Resilience of BST-2/Tetherin structure to single amino acid substitutions

Ian R. Roy[1], Camden K. Sutton[2], and Christopher E. Berndsen[3][4]

[1]Department of Health Sciences, James Madison University, Harrisonburg, VA 22807
[2]Department of Kinesiology, James Madison University, Harrisonburg, VA 22807
[3]Department of Chemistry and Biochemistry, James Madison University, 901 Carrier Dr. MSC 4501, Harrisonburg, VA 22087
[4]Center for Genome and Metagenome Studies, James Madison University, 901 Carrier Dr. MSC 4501, Harrisonburg, VA 22807

Corresponding author:
Christopher E. Berndsen[34]

Email address: berndsce@jmu.edu

ABSTRACT

Human Tetherin, also known as BST-2 or CD317, is a dimeric, extracellular membrane-bound protein that consists of N and C terminal membrane anchors connected by an extracellular domain. BST-2 is involved in binding enveloped viruses, such as HIV, and inhibiting viral release in addition to a role in NF-kB signaling. Viral tethering by Tetherin can be disrupted by the interaction with Vpu in HIV-1 in addition to other viral proteins. The structural mechanism of Tetherin function is not clear and the effects of human Tetherin mutations identified by sequencing consortiums are not known. To address this gap in the knowledge, we used data from the Ensembl database to construct and model known human missense mutations within the ectodomain to investigate how the structure of the ectodomain influences function. From the data, we identified an island of sequence stability within the ectodomain, which corresponds to a functionally and structurally important region identified in previous biochemical and biophysical studies. Most of the modeled mutations had little effect on the structure of the dimer and the coiled-coil, suggesting that the coiled-coil compensates for changes in primary structure. Thus, many of the functional defects observed in previous studies may not be due to changes in Tetherin structure, but in changes in protein-protein interactions or in aspects of Tetherin not currently understood. The lack of structural effects by mutations known to decrease function further illustrates the need for more study of the structure-function connection for this system. Finally, apparent flexibility in Tetherin sequence may allow for greater anti-viral activities with a larger number of viruses by reducing specific interactions with anti-Tetherin proteins, while maintaining virus restriction.

INTRODUCTION

Tetherin (BST-2, CD-317) is a membrane bound protein involved in a non-specific, immune response to enveloped viruses, such as HIV-1 and Ebola [Neil et al., 2008, Van Damme et al., 2008, Neil, 2013, Sauter, 2014]. During viral budding, one of the transmembrane anchors, usually the C-terminal anchor, becomes embedded in the virion particle membrane and prevents the diffusion of the virion away from the host cell [Perez-Caballero et al., 2009, Venkatesh and Bieniasz, 2013]. Viruses have developed numerous methods to evade Tetherin and this is an area of intense study [Sauter, 2014, McNatt et al., 2013]. Additionally, Tetherin is hypothesized to cause an inflammatory response via activation of the NF-kB pathway and may be involved in cell migration [Galão et al., 2012, Hotter et al., 2013]. The structural basis for these functions is not fully clear.

Structurally, Tetherin is a homo-dimer consisting of an N-terminal transmembrane helix and a C-terminal membrane anchor bridged by an alpha helical ectodomain (Figure 1) [Kupzig et al., 2003, Rollason et al., 2007, Andrew et al., 2011]. There are three disulfide bonds within the N-terminal portion of the ectodomain that connect to the two monomers of the dimer and appear to enhance the tensile strength of the protein [Du Pont et al., 2016]. Computational simulation of viral tethering has suggested

47   sequence encoded weaknesses in the C-terminal part of ectodomain that enable bending and flexing of
48   the ectodomain during the transition from membrane bound to bridging the virus and host cell [Ozcan
49   and Berndsen, 2017]. Despite these specific features, upwards of half of the coiled-coil region can be
50   removed without significant loss of function [Andrew et al., 2012].
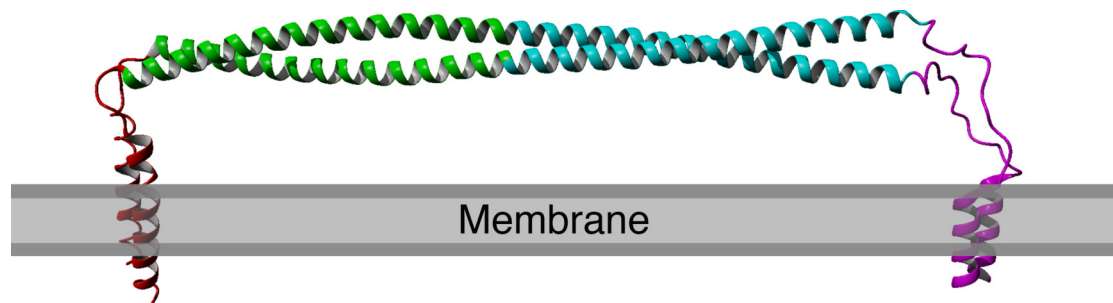
Figure 1. Structure of human Tetherin bound to the lipid membrane. Membrane anchors are shown in red (N) or magenta (C). The coiled-coil region within the ectodomain is shown in cyan. Image derived from the model produced by Ozcan and Berndsen, 2017. The C-terminal anchor is shown as an alpha helix, however there is evidence that in some forms of Tetherin, this may be a GPI anchor [Kupzig, et al., 2003, Rollason, et al., 2007]

51   A few previous studies of Tetherin structure and function have used either scanning mutagenesis or
52   analyzed a subset of known human mutations [Welbourn et al., 2015, Hammonds et al., 2012, Yang et al.,
53   2010]. Cysteine scanning mutagenesis within the ectodomain showed that mutations in the coiled-coil
54   region altered Tetherin dimerization and interfered with function, while cysteine mutations outside of this
55   region were often functional [Welbourn et al., 2015]. The new cysteines changed the orientation of the
56   helices and lead to changes in the folded state of Tetherin. Alanine scanning mutagenesis via changing
57   of the sequence four alanines at a time showed two areas of the Tetherin ectodomain were sensitive to
58   quadruple alanine mutation, however most mutations had no apparent effect [Hammonds et al., 2012].
59   The likely cause of functional defects was a change in protein localization [Hammonds et al., 2012].
60   Mutations that reversed the charge of three to six surface amino acids in the N-terminal portion of the
61   ectodomain did reduce viral tethering, however whether this was due to alterations in structure or virus
62   binding was unclear [Yang et al., 2010]. Deletions within the ectodomain also showed defects in viral
63   tethering [Yang et al., 2010]. However, a later study by Andrew and coworkers showed that deletions in
64   the ectodomain that alter the register of the alpha helix are detrimental while mutants with truncations
65   that maintain the orientation of the helices are functional [Andrew et al., 2012]. These previous examples
66   suggest that the Tetherin sequence is amenable to change as long as the alpha helical structure of the
67   ectodomain is maintained.
68   In order to more fully understand the structure-function connection in Tetherin, we analyzed structural
69   models of Tetherin containing one of 78 single nucleotide polymorphisms listed in the Ensembl database
70   [Aken et al., 2017]. In all cases, the mutations did not affect the global structure of Tetherin and instead
71   affected local structure primarily at the ends of the ectodomain. Mutations that occurred at the interface
72   of the Tetherin monomers induced greater changes in protein dynamics than those that occur on the
73   solvent-facing surface of the ectodomain. These data further support the idea that mutations in Tetherin
74   are tolerated provided that the alpha helices in the ectodomain remain intact.

METHODS

Production of Tetherin mutants

77   Tetherin mutants were generated from the membrane bound model of Tetherin described by [Ozcan and
78   Berndsen, 2017] using the list of human mutations found in Ensembl as of August 2018 [Aken et al.,
79   2017]. We wrote a YASARA macro to individually swap amino acids and minimally energy minimize the
80   resulting structure. All models were equilibrated for 10 to 20 nanoseconds in explicit water solvent and
81   a DOPC bilayer using an AMBER14 forcefield within YASARA Structure [Case et al., 2014, Krieger
82   and Vriend, 2015, Krieger et al., 2009]. Simulation conditions were 298 K, pH 7.4, 0.9% mass fraction

83 NaCl, at a density of 0.997 g/mL, with a 12 Angstrom non-bonded interaction cutoff, and a time step of
84 2.5 femtoseconds.

Analysis of Models

86 Mutant simulations were analyzed using existing macros within YASARA to calculate simulation energies,
87 root mean square fluctuation (RMSF), and the secondary structure over time. These macros were run on
88 the complete data set. Data for each mutation were summarized using Rmarkdown. Root mean square
89 deviation values were calculated using the Align function in YASARA structure. Global analysis of data
90 was performed in R using the components within the *tidyverse* package and the *cluster* package [Ross
91 et al., 2017, Wickham, 2009, Maechler et al., 2018]. Ridgeline plots and some figures were made using
92 the *ggridges* and *cowplot* packages of Claus Wilke [Wilke, 2018a,b].

1 RESULTS

1.1 Human missense mutations are absent from beginning of coiled-coil

95 The structure of Tetherin has been determined however the connection between the crucial functions of
96 Tetherin and this structure is less clear [Swiecki et al., 2011, Hinz et al., 2010, Schubert et al., 2010, Yang
97 et al., 2010]. The primary sequence of Tetherin is not well-conserved across species, although the general
98 organization appears to be [Heusinger et al., 2015, Blanco-Melo et al., 2016]. Areas of structural or
99 functional importance generally have a conserved protein sequence. Given the lack of apparent sequence
100 conservation between species, it is difficult to predict the important structural features of Tetherin. We
101 decided to compare sequences within the human species to identify areas of amino acid stability and
102 therefore relative importance. We retrieved variation data from the Ensembl database containing the
103 missense mutations within the ectodomain of Tetherin (Supplemental Table 1). We focused on the
104 mutations within the helical portion of the ectodomain because there is better, high quality structural
105 information on this region. Mapping the amino acid locations of SNPs showed that there was a lower
106 density of mutations between residues 93 and 117 (Figure 2A). We note that synonymous mutations,
107 which do not affect the amino acid sequence, still occur in this region, but amino acid changes appear to
108 be lacking (Figure 2A). Although the data set is limited in scope, the information suggests that amino
109 acids in this region, which includes one disulfide and the beginning of the coiled-coil region of the
110 ectodomain, are important for Tetherin structure or function. In addition to a general position search, we
111 mapped the mutations on the 3MQB structure and observed that 66 of 78 mutations appear to occur on
112 the exterior of the protein ("outside") rather than at the interface of the Tetherin dimer ("inside") (Figure
113 2B) [Yang et al., 2010]. Moreover, the frequency of many of these outside mutations is higher than for
114 inside mutations (Supplemental Table 1). The simplest interpretation of this distribution of mutations
115 is that the solvent and membrane facing surfaces of the protein are more tolerant of single amino acid
116 changes. This distribution supports the previously demonstrated role of dimerization in viral tethering
117 and stability, although dimerization is not entirely essential for tethering of all viruses [Andrew et al.,
118 2009, Sakuma et al., 2009, Schubert et al., 2010, Welbourn et al., 2015, Du Pont et al., 2016]. The 78
119 mutations occur at 59 unique sites within the protein and some positions have several polymorphisms
120 (summarized in Supplemental Table 1). We also observed that the areas of highest mutation density
121 were in the N-terminal portion of the ectodomain, which has been suggested to be more flexible and
122 is stabilized by the inter-molecular disulfides [Welbourn et al., 2015, Du Pont et al., 2016], or at the
123 C-terminal end of the coiled-coil region, which in structural simulations unwinds relatively easily as
124 it contains a non-canonical heptad [Andrew et al., 2012, Du Pont et al., 2016, Ozcan and Berndsen,
125 2017]. Therefore the locations of SNPs align well with previous functional and structural data on Tetherin.
126 However, the potential effects of these mutations on Tetherin structure are unclear.

1.2 Missense mutations do not alter overall structure

128 To determine the association between structure and function in Tetherin mutants, we computationally
129 made point mutations in the previously described full-length model of Tetherin at positions between
130 49 and 163, where the structure of Tetherin appears mostly helical [Ozcan and Berndsen, 2017]. All
131 models equilibrated within 10 to 20 ns of simulation. We then compared the root mean square deviation,
132 root mean square fluctuation, and secondary structure for all mutations to a similar simulation on a
133 model lacking mutations to identify structural and dynamic changes due to the mutation. The root mean
134 square deviation (RMSD) is a measure of the differences in the 3-D position of one or more elements of
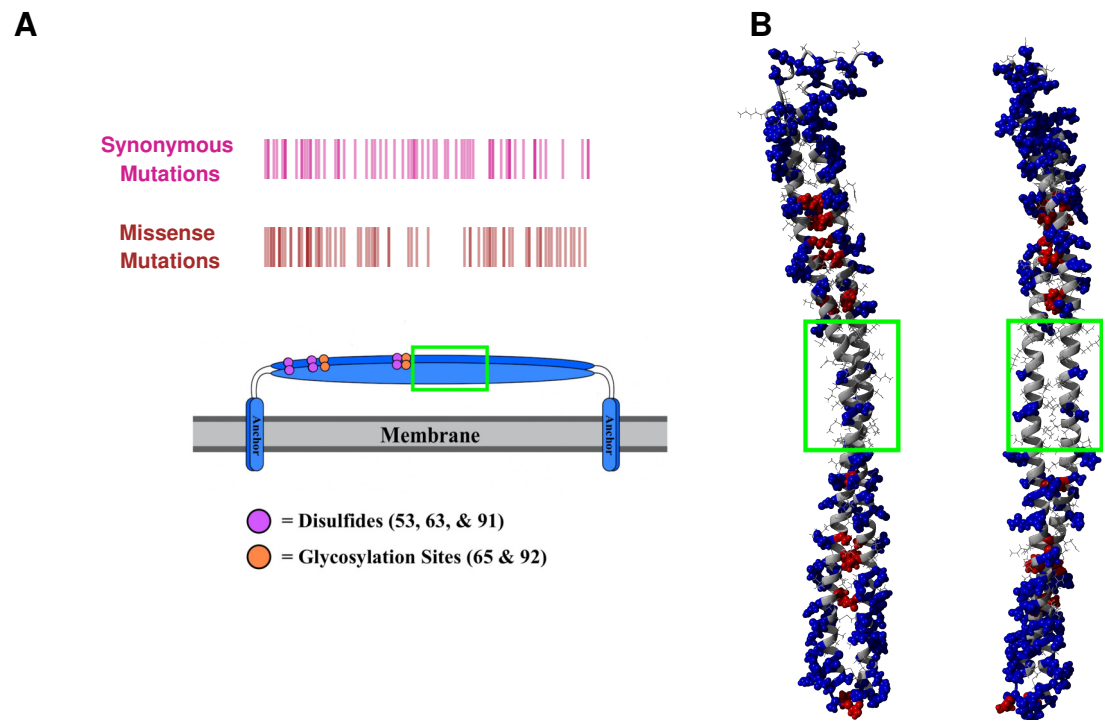
Figure 2. Human mutations within Tetherin (A) Cartoon showing domain organization and location of synonymous (purple) and missense mutations (red) in Tetherin (B) Model showing the location of the mutants on the inside (red) or outside (blue) of the Tetherin ectodomain. The green box in (A) and (B) indicates the location of amino acids 93-117

the ectodomain. We calculated the RMSD of the average structure from each mutant simulation from alignment of the ectodomain alone with a wild-type Tetherin that had also been equilibrated. The data were then plotted as RMSD vs. amino acid number as shown in Figure 3. Any large RMSD values would suggest a strong effect of the mutation on the overall Tetherin structure. We observed that there were distinctly large RMSD values across the ectodomain mutations and no single mutation disrupted the entire coiled-coil (Figure 3). These data suggest that mutations in Tetherin do not cause major changes to the helical structure.

1.3 Mutations alter local Tetherin structure

Single changes to the amino acid sequence may not cause drastic changes to protein structure. Therefore, we next analyzed our simulations for local changes in helical structure. Based on the distance between hydrogen bond donors and acceptors, amino acid positions were binned into types of secondary structure at 400 steps throughout the simulation. Representative plots of secondary structure across the simulation are shown in Figure 4. In addition, we simulated a point mutation known to affect the structure of Tetherin, Gly118Cys, as a positive control for structural disruptions [Welbourn et al., 2015]. The wild-type simulation shows a stable helical structure which is matched in behavior by the Asp129Glu and Asp103Asn mutations. The postive control, Gly118Cys, shows a chunk of 4 amino acids which are in a coiled structure, but the effects are limited to this region. Similarly, mutations Arg64Pro, Ala100Pro and Ile120Phe showed distinct regions where the secondary structure is altered compared to the wild-type as shown in Figure 4A. Proline is well-known to be disruptive to alpha helices because it is unable to form the hydrogen bonds characteristic of this secondary structure [Pace and Scholtz, 1998]. Position 64 within Tetherin is the most diversely mutated position in the data that we analyzed with 5 SNPs but only the Arg64Pro had any effects on structure (Supplemental Table 1). Phenylalanine is aromatic and occupies a larger volume than isoleucine suggesting the local secondary structure changes are due to the alterations in the interface between the two helices. We note that none of the changes in secondary
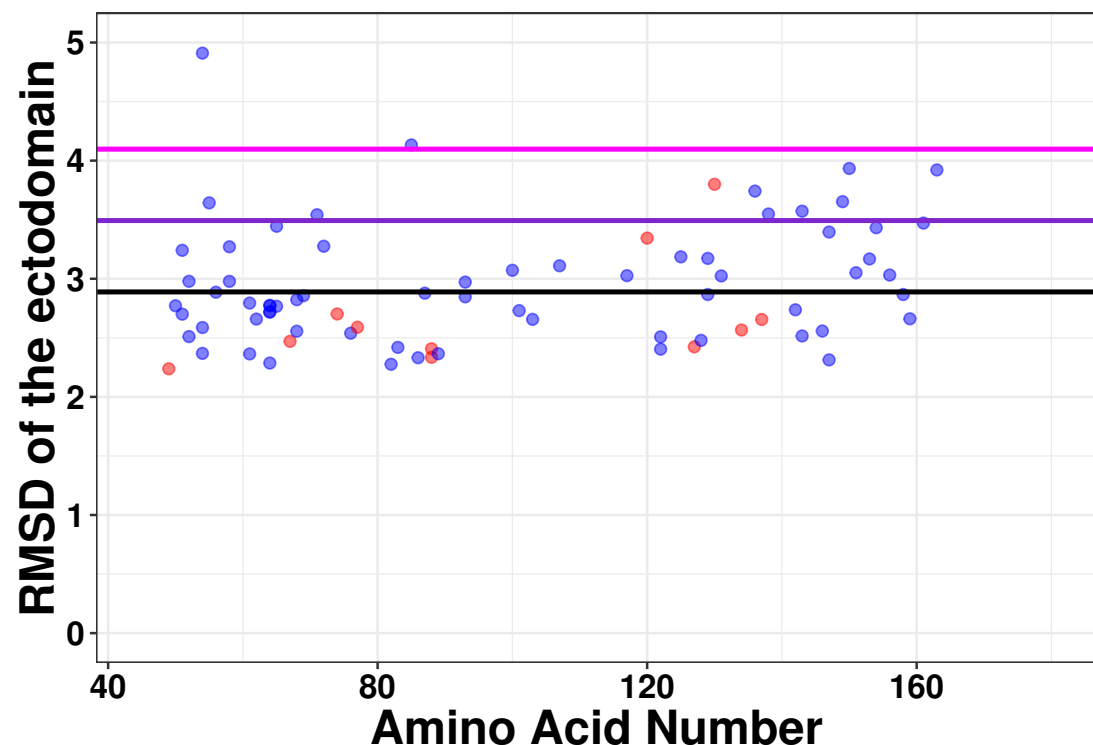
Figure 3. No large changes in structure induced by mutations as shown by the root mean square deviation plotted against amino acid number. Red points are those positions that are found on the inside of the Tetherin dimer, while blue points are those positions on the outside of the dimer. Black, purple, and magenta lines show the mean, mean + one and mean + two standard deviations of the RMSD values respectively. Mutations at positions C-terminal to amino acid 163 were not analyzed because this region appears unstructured in crystal structures [Yang, et al., 2010, Hinz, et al., 2010, Swiecki, et al., 2011]

159 structure spread along the coil. Instead, cracks in the helices stayed localized to a single area within the
160 ectodomain, supporting the biochemical and computational findings of [Du Pont et al., 2016]. We next
161 analyzed where the changes in structure were localized and plotted the number of times an amino acid
162 adopted a different environment versus position (Figure 4B). The effects on secondary structure appear
163 primarily at the ends of the coil, which are known to be unstructured, rather than proximal to the site of
164 substitution except for a few mutations (Figure 4B) [Swiecki et al., 2011]. These data suggest that the
165 potential disruptive effects of most mutations are minimized by the coiled-coil structure of Tetherin.
166   While most mutations had minimal effect on the helical structure of Tetherin, three mutations
167 Arg64Pro, Ala100Pro, and Ile120Phe did disrupt the helices. Arg64Pro and Ala100Pro induced alterations
168 in secondary structure near to the site of the polymorphism. In neither of these instances, do the alterations
169 in structure include more than two consecutive amino acids. Ile120Phe however induces a disruption in
170 secondary structure that is not directly adjacent to the mutation site and is maintained throughout the
171 simulation. For Ala100Pro mutation this disruption in secondary structure was reproduced while for
172 Ile120Phe, the disruption in secondary structure was inconsistent (Supplemental Figure 1). Visualization
173 of the energy minimum structure from the simulation for both the wild-type and Ile120Phe structure,
174 shows that Ile120 is an interior amino acid and the side chains could form a van der Waals contact (Figure
175 5A). However, in the structure of the Phe substitution, the side chain is displaced from the interface
176 (Figure 5A). Further comparison of the distance between the Tetherin monomers and at sites along the
177 helix shows that the Phe substitution noticeably changes the distribution of measurements on either side
178 of the mutation over a region extending at least 56 amino acids in length (Figure 5B). The alteration in the
179 amino acid distances appeared consistently in replicate simulations of this mutation despite the lack of
180 effect on secondary structure in some simulations (Supplemental Figure 2). In comparison, the Ala100Pro
181 mutation does not alter distribution of distances except for slight differences adjacent to the mutation
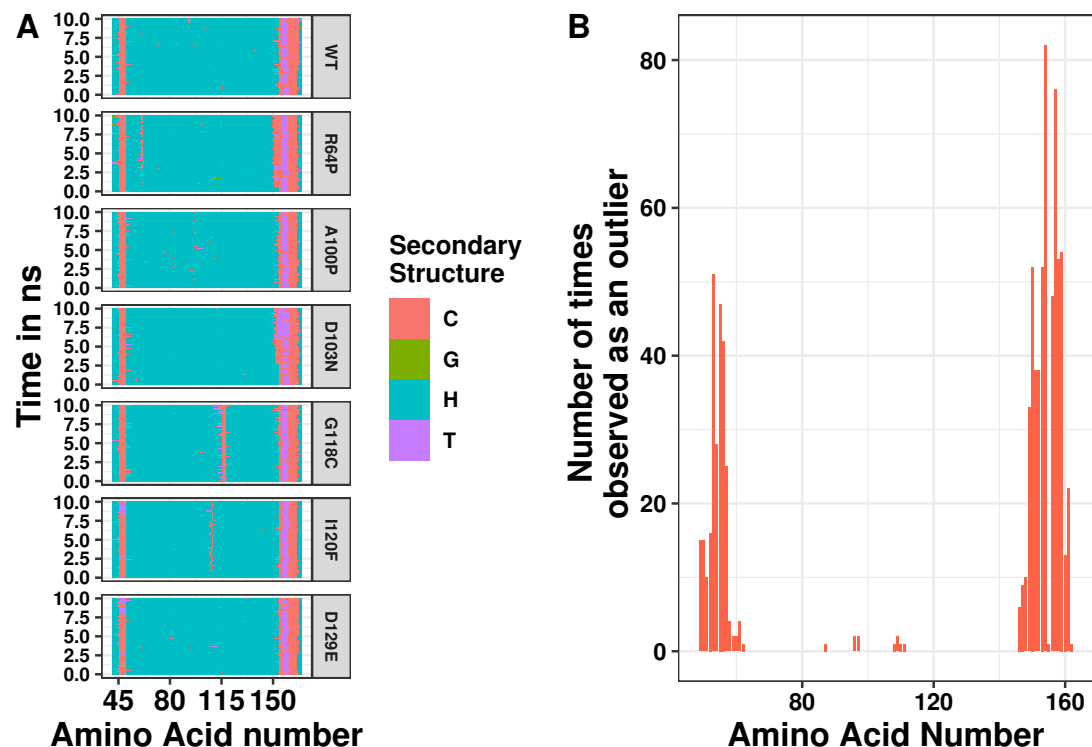
Figure 4. (A) Secondary structure of Tetherin mutant simulations over time for Arg64Pro, Ala100Pro, Gly118Cys, Ile120Phe, Asp129Glu, and Asp103Asn. Secondary structures were assigned by YASARA based on hydrogen bonding distances and psi angles at 25 picosecond intervals. 'H' stands for alpha helix, 'T' for turn, 'C' for coil, and 'G' for 3/10 helix. (B) Plot showing the cumulative number of times that the secondary structure was perturbed at each location for all simulations. Outliers were identified as sites that had distinct secondary structure more than two standard deviations above the average secondary structure value at that site across all simulations. A value of one means that the amino acid appeared as an outlier in 1 simulation.

site (Figure 5C). Observation of the individual measurement sites over time show that amino acids in the region of 105 to 123 quickly adjust to a new distance followed by slower adjustments at positions 81, 84 and 148 (Supplemental Figure 3). This behavior is not observed in the wild-type simulations (Supplemental Figure 3). These data suggest that the structure of Tetherin can tolerate many changes in the sequence and that for most changes that disrupt the helical structure there is limited spreading of breaks.

1.4 Effect of mutations on ectodomain dynamics

Protein structure is not just based on the static arrangement of amino acids, but the ensemble of structures that the protein can adopt, which is sometimes referred to as protein dynamics. To quantify dynamics, we analyzed the root mean square fluctuation (RMSF), which describes the amount of movement each amino acid undergoes during the simulation. We then correlated the per residue RMSF values between each mutation simulation and the wild-type simulation. We observed moderate correlations at most positions but wanted additional methods to analyze the shape of the cluster. Using the central point in the cluster would allow for easier visualization of how each mutation affects the global dynamics of the system and allow for easier comparison of the effects of mutations. Therefore, we next calculated the mean and medoid centers of the RMSF clusters. The data for the medoid and mean cluster centers in the mutants correlated well suggesting that our data are normally distributed (Figure 6). We plotted all the data determined by the k-means method at each amino acid position to determine if there were any trends or clustering. As seen in the right pane of Figure 6, mutations at the ends of the ectodomain tended to have larger k-mean values suggesting larger RMSF values and more dynamics in the simulation. Mutations on
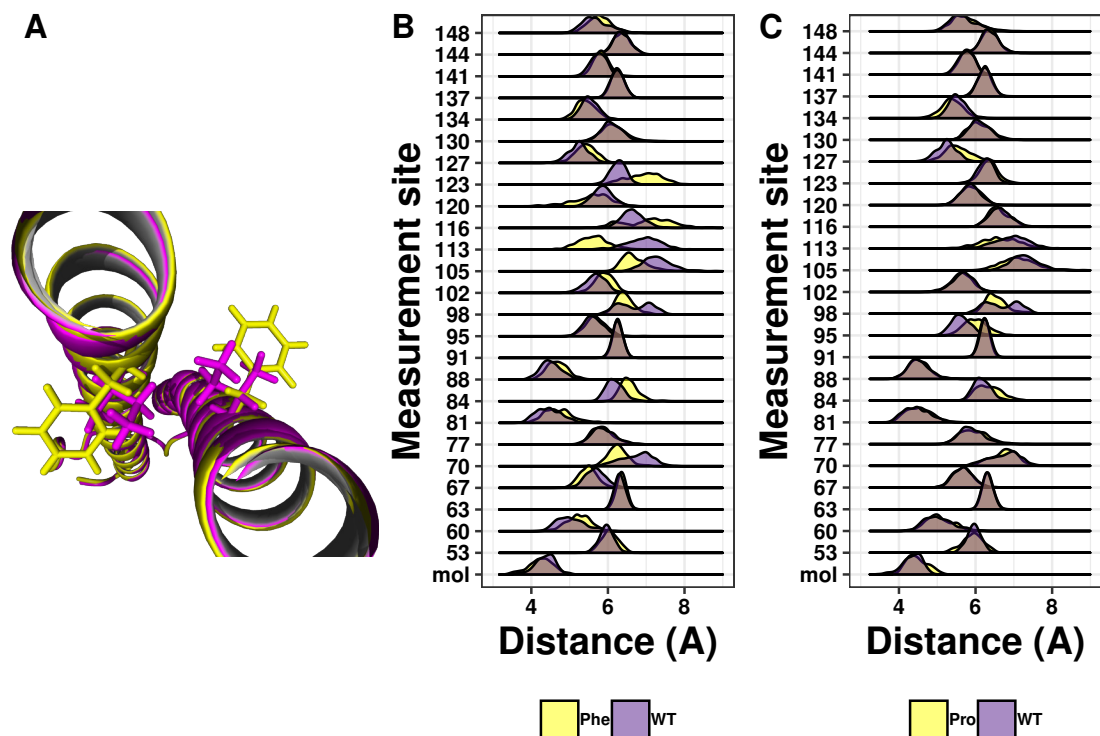
Figure 5. (A) Differences in side chain position between the wild-type Isoleucine (purple) and the mutation Phenylalanine (yellow) (B) Ridgeline plot showing the range of distances between the A and B molecules (mol) or the Calpha at the indicated positions comparing the Ile120Phe simulation (yellow) and the wild-type simulation (purple). Overlap is shown in pink. (C) Ridgeline plot showing the range of distances between the A and B molecules (mol) or the Calpha at the indicated positions comparing the Ala100Pro simulation (yellow) and the wild-type simulation (purple). Overlap between the two conditions is indicated by grey coloring.

the interface of the two helices did not increase dynamics relative to the outside mutations. These data are consistent with our previous finding that disruptions to secondary structure occurred preferentially toward the termini of the ectodomain and that the ectodomain can tolerate some changes to the structure.

1.5 Comparison with Tetherin mutations with known functional defects

Given the general lack of effect of point mutations on the structure of Tetherin, we next looked at known non-functional mutations described previously in the literature (Supplemental Table 2). We have shown that cysteine mutations or truncations that alter the orientation of the helices resulting in a poor interface are non-functional [Welbourn et al., 2015, Andrew et al., 2012]. However, others have shown multiple alanine or charge swap mutations can also reduce viral tethering [Yang et al., 2010, Hammonds et al., 2012]. Yang and coworkers proposed that swapping the charge of amino acids in the N-terminus altered the interaction of Tetherin with the lipid membrane. In Figure 7A, we suggest that this is unlikely because the mutation sites are not on a single face of the coiled-coil and not all sites interact with the membrane in our model. However, we cannot rule out a mechanism where the orientation of the helices relative to the membrane are maintained via interactions between these amino acids and the lipid heads. Simulations show a relatively stable distance between the coils with few disruptions outside of the sites of mutation (Figures 7B-C).

We next analyzed simulations containing the tetra-alanine and tetra-serine mutations of Tetherin by [Hammonds et al., 2012]. As with the charge swap mutations, the tetra-alanine mutations show essentially no structural changes (Figure 8). In contrast, the tetra-serine substitution showed large but localized changes in structure as assessed by distance measurements (Figure 8C). Given that the tetra-serine mutation sites are all at the interface of the coiled-coil, these data further link alterations of the coiled-coil
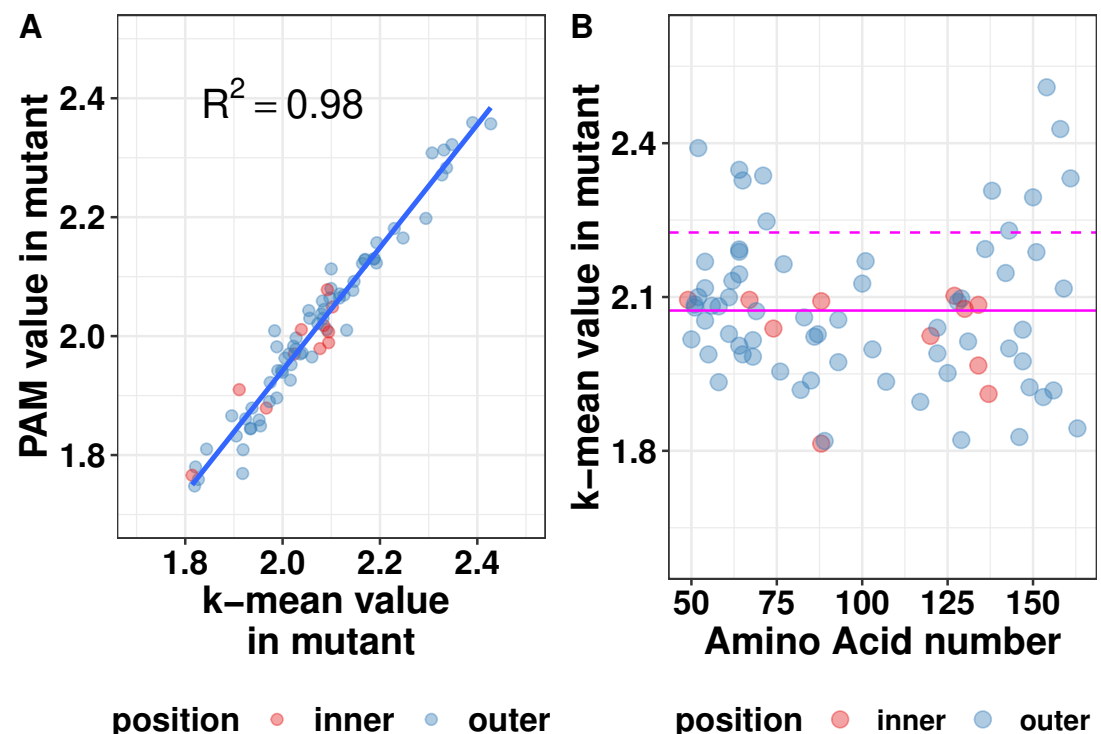
Figure 6. Mutation effects on protein dynamics (A) correlation of k-means and PAM values (B) scatterplot of k-mean values at each amino acid position. Magenta lines indicate the mean value (solid) and the mean plus one standard deviation (dashed) of the k-means

interface to defects in function [Welbourn et al., 2015, Andrew et al., 2012, Hammonds et al., 2012, Yang et al., 2010, Du Pont et al., 2016]. Additionally, these results match up to our results with the single point mutations showing limited structural effects induced by mutation.

DISCUSSION

The connection between Tetherin structure and function continues to be a significant question in the field. Putative orthologs of Tetherin show little sequence similarity and there are only structures for two orthologs [Hinz et al., 2010, Swiecki et al., 2011, Schubert et al., 2010, Yang et al., 2010, Heusinger et al., 2015, Blanco-Melo et al., 2016]. In order to gain more insight into the structural pliability of Tetherin, we analyzed the location and effects of Tetherin mutations found in the human genome. While the data set that we analyzed only includes data from those genome studies that deposit in Ensembl, this is the most comprehensive set of genomes available. Initial analysis showed a relative lack of missense mutations between amino acids 93 and 117. This section of Tetherin includes one cysteine which can form a disulfide, a glycosylation site, and also the first three heptads of the coiled-coil [Andrew et al., 2009, 2012]. While there are synonymous sequence changes in the DNA sequence along the entire length of the protein, this section lacks mutations that change the protein sequence. Our analysis of these mutations supports previous work showing this region as being important for flexibility or hinge motions [Hinz et al., 2010, Yang et al., 2010, Andrew et al., 2012, Ozcan and Berndsen, 2017]. Collectively, these data suggest that the flexibility of this region is tuned for a specific structure and/or function. Previous work suggested that this purpose could be to "break" in a controlled fashion during the transition from the membrane bound to the cell-virus bridging form [Ozcan and Berndsen, 2017]. However, more information is needed on the molecular functions of Tetherin to explore and test this idea fully.

Mapping of the mutations onto the structure of Tetherin showed that most mutations occur on amino acids on the exterior of the helices rather than at positions at the dimer interface (Figure 2) [Yang et al., 2010]. The importance of Tetherin dimerization for most, but not all, function is well-established,
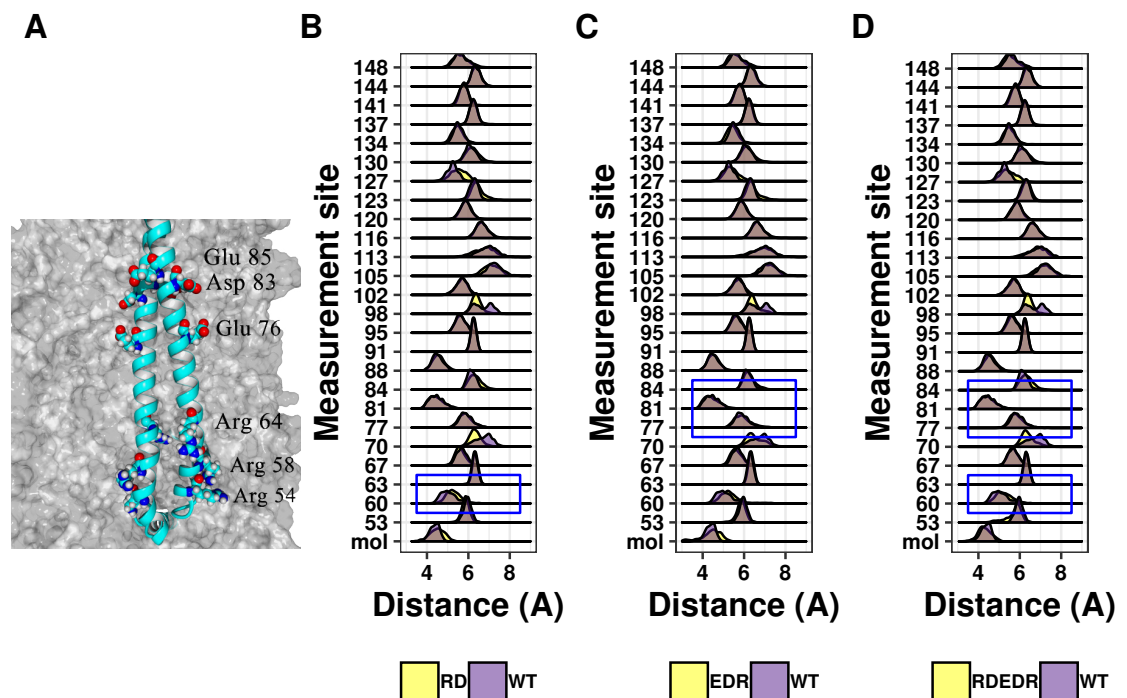
Figure 7. (A) Position of the charge swap mutations within the Tetherin coiled-coil (B)-(D) Ridgeline plots showing the range of distances between the A and B molecules (mol) or the Calpha at the indicated positions comparing the charge swap mutant simulations (yellow) and the wild-type simulation (purple). Overlap is shown in grey A blue box indicates the area where these mutants are within the sequence. (B) Arg54, Arg 58, Arg64 to aspartate. (C) Glu76, Asp83, Glu85 to arginine. (D) Both sets of mutants.

247 therefore fewer interior mutations was not surprising [Andrew et al., 2009, Sakuma et al., 2009, Schubert
248 et al., 2010, Welbourn et al., 2015, Du Pont et al., 2016]. Structurally, most amino acid substitutions,
249 except glycine and proline, have similar propensities to form an alpha helix, therefore a single amino
250 acid change regardless if it is on the inside or outside of the coil should not alter the helical structure of
251 Tetherin to a significant degree [Pace and Scholtz, 1998]. Why then is the inside vs outside distribution
252 interesting? Our data and analysis show that mutations on the inside of the coiled-coil occur less frequently
253 (Supplemental Table 1) and some show disruptions of the coiled-coil structure. If the mutations are likely
254 not affecting the helical structure, this suggests the lower mutation rate of inside amino acids may be to
255 avoid disrupting the interactions between the coils. One inside mutation that demonstrates this, Ile120Phe,
256 showed changes, although consistently, to the helical structure of the protein (Figure 4 and Supplemental
257 Figure 1). We linked changes that did appear to displacement of the side chain which led to changes in
258 the distance between amino acids in each monomer of the dimer (Figure 5 and Supplemental Figures 2
259 and 3). The apparent disruption of the Tetherin dimer interface and relative change in the orientation of
260 the monomers was transmitted down the coiled-coil. These data suggest that the Tetherin dimer limits
261 disruptions to localized regions and that more drastic reorganizations of the dimer are required to induce
262 unfolding of the structure, such as those from cysteine scanning mutagenesis performed by Welbourn
263 et al. [2015]. While not all inside mutations caused drastic effects, the database where we obtained our
264 SNP information likely does not contain many SNPs where the protein disrupted function enough to alter
265 organism survival. Comparison of our simulations to simulations of mutations with known functional
266 defects further demonstrated the relatively malleability of the Tetherin coiled-coil. The charge swap
267 mutations of Yang and coworkers showed minimal observable changes in structure or distance between
268 the monomers suggesting that the functional defects in these mutant proteins are not due to changes in the
269 structure of Tetherin (Figure 7) [Yang et al., 2010]. The proximity of these mutations to the glycosylation
270 sites and a region linked to trafficking of Tetherin may be the source of the functional defects as shown by
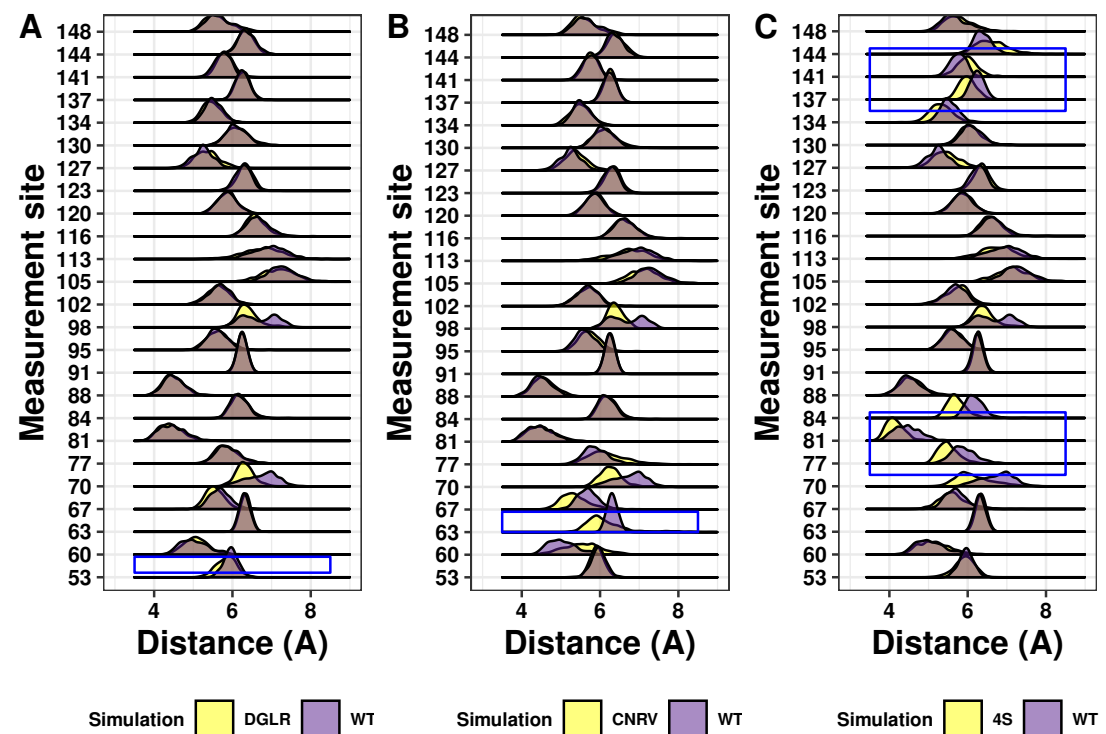
Figure 8. Ridgeline plots showing the range of distances between the Calpha at the indicated positions comparing the tetra-substituted simulation (yellow) and the wild-type simulation (purple). Overlap is shown in grey. Blue boxes indicate the region where the mutation occurs within the protein. (A) Asp55, Gly 56, Leu57, Arg58 to alanine (B) Cys63, Asn64, Arg65, Val66 to alanine. (C) Val74, Val84, Leu137, Leu144 to serine.

mutations in this region by [Yang et al., 2010, Hinz et al., 2010, Hammonds et al., 2012, Waheed et al., 2018]. In the region between residues 54 and 85, there are 28 missense SNPs listed in Ensembl including two at the glycosylation site at position 65 (Figure 2A and Supplemental Table 1). Only one (Arg64Pro) showed any effects on structure (Figure 4A), further suggesting that the functional defects in this region are due to non-structural changes in the protein.

Analysis of alanine scanning or tetra-serine mutations from Hammonds and coworkers showed localized effects on the orientation of the helices (Figure 8) [Hammonds et al., 2012]. The most apparent was for the tetra-serine mutation, which changed four interface amino acids to serine and showed noticiable changes in the distance between the helices local to the sites of mutation (Figure 8). The tetra-serine mutation did not however alter secondary structure as we observed for the Ile120Phe, Arg64Pro, or Ala100Pro mutations (Figure 4). In the two regions encompassed by the tetra-serine mutations, there are a total of 10 SNPs in Ensembl, although none of these mutations showed apparent effects on Tetherin structure. Given the similarity in effects on the orientation of the helices between the Ile120Phe mutation and the tetra-serine mutation, it is tempting to speculate that the Ile120Phe mutant likely has reduced function, however this proposal awaits experimental confirmation of the computational predictions in this present study.

In addition to the genetic stability of the middle section of Tetherin, we observed that most mutations had little if any apparent effect on the overall structure of Tetherin. Calculation of the RMSD between the average structure during simulations and the wild-type protein showed no trends or large changes in the structure. Supporting these data were analysis of the secondary structure at each amino acid which showed most mutations appears to affect the ends of the ectodomain, which are known to be largely unstructured [Hinz et al., 2010, Andrew et al., 2012, Ozcan and Berndsen, 2017, Swiecki et al., 2011, Schubert et al., 2010, Yang et al., 2010]. Given the importance of the alpha helical structure of Tetherin to tensile strength, these findings are not surprising, although certainly unexpected given the number

and diversity of mutations analyzed [Du Pont et al., 2016, Andrew et al., 2012]. This is most apparent in the Arg64Pro and Ala100Pro mutations, which we expected would cause drastic changes in Tetherin structure and dynamics. While the proline mutation does cause local changes in secondary structure (Figure 4), the effects of this mutation are localized to a few amino acids proximal to the mutation site. The detrimental effects of proline in helices are well-known, however the strength of the coiled-coil and the disulfides appears to prevent spreading of this "crack" in the helix [Pace and Scholtz, 1998, Du Pont et al., 2016]. Previously, Sauter and coworkers analyzed the functional effects of seven rare mutations in Tetherin from the human genome [Sauter et al., 2013]. For the mutations located in the ectodomain, the Asn49Ser mutation affected NF-κB activation, possibly due to changes in cell surface expression while the other mutations had no effects on Tetherin function [Hammonds et al., 2012, Sauter et al., 2013]. The Asn49Ser mutation had no apparent effects on structure in the present study suggesting that the functional or expression defects are due to another unknown aspect of Tetherin.

The computational approach used by this study showed that for a limited test set most of the effects on Tetherin structure were reproducible. However, the limitations of computational approaches are clear and we urge against overly broad interpretations of the data.

2 CONCLUSIONS

Immediately prior to submission of this paper in 2019, we searched ClinVar for disease causing mutations in Tetherin and we did not find any single nucleotide polymorphisms in the ectodomain with known links to diseases or pathogenic phenotypes. These data lead to the idea that mutations are tolerated within Tetherin as long as they do not affect the overall helical structure and/or dimerization. Given that Tetherin has no known enzymatic functions and the interactions of Tetherin with viruses like HIV-1 can rapidly adapt to immune system, this sequence flexibility is likely advantageous as it could decrease specific interactions with proteins that antagonize Tetherin anti-viral activities. The structural stability of the coiled-coil structure in the face of a changing primary structure would permit viral tethering, while allowing evasion of viral proteins.

ACKNOWLEDGMENTS

We thank the James Madison University Department of Chemistry and Biochemistry, Yasmeen Shorish, Kadir A. Ozcan, and the 4 – VA Organization for the resources, professional support, and models used throughout the project. We also thank the CHEM 361 students in Biochemistry 1 during Fall 2017 for their semester long work which helped shape the direction of the investigation.

REFERENCES

Bronwen L. Aken, Premanand Achuthan, Wasiu Akanni, M. Ridwan Amode, Friederike Bernsdorff, Jyothish Bhai, Konstantinos Billis, Denise Carvalho-Silva, Carla Cummins, Peter Clapham, Laurent Gil, Carlos García Girón, Leo Gordon, Thibaut Hourlier, Sarah E. Hunt, Sophie H. Janacek, Thomas Juettemann, Stephen Keenan, Matthew R. Laird, Ilias Lavidas, Thomas Maurel, William McLaren, Benjamin Moore, Daniel N. Murphy, Rishi Nag, Victoria Newman, Michael Nuhn, Chuang Kee Ong, Anne Parker, Mateus Patricio, Harpreet Singh Riat, Daniel Sheppard, Helen Sparrow, Kieron Taylor, Anja Thormann, Alessandro Vullo, Brandon Walts, Steven P. Wilder, Amonida Zadissa, Myrto Kostadima, Fergal J. Martin, Matthieu Muffato, Emily Perry, Magali Ruffier, Daniel M. Staines, Stephen J. Trevanion, Fiona Cunningham, Andrew Yates, Daniel R. Zerbino, and Paul Flicek. Ensembl 2017. *Nucleic Acids Res.*, 2017. ISSN 13624962. doi: 10.1093/nar/gkw1104.

Amy J Andrew, Eri Miyagi, Sandra Kao, and Klaus Strebel. The formation of cysteine-linked dimers of BST-2/tetherin is important for inhibition of HIV-1 virus release but not for sensitivity to Vpu. *Retrovirology*, 6:80, jan 2009. ISSN 1742-4690. doi: 10.1186/1742-4690-6-80.

Amy J Andrew, Eri Miyagi, and Klaus Strebel. Differential effects of human immunodeficiency virus type 1 Vpu on the stability of BST-2/tetherin. *J. Virol.*, 85(6):2611–2619, 2011. ISSN 1098-5514. doi: 10.1128/JVI.02080-10.

Amy J. Andrew, Christopher E. Berndsen, Sandra Kao, and Klaus Strebel. The size and conservation of a coiled-coil structure in the ectodomain of human BST-2/tetherin is dispensable for inhibition of HIV-1

virion release. *J. Biol. Chem.*, 287(53):44278–44288, 2012. ISSN 00219258. doi: 10.1074/jbc.M112. 418822.

Daniel Blanco-Melo, Siddarth Venkatesh, and Paul D. Bieniasz. Origins and Evolution of tetherin, an Orphan Antiviral Gene. *Cell Host Microbe*, 20(2):189–201, 2016. ISSN 19346069. doi: 10.1016/j. chom.2016.06.007.

D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. III Cheatham, T.A. Darden, R.E. Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, and P.A. Kollman. *Amber14 Reference manual*. 2014. ISBN 9780874216561. doi: 10.1007/s13398-014-0173-7.2.

Kelly E. Du Pont, Aidan M. McKenzie, Oleksandr Kokhan, Isaiah Sumner, and Christopher E. Berndsen. The Disulfide Bonds within BST-2 Enhance Tensile Strength during Viral Tethering. *Biochemistry*, 55 (6):940–947, 2016. ISSN 15204995. doi: 10.1021/acs.biochem.5b01362.

Rui Pedro Galão, Anna Le Tortorec, Suzanne Pickering, Tonya Kueck, and Stuart J.D. Neil. Innate sensing of HIV-1 assembly by tetherin induces NFκB-dependent proinflammatory responses. *Cell Host Microbe*, 12(5):633–644, 2012. ISSN 19313128. doi: 10.1016/j.chom.2012.10.007.

Jason Hammonds, Lingmei Ding, Hin Chu, Ken Geller, Andrew Robbins, Jaang-Jiun Wang, Hong Yi, and Paul Spearman. The tetherin/BST-2 coiled-coil ectodomain mediates plasma membrane microdomain localization and restriction of particle release. *J. Virol.*, 86(4):2259–72, feb 2012. ISSN 1098-5514. doi: 10.1128/JVI.05906-11.

Elena Heusinger, Silvia F. Kluge, Frank Kirchhoff, and Daniel Sauter. Early Vertebrate Evolution of the Host Restriction Factor Tetherin. *J. Virol.*, 89(23):12154–12165, 2015. ISSN 0022-538X. doi: 10.1128/ JVI.02149-15. URL http://jvi.asm.org/lookup/doi/10.1128/JVI.02149-15.

Andreas Hinz, Nolwenn Miguet, Ganesh Natrajan, Yoshiko Usami, Hikaru Yamanaka, Patricia Renesto, Bettina Hartlieb, Andrew A. McCarthy, Jean Pierre Simorre, Heinrich Göttlinger, and Winfried Weissenhorn. Structural basis of HIV-1 tethering to membranes by the BST-2/tetherin ectodomain. *Cell Host Microbe*, 7(4):314–323, 2010. ISSN 19313128. doi: 10.1016/j.chom.2010.03.005.

Dominik Hotter, Daniel Sauter, and Frank Kirchhoff. Emerging role of the host restriction factor tetherin in viral immune sensing. *J. Mol. Biol.*, 425(24):4956–4964, 2013. ISSN 00222836. doi: 10.1016/j.jmb.2013.09.029.

Elmar Krieger and Gert Vriend. New ways to boost molecular dynamics simulations. *J. Comput. Chem.*, 36(13):996–1007, 2015. ISSN 1096987X. doi: 10.1002/jcc.23899.

Elmar Krieger, Keehyoung Joo, Jinwoo Lee, Jooyoung Lee, Srivatsan Raman, James Thompson, Mike Tyka, David Baker, and Kevin Karplus. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8, 2009. ISSN 08873585.

Sabine Kupzig, Viktor Korolchuk, Ruth Rollason, Anna Sugden, Andrew Wilde, and George Banting. Bst-2/HM1.24 is a raft-associated apical membrane protein with an unusual topology. *Traffic*, 4(10): 694–709, 2003. ISSN 13989219. doi: 10.1034/j.1600-0854.2003.00129.x.

Martin Maechler, Peter Rousseeuw, Anja Struyf, Mia Hubert, and Kurt Hornik. cluster: Cluster Analysis Basics and Extensions, 2018.

Matthew W McNatt, Trinity Zang, and Paul D Bieniasz. Vpu binds directly to tetherin and displaces it from nascent virions. *PLoS Pathog.*, 9(4):e1003299, apr 2013. ISSN 1553-7374. doi: 10.1371/journal. ppat.1003299.

Stuart J.D. Neil. The Antiviral Activities of Tetherin. *Intrinsic Immun.*, 371:67–104, 2013. doi: 10.1007/978-3-642-37765-5.

Stuart J.D. Neil, Trinity Zang, and Paul D. Bieniasz. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature*, 451(7177):425–430, 2008. ISSN 14764687. doi: 10.1038/nature06553.

Kadir A. Ozcan and Christopher E. Berndsen. Bending of the BST-2 coiled-coil during viral budding. *Proteins Struct. Funct. Bioinforma.*, 85(11):1–7, 2017. ISSN 08873585. doi: 10.1002/prot.25362.

C Nick Pace and J Martin Scholtz. A helix propensity scale based on experimental studies of peptides and proteins. *Biophys. J.*, 75(1):422–7, jul 1998. ISSN 0006-3495. URL http://www.ncbi.nlm.nih.gov/pubmed/9649402http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1299714.

David Perez-Caballero, Trinity Zang, Alaleh Ebrahimi, Matthew W McNatt, Devon a Gregory, Marc C Johnson, and Paul D Bieniasz. Tetherin inhibits HIV-1 release by directly tethering virions to cells. *Cell*, 139(3):499–511, oct 2009. ISSN 1097-4172. doi: 10.1016/j.cell.2009.08.039.

Ruth Rollason, Viktor Korolchuk, Clare Hamilton, Peter Schu, and George Banting. Clathrin-mediated endocytosis of a lipid-raft-associated protein is mediated through a dual tyrosine motif. *J Cell Sci*, 120 (Pt 21):3850–8, 2007. ISSN 0021-9533. doi: 10.1242/jcs.003343. URL http://www.ncbi.nlm.nih.gov/pubmed/17940069.

Zev Ross, Hadley Wickham, and David Robinson. Declutter your R workflow with tidy tools. *PeerJ Prepr.*, 2017. ISSN 2167-9843. doi: 10.7287/peerj.preprints.3180v1.

Toshie Sakuma, Akira Sakurai, and Jiro Yasuda. Dimerization of tetherin is not essential for its antiviral activity against Lassa and Marburg viruses. *PLoS One*, 4(9):1–6, 2009. ISSN 19326203. doi: 10.1371/journal.pone.0006934.

Daniel Sauter. Counteraction of the multifunctional restriction factor tetherin. *Front. Microbiol.*, 5(April): 163, jan 2014. ISSN 1664-302X. doi: 10.3389/fmicb.2014.00163.

Daniel Sauter, Dominik Hotter, Susanne Engelhart, Fabian Giehler, Arnd Kieser, Christian Kubisch, and Frank Kirchhoff. A rare missense variant abrogates the signaling activity of tetherin/BST-2 without affecting its effect on virus release. *Retrovirology*, 10(1):1–10, 2013. ISSN 17424690. doi: 10.1186/1742-4690-10-85.

Heidi L Schubert, Qianting Zhai, Virginie Sandrin, Debra M Eckert, Mitla Garcia-Maya, Louise Saul, Wesley I Sundquist, Roberto A Steiner, and Christopher P Hill. Structural and functional studies on the extracellular domain of BST2/tetherin in reduced and oxidized conformations. *Proc. Natl. Acad. Sci. U. S. A.*, 107(42):17951–6, oct 2010. ISSN 1091-6490. doi: 10.1073/pnas.1008206107.

Melissa Swiecki, Suzanne M. Scheaffer, Marc Allaire, Daved H. Fremont, Marco Colonna, and Tom J. Brett. Structural and biophysical analysis of BST-2/tetherin ectodomains reveals an evolutionary conserved design to inhibit virus release. *J. Biol. Chem.*, 286(4):2987–2997, 2011. ISSN 00219258. doi: 10.1074/jbc.M110.190538.

Nanette Van Damme, Daniel Goff, Chris Katsura, Rebecca L. Jorgenson, Richard Mitchell, Marc C. Johnson, Edward B. Stephens, and John Guatelli. The Interferon-Induced Protein BST-2 Restricts HIV-1 Release and Is Downregulated from the Cell Surface by the Viral Vpu Protein. *Cell Host Microbe*, 3(4):245–252, 2008. ISSN 19313128. doi: 10.1016/j.chom.2008.03.001.

Siddarth Venkatesh and Paul D Bieniasz. Mechanism of HIV-1 Virion Entrapment by Tetherin. *PLoS Pathog.*, 9(7):e1003483, jul 2013. ISSN 1553-7374. doi: 10.1371/journal.ppat.1003483.

Abdul A. Waheed, Ariana Gitzen, Maya Swiderski, and Eric O. Freed. High-mannose but not complex-type glycosylation of tetherin is required for restriction of HIV-1 release. *Viruses*, 10(1):1–19, 2018. ISSN 19994915. doi: 10.3390/v10010026.

Sarah Welbourn, Sandra Kao, Kelly E. Du Pont, Amy J. Andrew, Christopher E. Berndsen, and Klaus Strebel. Positioning of cysteine residues within the N-terminal portion of the BST-2/Tetherin ectodomain is important for functional dimerization of BST-2. *J. Biol. Chem.*, 290(6):jbc.M114.617639, 2015. ISSN 0021-9258. doi: 10.1074/jbc.M114.617639.

438 Hadley Wickham. *Elegant graphics for data analysis.* Springer, 2009. ISBN 978-0-387-98140-6. doi:
439   10.1007/978-0-387-98141-3.

440 Claus O Wilke. cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2', 2018a. URL
441   `https://cran.r-project.org/web/packages/cowplot/index.html`.

442 Claus O Wilke. ggridges: Ridgeline Plots in 'ggplot2', 2018b. URL `https://cran.r-project.`
443   `org/web/packages/ggridges/index.html`.

444 Haitao Yang, Jimin Wang, Xiaofei Jia, Matthew W McNatt, Trinity Zang, Baocheng Pan, Wuyi Meng,
445   Hong-Wei Wang, Paul D Bieniasz, and Yong Xiong. Structural insight into the mechanisms of
446   enveloped virus tethering by tetherin. *Proc. Natl. Acad. Sci. U. S. A.*, 107(43):18428–32, 2010. ISSN
447   1091-6490. doi: 10.1073/pnas.1011485107.