expression vectors containing NCAM1 (IG1-2, IG1-5, IG1-5 Fn1, IG1-5 Fn1-2) C-terminal truncations were expressed and purified as described above for the RABV G, with the second step of the purification consisting of Size-Exclusion Chromatography on an Äkta Go system using a Superdex® 200 high performance column pre-equilibrated with SEC buffer (pH 5,8)

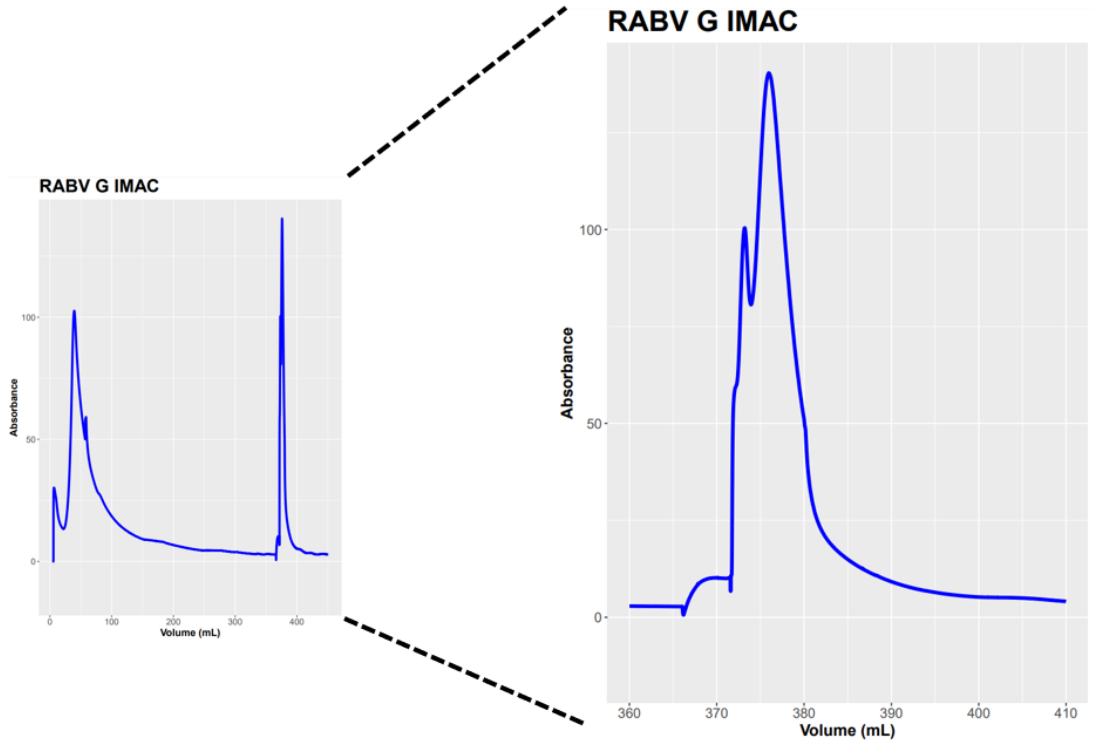
Co-Immunoprecipitation

Material used was cell supernatant of HEK293E+ and HEK293ES cells transfected as described previously with RABV G. After previously described IMAC purification, in-batch incubation of diluted IMAC elution with washed Streptactin Sepharose® beads was performed. Beads were isolated through centrifugation and incubated in 200 µL volume of SEC buffer (pH 7,5) buffer (150 mM NaCl, 20 m MES, 2 mM $CaCl_2$ and pH 7,5) containing 1 µM binding partners (Fab 17C7, Fab 6G11, NCAM1 IG1-2, NCAM1 IG1-5 Fn1, NCAM1 IG1-5 Fn1-2, p75) overnight at 4°C with overhead rotation. Beads were washed three times with SEC. Elution was performed in two steps using SEC buffer with 10 mM desthiobiotin and evaluated on SDS-PAGE followed by Western Blot with mouse Anti-His (Qiagen penta-His) and secondary anti-mouse fused to horse radish peroxidase.

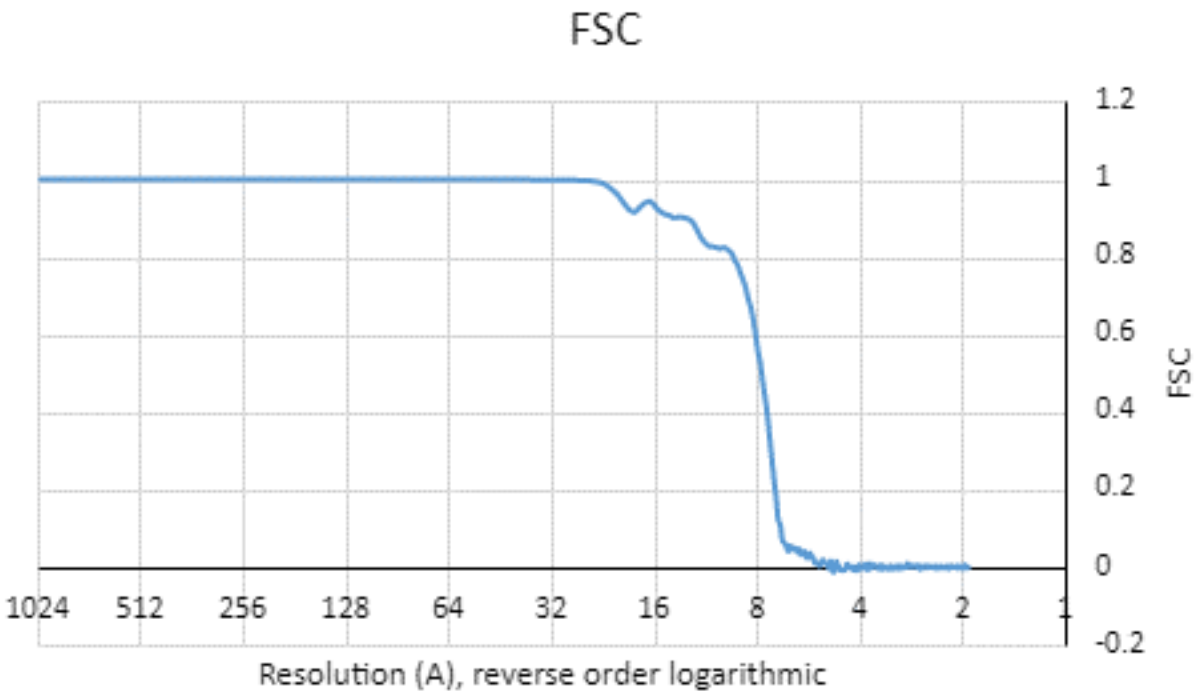
HDX-MS of RABV G 2P mutant in combination with Fabs

The initial identifications were done in triplo with 10 µL of purified RABV G 2P-mutant (100 pmol) diluted in 30 µL 100 mM Tris, 150 mM NaCl at pH 7,5. Also added was 20 µL 6 M Urea quench solution with 300 mM TCEP at pH 2,5. Immediately after quench sample was injected into 50 µL sample loop of nanoACQUITY UPLC System combined with HDX technology. Sample was loaded onto pepsin column for digestion at 20 °C followed up by trapping in-line at 0,5 °C in Waters Van Guard™ BEH C18 trap column for 3 min. Separation was performed using Waters Acquity UPLC BEH C18 analytical column in 15 minute gradient increasing acidified acetonitrile (0.1% formic acid) percentage from eight to 95% while decreasing acidified Milli-Q water. Peptides were measured with XEVO G2 mass spectrometer over mass range of 50 to 2000 m/z. Peptide identifications were performed with Waters PLGS 3.0.1 software including methionine oxidation and N-linked glycans as possible modifications. Peptide spectra were further processed using DynamX 3.0 selecting peptides with an intensity over 1000, a minimum length of 4 amino acids and presence in at least two of the three replicate measurements. Inspection of peptides resulted in removal of extremely long peptides as well as peptides with overlapping spectra interfering with the assignments. These peptides were tracked during the actual measurements of 70 pmol RABV G 2-P mutant and 200 pmol of Fab 17C7 or 80 pmol RABV G 2P mutant and 160 pmol of Fab 6G11 which were incubated at room temperature for 30 minutes with constant shaking for Fabs to bind prior to experiment. The samples were incubated at four different time points: 10 seconds, 1 minute, 1 hour and 4 hours with deuterium at a concentration of 75% before quenching with the urea quenching buffer. The same filtering and inspection of peptides occurred followed by statistical analysis in Excel using a student T-test and excluding peptides with a relative uptake of deuterium with a $p > 0,01$.

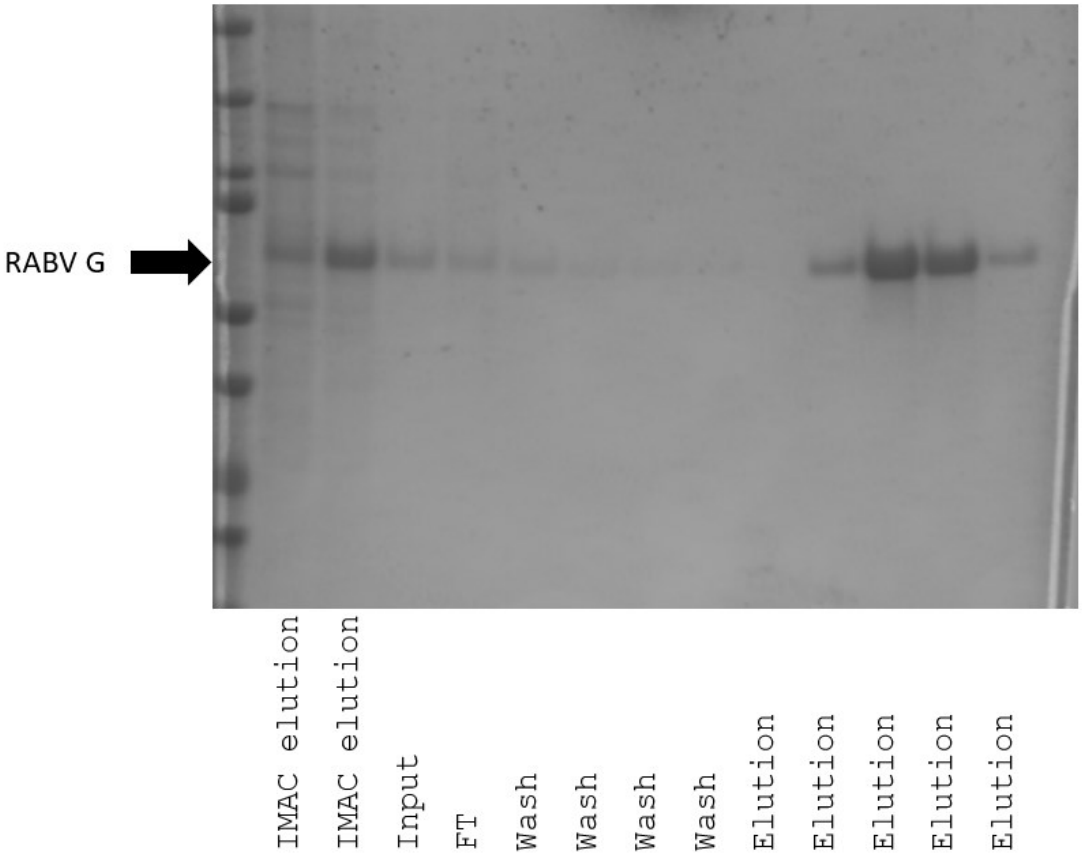
Supplementary figures



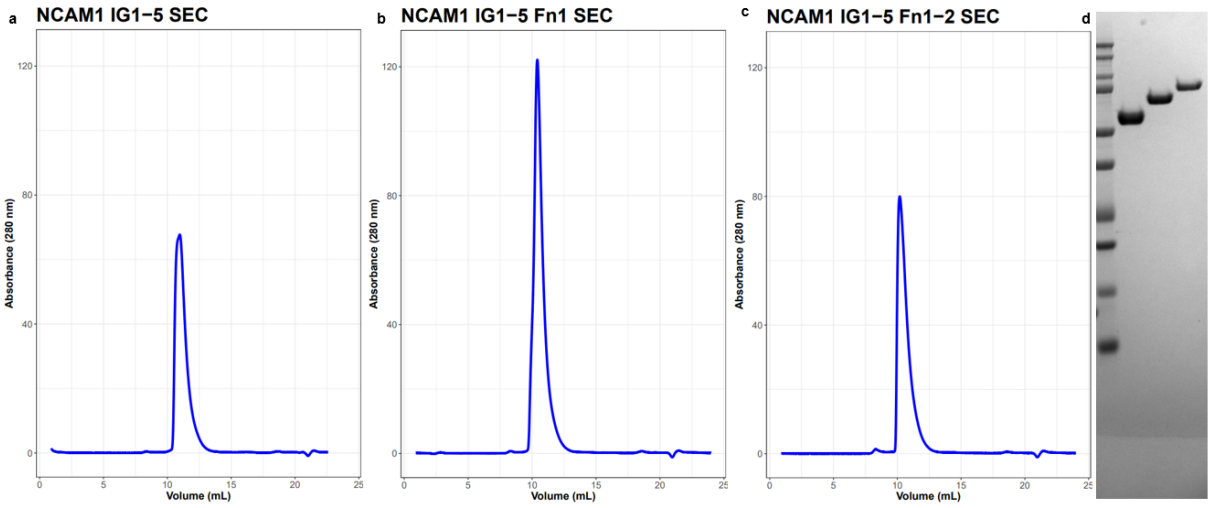
Supplementary figure 1 | IMAC step of RABV G post-fusion purification



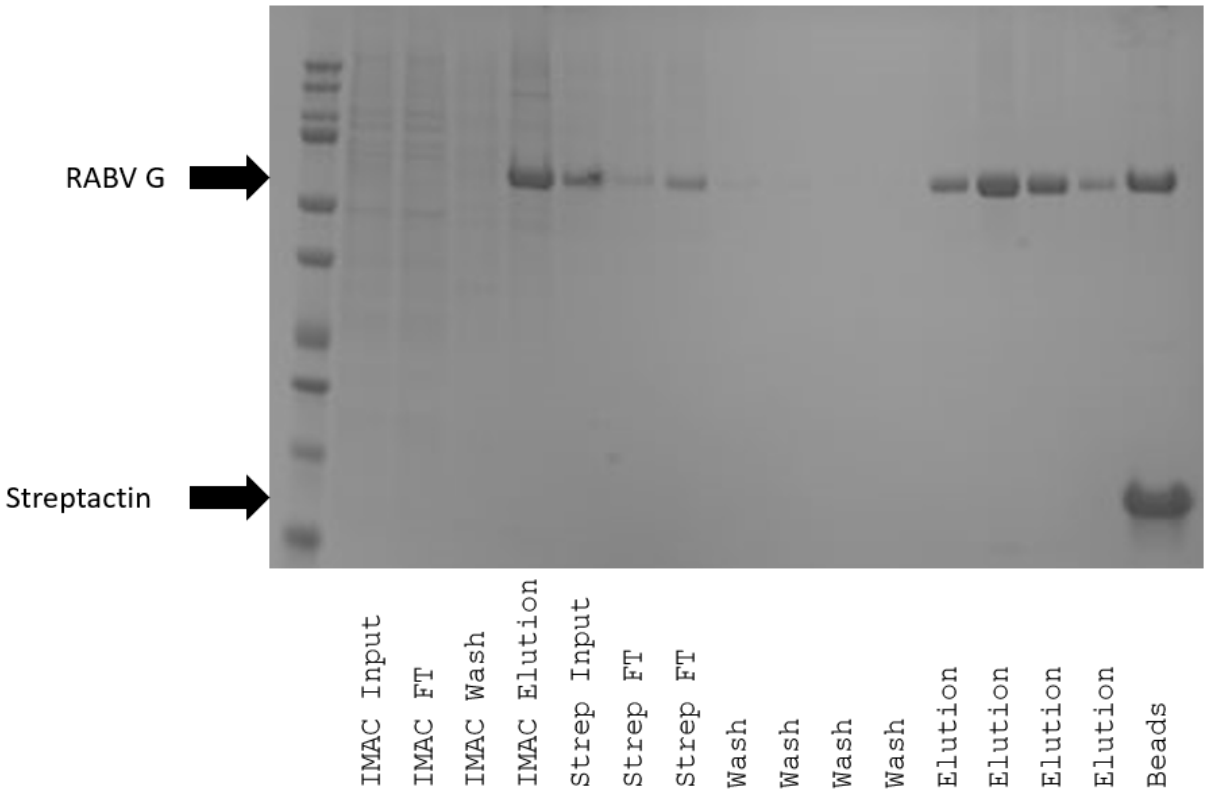
Supplementary Figure 2 | Fourier Shell Correlation of best 3D class



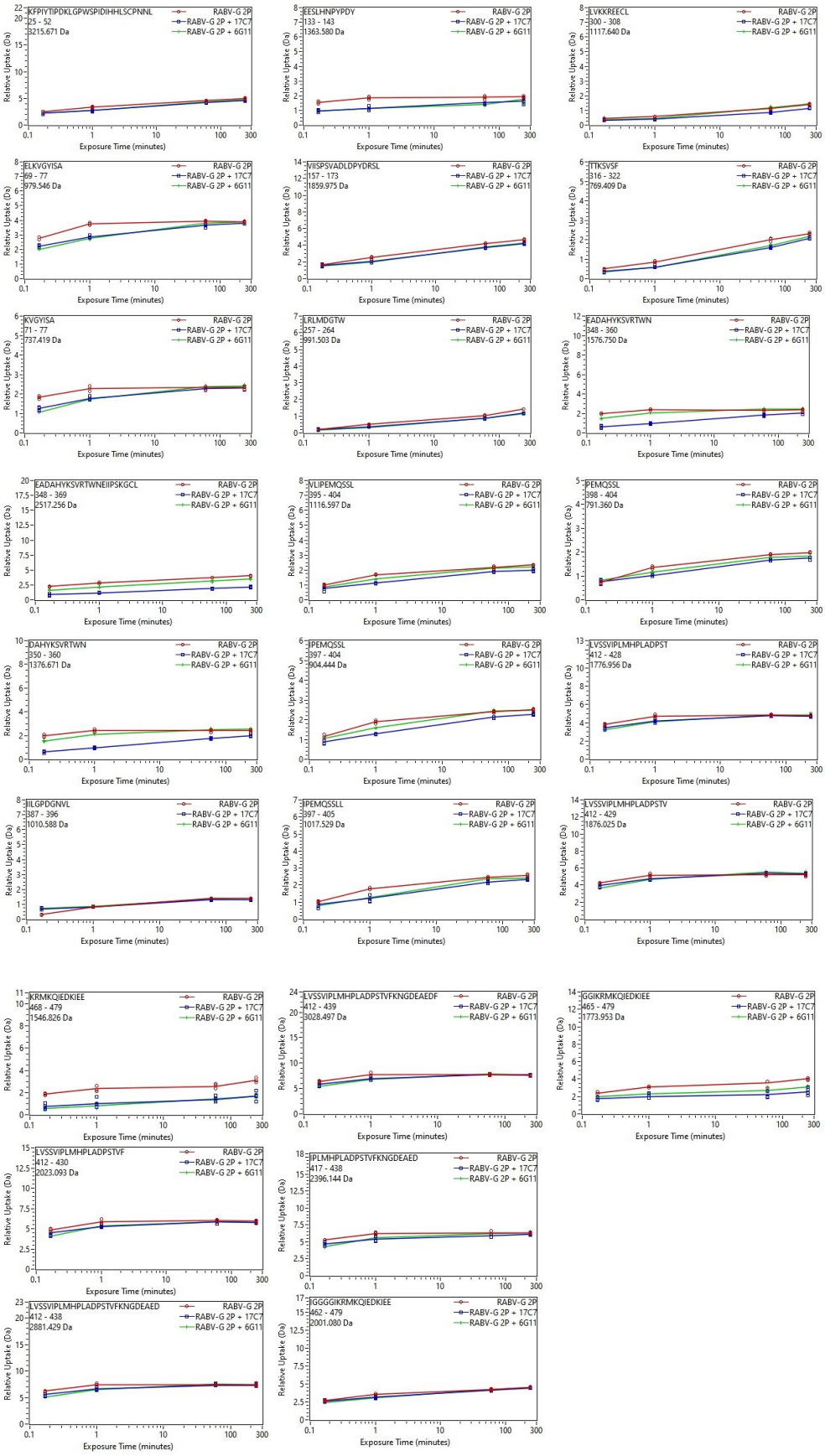
Supplementary figure 3 | SDS-PAGE of Strep-tag step of Large-Scale purification of purification of RABV G pre-fusion



Supplementary figure 4 | Size Exclusion Chromatograms of NCAM1 truncations.



Supplementary figure 5 | SDS-PAGE of Strep-tag step of purification of Large-Scale purification of 2P-mutant E288P, L290P



Supplementary figure 6 | Relative uptake of peptides of RABV G 2P-mutant per timepoint

References

[i] World Health Organization, https://www.who.int/rabies/epidemiology/en/ (30-04-2021)
[ii] Meslin, F. X., & Briggs, D. J. (2013). Eliminating canine rabies, the principal source of human infection: what will it take?. *Antiviral research*, *98*(2), 291-296.
[iii] Fisher, C. R., & Schnell, M. J. (2018). New developments in rabies vaccination. *Revue scientifique et technique (International Office of Epizootics)*, *37*(2), 657-672.
[iv] Liu, Q., & Ertl, H. C. (2012). Preventative childhood vaccination to rabies. *Expert opinion on biological therapy*, *12*(8), 1067-1075.
[v] Fooks, A. R., Banyard, A. C., Horton, D. L., Johnson, N., McElhinney, L. M., & Jackson, A. C. (2014). Current status of rabies and prospects for elimination. *The Lancet*, *384*(9951), 1389-1399.
[vi] Guo, Y., Duan, M., Wang, X., Gao, J., Guan, Z., & Zhang, M. (2019). Early events in rabies virus infection—Attachment, entry, and intracellular trafficking. *Virus research*, *263*, 217-225.
[vii] Prehaud, C., Coulon, P.., LaFay, F., Thiers, C., & Flamand, A. (1988). Antigenic site II of the rabies virus glycoprotein: structure and role in viral virulence. *Journal of virology*, *62*(1), 1-7.
[viii] Liu, X., Yang, Y., Sun, Z., Chen, J., Ai, J., Dun, C., ... & Guo, X. (2014). A recombinant rabies virus encoding two copies of the glycoprotein gene confers protection in dogs against a virulent challenge. *PloS one*, *9*(2), e87105.
[ix] Schnell, M. J., McGettigan, J. P., Wirblich, C., & Papaneri, A. (2010). The cell biology of rabies virus: using stealth to reach the brain. *Nature Reviews Microbiology*, *8*(1), 51-61.
[x] Lyles, D. S., McKenzie, M., & Parce, J. W. (1992). Subunit interactions of vesicular stomatitis virus envelope glycoprotein stabilized by binding to viral matrix protein. *Journal of virology*, *66*(1), 349-358.
[xi] Préhaud, C., Wolff, N., Terrien, E., Lafage, M., Mégret, F., Babault, N., ... & Lafon, M. (2010). Attenuation of rabies virulence: takeover by the cytoplasmic domain of its envelope protein. *Science signaling*, *3*(105), ra5-ra5.
[xii] Coll, J. M. (1995). The glycoprotein G of rhabdoviruses. *Archives of virology*, *140*(5), 827-851.
[xiii] Sissoeff, L., Mousli, M., England, P., & Tuffereau, C. (2005). Stable trimerization of recombinant rabies virus glycoprotein ectodomain is required for interaction with the p75NTR receptor. *Journal of general virology*, *86*(9), 2543-2552.
[xiv] Thoulouze, M. I., Lafage, M., Schachner, M., Hartmann, U., Cremer, H., & Lafon, M. (1998). The neural cell adhesion molecule is a receptor for rabies virus. *Journal of virology*, *72*(9), 7181-7190.
[xv] Wang, J., Wang, Z., Liu, R., Shuai, L., Wang, X., Luo, J., ... & Bu, Z. (2018). Metabotropic glutamate receptor subtype 2 is a cellular receptor for rabies virus. *PLoS pathogens*, *14*(7), e1007189.
[xvi] Lentz, T. L., Burrage, T. G., Smith, A. L., Crick, J., & Tignor, G. H. (1982). Is the acetylcholine receptor a rabies virus receptor?. *Science*, *215*(4529), 182-184.
[xvii] Belot, L., Albertini, A., & Gaudin, Y. (2019). Structural and cellular biology of rhabdovirus entry. *Advances in virus research*, *104*, 147-183.
[xviii] Gaudin, Y., Ruigrok, R. W., Tuffereau, C., Knossow, M., & Flamand, A. (1992). Rabies virus glycoprotein is a trimer. *Virology*, *187*(2), 627-632.
[xix] Baquero, E., Albertini, A. A., Vachette, P., Lepault, J., Bressanelli, S., & Gaudin, Y. (2013). Intermediate conformations during viral fusion glycoprotein structural transition. *Current opinion in virology*, *3*(2), 143-150.
[xx] Wojczyk, B. S., Stwora-Wojczyk, M., Shakin-Eshleman, S., Wunner, W. H., & Spitalnik, S. L. (1998). The role of site-specific N-glycosylation in secretion of soluble forms of rabies virus glycoprotein. *Glycobiology*, *8*(2), 121-130
[xxi] Yang, F., Lin, S., Ye, F., Yang, J., Qi, J., Chen, Z., ... & Lu, G. (2020). Structural analysis of rabies virus glycoprotein reveals pH-dependent conformational changes and interactions with a neutralizing antibody. *Cell host & microbe*, *27*(3), 441-453.
[xxii] Roche, S., Bressanelli, S., Rey, F. A., & Gaudin, Y. (2006). Crystal structure of the low-pH form of the vesicular stomatitis virus glycoprotein G. *Science*, *313*(5784), 187-191.
[xxiii] Schorcht, A., Cottrell, C. A., Pugach, P., Ringe, R. P., Han, A. X., Allen, J. D., ... & Sanders, R. W. The glycan hole area of HIV-1 envelope trimers contributes prominently to the induction of autologous neutralization. *Journal of virology*, JVI-01552.