Porcine circovirus 2 uses a multitude of weak binding sites to interact with heparan sulfate, and the interactions do not follow the symmetry of the capsid

Sonali Dhindwal[a], Bryant Avila[a], Shanshan Feng[a,b], Reza Khayat[a,b,#]

Running Head: Porcine circovirus 2 cellular recognition

a. Department of Chemistry and Biochemistry, The City College of New York, New York, NY 10031

b. Graduate Program in Biochemistry, The Graduate Center of the City University of New York, New York, NY 10016

Address correspondence to Reza Khayat, rkhayat@ccny.cuny.edu

Abstract

Porcine circovirus 2 is the smallest pathogenic virus capable of autonomous replication within its host. Infections result in immunosuppression and subsequent death of the host, and are initiated via the attachment of the PCV2 icosahedral capsid to heparan sulfate and chondroitin sulfate B glycosaminoglycans on the cell surface. However, the underlying mechanism of structural recognition remains to be explored. Using heparin, a routinely used analog of heparan sulfate, we demonstrate that increasing lengths of heparin exhibit greater affinity towards PCV2. Our competition assays indicate that dextra sulfate (8kDa) has higher affinity than heparin (12kDa), chondroitin sulfate B (41kDa) hyaluronic acid (1.6MDa), and dextran (6kDa) for PCV2. This suggests that polymers high in sulfate content are capable of competing with the PCV2-heparan sulfate interaction, and thus have the potential to inhibit PCV2 infection. Finally, we visualize the interaction between heparin and the PCV2 capsid using cryo-electron microscopy single particle analysis, symmetry expansion, and focused classification. The image reconstructions provide the first example of an asymmetric distribution of heparin on the surface of an icosahedral virus capsid. We demonstrate that each of the 60 capsid subunits that generate the *T*=1 capsid can bind heparin via one of five binding sites. However, not all of the binding sites are occupied by heparin and only one- to two-thirds of the binding sites are occupied. The binding sites are defined by arginine, lysine, and polar amino acids. Mutating the arginine, lysine, and polar amino acids to alanine diminishes the binding capacity of PCV2 to heparin.

Keywords. Circovirus, heparan sulfate, chondroitin sulfate B, glycosaminoglycan, heparin, cellular attachment, electron microscopy, symmetry expansion, focused classification.

39 **Importance:**

40 It has been demonstrated that porcine circovirus 2 (**PCV2**) attaches to cells via heparan sulfate (**HS**) and
41 chondroitin sulfate B (**CSB**) glycosaminoglycans; however, the underlying structural mechanism describing the
42 HS/CSB recognition by PCV2 remains to be explored. We use cryo-electron microscopy with single particle
43 analysis, symmetry expansion, and focused classification to visualize the interaction between the PCV2 capsid
44 and heparin, an analog of heparan sulfate, to better than 3.6Å resolution. We observe that the interaction
45 between the PCV2 and heparin does not adhere to the icosahedral symmetry of the capsid. To the best of our
46 knowledge, this is the first example where the interaction between heparin and an icosahedral capsid does not
47 follow the symmetry elements of the capsid. Our findings also suggest that anionic polymers such as dextran
48 sulfate may act to inhibit PCV2 infection.

49

50 **Introduction**

51 Porcine circovirus (**PCV**) belongs to the *Circovirus* genus of the *Circoviridae* family (1, 2). Members of this family
52 are non-enveloped viruses and possess a circular single stranded DNA (**ssDNA**) genome. Circoviruses are widely
53 distributed in nature and infect terrestrial, avian, and aquatic members of the animal kingdom (3, 4). Three
54 genotypes of PCV have been identified –PCV1, PCV2 and PCV3. PCV1 (1,759nt) was first detected in porcine
55 kidney (PK-15) cell lines, and later found to be a nonpathogenic virus (5, 6). PCV2 (1,767nt to 1,768nt) is
56 morphologically similar to but genetically and antigenically distinct from PCV1, and was isolated from pigs with
57 Postweaning Multisystemic Wasting Syndrome (**PMWS**) (3, 7-10). PMWS, later named to porcine circovirus-
58 associated disease (PCVAD) or porcine circovirus disease (PCVD), culminates in the immunosuppression of the
59 host and death from secondary infection (11-14). Autopsy of infected pigs identifies PCV2 in nearly every tissue,
60 indicating that it has a broad tissue tropism (1, 2, 11, 15). The promiscuous nature of PCV2 is further exhibited
61 by its ability to infect and induce its pathogenic phenotype in rodents and bovine living in the vicinity of infected
62 farms, and BALB/c mice and human cells in the laboratory (3, 4, 16-19). PCV3 (2,000nt) was recently identified
63 and shown to be associated with porcine dermatitis, reproductive failure, and nephropathy syndrome (5, 6, 20).

64 PCV2 is the smallest pathogenic virus capable of replicating in cells without the need for additional viruses (3, 7-
65 10, 14). Its ~1.7knt ambisense genome encodes for a replicase (ORF1) responsible for the rolling circle
66 replication of the genome, a capsid protein (ORF2) responsible for forming the capsid and enclosing the
67 genome, and an ORF3 and ORF4 which may be responsible for causing cellular apoptosis and the pathogenic
68 nature of PCV2 (5, 11-14, 21-24). PCV2 has been shown to initiate cellular infection via attachment to the
69 glycosaminoglycans (**GAG**s) heparan sulfate and chondroitin sulfate B (25). Heparan sulfate (**HS**) and
70 chondroitin sulfate B (**CSB**) are ubiquitously expressed on mammalian cells and act as attachment factors for a
71 variety of macromolecules such as proteases, chemokines, receptors, and pathogens (26, 27). Heparan sulfate is
72 a 30 to 70 kDa linear polysaccharide (40-300 sugar residues and approximately 20-150nm long) composed of
73 alternating sulfated domains (**NS**) and unsulfated domains (**NA**) (26, 28). The NS domains are composed of
74 three to eight repeating disaccharides of L-Iduronic acid (**IdoA**) and D-Glucosamine (**GlcN**) (**Figure S1**). An NS
75 disaccharide can possess two to three sulfates. The NA domains are composed of two to twelve repeating
76 disaccharides of N-Acetyl-D-Glusomaine (**GlcNAc**) and D-Glucoronic acid (**GlcA**) (**Figure S1**) (29). Thus a single

77 chain of HS is composed of multiple NS and NA domains. The NS domain is similar to the heparin structure, and
78 heparin is routinely used as a reagent to study the interaction between the NS domains and their interacting
79 partners (28). Heparin is not a cellular attachment factor, but stored in the granules of mast cells and released
80 into the vascular system as an anti-inflammatory agent (30). CSB, also a linear polysaccharide, is composed of
81 the repeating disaccharide N-Acetyl-Galactosamine (**GalNAc**) and IdoA. A CSB disaccharide has one to two
82 sulfates (27).

83 PCV2 has been shown to internalize via two distinct pathways: clathrin-mediated endocytosis in monocytic (31)
84 and dendritic (1) cells, and caveolae-, clathrin and dynamin-independent small GTPase regulated pathways in
85 epithelial cell lines of porcine kidney (PK15), swine kidney (SK) and swine testicles (ST) (3). The PCV2
86 nucleocapsid has been shown to gain cellular entry by escaping the acidified endosome-lysosome of 3D4/31
87 cells (31). A bipartite nuclear localization signal (**NLS**) at the N-terminus of the capsid protein has been
88 implicated to guide the nucleocapsid into the nucleus for genome replication (32). The PCV2 replicase (ORF1)
89 recruits the cellular machinery to initiate rolling circle replication of the PCV2 genome in the nucleus (33). The
90 newly synthesized PCV2 capsid protein is transported into the nucleus, via its NLS, for genome encapsidation
91 and assembly of infectious virion. Virion then egress from the cell to initiate another cycle of infection (3, 8, 10).

92 The PCV capsids are 20nm in diameter, possess icosahedral symmetry, and are composed of 60-copies (*T=1*) of a
93 capsid protein (34, 35). From here forth we use the terminology "capsid subunit" to refer to the capsid protein
94 in the context of the assembled capsid. The crystal structure of the PCV2 virus like particle (**VLP**) visualized the
95 capsid subunit fold to be that of the canonical viral jelly-roll, first observed for the capsid subunit of tobacco
96 bushy stunt virus (35, 36). The subunit can be described as a β-sandwich, or β-barrel, with a right hand twist
97 (**Figure 1A**). The β-sandwich is composed of two β-sheets. One β-sheet is made of the four antiparallel β-
98 strands BIDG, and the second β-sheet is composed of the four antiparallel β-strands CHEF (**Figure 1A**). The
99 interface between the two β-sheets is the hydrophobic core of the protein. Loops connecting the strands
100 generate the surface topography of the VLP. Strands BC, DE, FG and HI are short and surround the 5-fold vertex
101 of the VLP (**Figure 1B**). Loops connecting strands CD, EF and GH are longer. Loop CD from neighboring subunits

102 contact one another across the icosahedral 2-fold axes of symmetry, and loop GH from neighboring subunits
103 contact one another across the icosahedral 3-fold axes of symmetry (**Figure 1B**).

104 Because the structural mechanism of cellular recognition is not understood, we use combination of cryo-
105 electron microcopy, biochemistry, and site directed mutagenesis to describe this interaction. We use heparin as
106 an analog for studying the interaction between heparan sulfate and PCV2. We ask, does the length of heparin
107 affect its affinity for PCV2, what chemical moieties of GAGs interact with PCV2, where are the heparin binding
108 sites on PCV2, does binding of heparin follow the icosahedral symmetry of the capsid, and does PCV2 undergo a
109 conformational change upon binding of heparin? Our biochemical results demonstrate that longer
110 oligosaccharides of heparin bind with greater affinity to the capsid, and that that the interaction is
111 predominantly driven by the sulfates of the polysaccharide. Our structural results visualize heparin to bind one
112 of five binding sites per capsid subunit, and the interaction to not adhere to the capsid's icosahedral symmetry.
113 Finally, binding of heparin does not induce the Cα backbone to undergo a conformational change. To our
114 knowledge, this is the first example of a non-symmetric distribution of heparin on an icosahedral virus.
115 Furthermore, our studies provide the first structural study of a member of the *Circovirus* genus interacting with
116 the cellular attachment factor of a cell to initiate infection. The knowledge gained in this study can pave the
117 path for developing molecules to interfere with this interaction and inhibit PCV2 infection.

118 **Results**

119 *The interaction between PCV2 and heparin is reversible, dependent on the size of heparin, and primarily dictated*
120 *by sulfates*

121 To study the interaction between heparan sulfate and PCV2 we used an *in vitro* binding assay that involves
122 interacting PCV2 virus like particles (**VLP**) with chromatography sorbent conjugated to 15kDa porcine intestinal
123 mucosa heparin. Heparin is routinely used as an analog of heparan sulfate (**HS**) for studying the interaction
124 between macromolecules and HS (26, 29). This is because the structure of heparin is similar to the NS domain of
125 HS (**Figure S1**). We first determined the concentration of Baculovirus expressed PCV2 VLP (GenBank:
126 ACA51584.1) and time necessary to interact with the sorbent to achieve a robust readout (**Figure 2A**). Two
127 concentrations of VLP were used (370 nM and 92 nM). Maximum binding occurred within 30 minutes for the

128 lower concentration of VLP, and within 3 hours for the higher concentration of VLP. We chose to use the higher
129 concentration of VLP (370 nM) for our subsequent studies because of its improved signal in our binding assays
130 (**Figure S2A).** We then examined if the interaction was reversible by incubating the PCV2 bound sorbent with
131 increasing concentrations of 12kDa porcine intestinal mucosa heparin (**Figure 2B, S2B**). Elution of PCV2 from the
132 sorbent demonstrates that the interaction is reversible.

133 To determine the chemical moieties of GAGs that PCV2 recognizes for cellular attachment, we employed a
134 competition assay to identify the effectiveness of different ligands to compete with the mentioned
135 chromatography sorbent for binding PCV2 (**Figure 2C**). We examined the length of heparin, the pyranose
136 structure, sulfate content, and branched polysaccharide chains. The results demonstrate that while the
137 efficiency of heparin dp6 (6-hexoses) (**Figure 2C, S2C**), dp10 (10-hexoses) (**Figure 2C, S2D**) and dp20 (20-
138 hexoses) (**Figure 2C, S2E**) to inhibit 15kDa heparin from binding to PCV2 increases, they are not nearly as
139 effective as dp36 (36-hexoses, 12kDa) heparin (**Figure 2C, S2F**). A plausible explanation could be that there are
140 multiple weak binding sites on the PCV2 capsid, and that longer oligosaccharides have higher affinity to PCV2
141 because they can bind to multiple sites concurrently. It has been shown that heparin fragments below dp18 are
142 helical, extended and rigid but become more bent with increasing size (37). Thus the dp6 and dp10 may bind to
143 a few binding sites on a PCV2 facet, while the dp20 and dp36 heparin may bend and bind to multiple facets on
144 PCV2.

145 To address if a polysaccharide with high sulfate content can interact with PCV2, we examined the ability of
146 dextran sulfate (8kDa) to inhibit the interaction between PCV2 and the chromatography sorbent (**Figure S2G**).
147 The dextran sulfate (8kDa) used in our studies is a highly branched polysaccharide, with glucose repeating units
148 connected via 1→4 linkages and branching points at 1→6. There are 2.8 sulfates per glucose monomer for this
149 product (personal communication with Alfa Aesar manufacturer). Dextran sulfate is the ligand most efficient at
150 inhibiting the interaction between PCV2 and 15kDa heparin (**Figure 2C**). However dextran, which has the same-
151 branched structure as dextran sulfate, appears to slightly enhance the affinity of the VLP for the sorbent (**Figure**
152 **S2H**). This may be due to dextran acting as a crowding agent (38). The ability of dextran sulfate to inhibit the
153 PCV2-heparin interaction while dextran slightly enhances the interaction suggests that the sulfates are sufficient

154 for binding to PCV2, and that the backbone polysaccharide common to dextran and dextran sulfate does not
155 appear to interact with PCV2.

156 To address if longer GAGs with lower sulfate content exhibit reduced affinity for PCV2, we examined the ability
157 of 1.6MDa hyaluronic acid (**HA**) and 41kDa chondroitin sulfate B (**CSB**) glycosaminoglycans to inhibit the
158 interaction between PCV2 and 15kDa heparin. HA demonstrates no affinity for PCV2 (**Figure 2C, S2I**). HA
159 possesses a single carboxylate on each of its repeating disaccharide units. The hydrogen bonds formed within
160 and between its repeating disaccharides produce stiff helical polysaccharide chains that entangle at very low
161 concentrations (39). Thus, the carboxylate, stiff, and entangled properties of HA may not provide the
162 appropriate interface for PCV2 to bind. The CSB used in our studies is longer and more flexible (40) than the
163 12kDa heparin we used; however, CSB demonstrates a much lower affinity for PCV2 than heparin. This may be
164 because of the lower sulfate content of CSB (an average of 1.1 sulfate per repeating disaccharide) when
165 compared to heparin (an average of 2.4 sulfates per repeating disaccharide) (**Figure 2C, S2J**).

166 Collectively, our data indicate that the PCV2-heparin interaction is reversible, the strength of the interaction
167 increases with longer lengths of heparin, oligosaccharides high in sulfate content bind with greater affinity, long
168 oligosaccharides with low sulfate content bind weakly, and that the interaction between PCV2 and the
169 polysaccharide is predominantly driven by sulfates on the polysaccharide.

170 *A capsid subunit can bind heparin via multiple sites*

171 To identify the binding sites of heparin on PCV2 we determined the cryo-EM image reconstruction of PCV2 in the
172 absence and presence of 12kDa (dp36) porcine intestinal mucosa heparin at pH 7. The purpose of the
173 unliganded image reconstruction was to identify if a conformational change occurred in the subunit upon
174 binding to heparin. We generated icosahedral image reconstruction of the unliganded and heparin-liganded
175 PCV2 to 3.3Å and 2.8Å resolution, respectively. The image reconstructions are of suitable quality to allow us to
176 confidently model the atomic coordinates of PCV2 (**Table 1, Figure 3A**). The refined atomic models for the two
177 image reconstructions overlay with a Cα RMSD of 0.27Å, and an all atom RMSD of 0.98Å (41). These differences
178 are small and can be a result of modeling error at these resolutions (**Figure 3B**) (42, 43). The symmetrized

179 heparin-liganded image reconstruction does not allow us to confidently identify the molecular envelopes for
180 heparin.

181 We postulated that if each subunit possessed multiple binding sites for heparin but only one site could be
182 occupied per subunit then each capsid might be composed of a unique combination of subunits liganded to
183 heparin; thus, icosahedral-averaging strategy could eliminate the molecular envelope of heparin. To test this
184 possibility, we employed the symmetry expand capability of Relion on the alignment parameters from the image
185 reconstruction where icosahedral symmetry (I1) was imposed (44). Briefly, the symmetry expansion protocol
186 generates a copy of each particle for each symmetry operator used during the symmetrized image
187 reconstruction. In our case, 60 equivalent particles were generated for each imaged VLP because 60 matrices
188 describe the I1 symmetry used by Relion. The orientation parameters of these particles are determined by
189 modifying the orientation parameters of the parent particle by the 60 symmetry operators that describe the I1
190 symmetry. Consequently, a single imaged VLP that was assigned an orientation parameter for I1 symmetry is
191 replaced with 60 copies of the VLP assigned orientation parameters for C1 symmetry. These 60 orientation
192 parameters are related to one another via I1 symmetry. The consequence of this strategy is to align the
193 projections of all the subunits from the VLP to one another (45). We then used Frealign to perform focused
194 classification of a subunit (46, 47). Seven of the ten classes exhibit molecular envelopes distinct from the
195 symmetrized image reconstruction (**Figure 4**). To generate a difference map for each class average we
196 subtracted from each class average an image reconstruction generated using the same number of particles as
197 the class average –these particles were randomly selected from the symmetry expanded data set. Calculation of
198 local resolution for each of the ten classes, with MonoRes, suggests that the resolution of the peaks vary from
199 7Å to 15Å (**Figure S3**) (48). The difference peaks were low-pass filtered to 7Å resolution and rendered at 27-
200 sigma (**Figure 4C**). We used 27-sigma as the criteria to remove small and spurious density.

201 The difference peaks may correspond to segments of the 12kDa heparin molecule incubated with PCV2. We
202 note that the amorphous shapes of these peaks indicate that heparin may use multiple binding modes/poses to
203 bind to the capsid. Indeed, heparin is known to use multiple binding modes/poses to interact with proteins (28).
204 We modeled segments of heparin as rigid bodies into the difference peaks to see if the models were chemically

205 reasonable. We then examined if mutation of the basic and polar amino acids that interact with the sulfates in
206 our models diminished the mutant VLP's affinity for heparin. The VLP for these studies were generated using an
207 *in vitro* assembly protocol of capsid protein expressed in *E. coli* –see Materials and Methods.

208 Class 1 (11% of subunits) does not possess a difference peak and thus represents an unliganded subunit. The
209 difference peak for class 2 (9% of subunits) is located near Arg73 and Lys132 of a subunit. A four-hexose
210 segment of heparin can be modeled into this difference peak (**Figure 5A**). Mutation of Arg73 to Ala diminishes
211 the capsid-heparin affinity to 68% of wild type, whereas mutation of Lys132 to Ala abolishes capsid assembly
212 (**Figure 5F**). The difference peak for class 3 (9% of subunits) is located near Lys58 and Arg59, but is too small to
213 accommodate heparin. The difference peak for class 4 (11% of subunits) is near Arg89, and next to the
214 icosahedral 2-fold axis of symmetry. A three-hexose segment of heparin can be modeled into this density
215 (**Figure 5B**). Mutation of Arg89 to Ala diminishes the capsid-heparin affinity to 46% of wild type (**Figure 5F**). The
216 difference peak for class 5 (10% of subunits) is extremely weak. The difference peak for class 6 (9% of subunits)
217 runs between adjacent facets. It is located near Arg186 and Lys227 of the adjacent subunit. A five-hexose
218 segment of heparin can be modeled into this difference peak. A number of hydrogen bonds are also possible
219 between the sulfates of heparin and the side chains of Thr131, Tyr138, and Ser86 (**Figure 5C**). Mutation of
220 Arg186 to Ala diminishes the capsid-heparin affinity to 78% of wild type, whereas mutation of Lys227 to Ala
221 abolishes capsid assembly (**Figure 5F**). The difference peak for class 7 (14% of subunits) is distal to Arg/Lys, but a
222 five-hexose segment of heparin modeled into the difference peak suggests that hydrogen bonds can form
223 between the sulfates of heparin and the side chains of Asn128, Gln188, and the three Tyr173 at the icosahedral
224 3-fold axis (**Figure 5D**). Mutation of Asn128 or Gln188 to Ala diminishes the capsid-heparin affinity to 85% and
225 84% of wild type, respectively (**Figure 5F**). The difference peak for class 8 (12% of subunits) is also very weak.
226 The difference peak for class 9 (7% of subunits) is located near Arg73 and Lys132 of the neighboring subunit, and
227 is equivalent to the peak of class 2. This class has been identified by our procedure because the mask used for
228 classification includes regions of neighboring subunits that are related to one another via the icosahedral
229 symmetry of the capsid. Given that this is a redundant class, we exclude it from further analysis. The difference
230 peak for class 10 (8% of subunits) is located near the icosahedral 5-fold axis, and surrounded by Lys58 and Arg59
231 of one subunit and Lys63 of a neighboring subunit. A sulfate was observed at this location in the crystal

232 structure of PCV2 (35). A two-hexose segment of heparin can be modeled into this density (**Figure 5E**). Mutation
233 of either Lys58 or Arg59 to Ala diminishes PCV2-heparin interaction to 69%, while mutation of Lys63 diminishes
234 the affinity to 38% (**Figure 5F**).

235 To compare the difference peaks on the surface of the PCV2 capsid, we generated a stereographic projection of
236 the PCV2 surface with the program RIVEM (**Figure 5G**) (49). Overlaid onto this image are contour plots of the
237 symmetrized difference peaks, the sulfates observed in the PCV2 VLP crystal structure, and the 532- icosahedral
238 symmetry elements (35). The difference peaks have been symmetrized to visualize all the binding sites available
239 on the surface of the capsid. The sulfates observed in the crystal structure are located near the difference peaks
240 from classes 4 and 10. The stereographic projection demonstrates that the differences peaks from classes 6 and
241 7 partially overlap and share a number of amino acids that include Asn128 and Gln188.

242 The resolution of the image reconstructions from each class is of sufficient quality to refine the atomic models.
243 Refining the atomic model generated from the image reconstruction with imposed icosahedral symmetry into
244 the subunit from each of the seven classes does not identify significant movements in the main chain atoms that
245 could be associated with a conformational change upon binding to heparin. The largest Cα RMSD between
246 subunits from two classes is 0.313Å.

247 *The occupied heparin binding sites do not follow the icosahedral symmetry of the capsid*

248 The result of focused classification is to assign each of the sixty subunits belonging to an imaged VLP to one of
249 ten classes. We used this information along with the inverse of the Relion I1 matrices to position the difference
250 peaks onto the originating VLPs (**Figure 6**). We refer to these models as composite VLP, as they are composed of
251 a symmetrized PCV2 image reconstruction and the positioned difference peaks. The number of difference peaks
252 present per VLP follows a Gaussian distribution: with VLP possessing a minimum of 18 peaks (**Figure 6A, S5**), an
253 average of 31 peaks (**Figure 6B, S5**), and maximum of 43 peaks (**Figure 6C, S5**). The dp36 heparin used to
254 generate this data prompted us to look for occupied binding sites that could accommodate its physicochemical
255 properties (**Figure 6**). Indeed, we can manually model dp24 and dp36 heparin (PDB entries 3IRJ and 3IRL) as
256 rigid bodies into these composite VLPs (**Figure 6A, 6B, 6C**). The bend in the dp36 heparin allows it to bind
257 binding sites distributed across two adjacent facets of the capsid.

258 **Discussion**

259 The cellular attachment factors responsible for initiating infection of porcine monocytic 3D4/31 and porcine
260 kidney epithelial PK-15 cells by PCV2 were demonstrated to be heparan sulfate (**HS**) and chondroitin sulfate B
261 (**CSB**) using competition assays, enzymatic removal of HS and CSB, and mutant CHO cells deficient in HS (*pgs*D-
262 677) or HS and CSB (*pgs*A-745) (25). Both HS and CSB are long linear polysaccharides known as
263 glycosaminoglycans (**GAGs**). HS is composed of two alternating domains: a variably sulfated and rigid NS domain
264 and a non-sulfated and flexible NA domain (26, 29, 50). Three to eight repeating disaccharides generate the NS
265 domain, and two to twelve repeating disaccharides generate the NA domain. The NS domain has a chemical
266 structure similar to heparin with a length of dp6 (6-hexoses) to dp16 (16-hexoses) (**Figure S1)**. The NS domains
267 bind to a plethora of ligands that include, but are not limited to, cytokines, enzymes and their inhibitors,
268 receptors, chemokines, and pathogens (26). A number of non-enveloped icosahedral viruses are known to use
269 HS for cellular attachment. Examples include serotypes of adenovirus (51), serotypes of adeno-associated virus
270 (52), foot-and-mouth disease virus (53), and human papillomavirus (54). Here we note that while heparin is
271 routinely used as an analogue of HS (29), it is not a cellular attachment factor. Heparin is stored in the secretory
272 granules of mast cells and released into the vascular system as an anti-inflammatory agent (30). An extensive
273 review by Xu and Esko describes that a large portion of the binding energy between heparin and proteins can be
274 attributed to the ionic interaction between the Arg/Lys amino acids of proteins and the sulfates of heparin.
275 However, polar amino acids such as Asn, Gln and His can also make hydrogen bonds with the sulfates of heparin,
276 and can contribute 20-70% of the binding energy. Analysis of structures deposited into the PDB indicates that
277 heparin binding sites vary in size (200Å^2 to 360Å^2), number of lysine and arginine residues (0 to 7), and number
278 of polar amino acids (0 to 5). Interestingly, the size of the interface and number of basic residues in the binding
279 site do not predict the affinity of the interaction. Protein-heparin binding strength also vary from 0.3nM to 7µM
280 (K_d). The majority of the heparin binding amino acids are present in the loops of the proteins (28).

281 We have used an *in vitro* competition assay to demonstrate that PCV2 virus like particles (**VLP**) bind to increasing
282 lengths of heparin fragments with greater affinity. The binding trend suggests that the PCV2 VLP may possess
283 multiple weak binding sites such that longer fragments of heparin are able to bind additional sites for producing

284 greater affinity –a phenomenon referred to as avidity (55). Our assays also demonstrate that PCV2 has a greater
285 affinity towards dp36 heparin (12kDa) than the much longer 41kDa CSB. This is in agreement with the cellular
286 competition assays that demonstrate the incubation of PCV2 with heparin reduces infection greater than
287 incubation of PCV2 with CSB (25). Moreover, infection of Chinese hamster ovary (**CHO**) mutant cell lines *pgs*A-
288 677 (lacking HS) was shown to be reduced to ~28% that of wild-type, while infection of the mutant cell lines
289 *pgs*A-745 (lacking HS and CSB) was reduced to ~10% that of wild-type. Suggesting that HS may be the preferred
290 attachment factor on the surface of the cell.

291 Our *in vitro* competition assays also demonstrate that PCV2 VLP exhibit greater affinity towards sulfated
292 polysaccharides. In particular, the observation that PCV2 VLP binds to dextran sulfate (**DS**) 8K but not to dextran
293 6K suggest that PCV2 recognizes and interacts with the sulfates of the polysaccharide. Indeed, our competition
294 experiments demonstrate that DS is the most potent of competitors tested in our assays, suggesting that DS
295 may be used as an inhibitor of PCV2 infection.

296 Our structural data identify six class averages where additional molecular envelopes modify the molecular
297 envelope of the capsid subunits. Subtracting the molecular envelope of an unmodified capsid from these class
298 averages generates six difference peaks. The difference peaks are large enough to accommodate heparin
299 oligosaccharides. Heparin oligosaccharides can be modeled into the difference peaks such that their sulfates are
300 in close proximity to amino acids Lys58, Arg59, Lys63, Arg73, Arg89, Lys132, Asn128, Arg186, Gln188, and
301 Lys227. Single point substitution of these amino acids to Ala diminish the capsid's affinity for heparin, and thus
302 provide biochemical support that the difference peaks represent bound heparin. A single mutation is not
303 sufficient to tremendously diminish binding of heparin; therefore, we anticipate that multiple mutations may be
304 necessary to eliminate the interaction. Comparison of the difference peaks suggests that the peaks belonging to
305 two class averages are equivalent. Thus heparin can bind to one of five binding sites per PCV2 capsid subunit,
306 such that a PCV2 VLP can possess a maximum of 60 sites occupied by heparin. Given that each difference peak
307 originates from a subunit at a specific position in a VLP, the difference peak can be mapped to that position in
308 the VLP to generate a heparin-liganded molecular envelope for each imaged VLP. These molecular envelopes
309 demonstrate that heparin asymmetrically decorates the surface of the capsid rather than the symmetric

310 decoration demonstrated for a number of icosahedral viruses. This may be because prior structural studies of
311 icosahedral viruses in complex with heparin utilized averaging techniques that may have eliminated the
312 asymmetric distribution of heparin on the capsid surface (52, 53, 56-59). The packing of capsids into a crystal for
313 X-ray crystallography inherently imposes the averaging of structures (60). Similarly, cryo-EM image
314 reconstructions with imposed icosahedral symmetry average non-icosahedral entities to symmetric structures
315 with reduced occupancy or amorphous blobs (e.g. the molecular envelope of vial genomes). However, the use
316 of symmetry expansion and focused classification forgo symmetry averaging to provide a more detailed analysis.

317 Our modeling suggests that the reported solution structures of dp24 and dp36 heparin (37) can bind to adjacent
318 facets of the PCV2 capsid. This is primarily because of the bent structure of these oligosaccharides (37).
319 Consequently, it is anticipated that shorter heparin lacking such a bend will be unable to interact with adjacent
320 facets. The longest NS domain of HS is reported to be eight disaccharides (dp16) (28) and thus unlikely to bend
321 and bind to adjacent facets. This creates the likelihood that an NS domain binds to a single facet and the
322 neighboring flexible NA domain allows the HS chain to bend for the subsequent NS domain to bind to an
323 additional facet on the surface of PCV2. Heparan sulfate possesses 40-300 sugar molecules (20-150nm long),
324 and is composed of multiple alternating NS (dp6-dp16) and NA domains (dp4-dp24) (26, 28, 50).

325 In conclusion, we demonstrate that PCV2 capsids bind increasing lengths of heparin with greater affinity, and
326 that this interaction is predominantly driven by the affinity for the sulfates than the pyranose. Our structural
327 studies visualize heparin to bind one of five binding sites per subunit, and the binding of heparin to the capsid to
328 not adhere to the icosahedral symmetry of the capsid. Binding of heparin does not induce a conformational
329 change in the capsid. We also show that longer heparin oligosaccharides (dp24 and dp36) can be modeled to
330 bind adjacent icosahedral facets. Our studies strongly imply that polymers of sufficient length and sulfate
331 content, including heparin 12kDa and dextran sulfate 8kDa, may act as inhibitors of PCV2 infection, and that
332 such inhibitors may act as therapeutics in swine farms where PCV2 infection is present. Future studies include
333 assessing if capsid dynamics are altered upon binding of heparin, if certain binding sites are dispensable for
334 binding to heparin and heparan sulfate, and if the asymmetric distribution of heparin on the capsid surface is
335 important for the life-cycle of the virus.

336

337 **Materials and Methods**

338 *PCV2 VLP expression and purification*

339 The codon sequence for the PCV2 capsid protein (GeneBank ACN59889.1) was optimized by GenScript for
340 Baculovirus expression system. GenScript also generated the Baculovirus stock used in these studies.
341 *Trichoplusia ni* insect cells were propagated in ESF921 medium (Expression System) and infected with the
342 Baculovirus at a multiplicity of infection (MOI) of 0.1. Infected cells were grown for 5 days (50% viability
343 according to Trypan blue staining), and harvested via centrifugation at 3,000xg for 20 minutes.

344 PCV2 VLP was purified from the nucleus of the infected cells. Cells from 500ml of media were gently centrifuged
345 at 800xg for 10 minutes at 4°C. The cell pellet was suspended and lysed in 20mM 4-(2-hydroxyethyl)-1-
346 piperazineethanesulfonic acid (HEPES) pH 7.5, 250mM Sodium Chloride (NaCl), 2mM Magnesium Chloride
347 ($MgCl_2$), 1mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), 1% (v/v) NP-40, and placed on a revolver at
348 4°C for 30 minutes. The lysate was centrifuged for 30 minutes, 720xg, at 4°C to pellet the nuclei. The nuclei
349 pellet was suspended in nuclear lysis buffer: 20mM HEPES pH 7.5, 500mM NaCl, 2mM $MgCl_2$, 0.2mM TCEP, 2mM
350 Deoxycholate, 1µl salt activated nuclease (Sigma-Aldrich, SRE0015-5KU) and lysed using a sonicator. The lysate
351 was centrifuged at 28,000xg for 30 minutes at 4°C. The supernatant was overlaid on a 30% sucrose cushion and
352 centrifuged at 184,000xg, 11°C, 2.5 hours using a 45Ti Beckman Coulter rotor. The pellet was suspended in 2ml
353 of nucleus lysis buffer and processed on a discontinuous 20 to 70% sucrose gradient. The gradient was
354 centrifuged using a SW40 Beckman Coulter rotor at 160,000xg, 4°C, and 18 hours. The fraction containing the
355 VLP was visually identified using light scattering and SDS-PAGE. The fraction containing VLP was dialyzed
356 overnight in 20mM HEPES pH 7.0, 250mM NaCl, 0.2mM TCEP, and 2mM ethylenediaminetetraacetate (EDTA).
357 Samples were concentrated using a 100kDa MWCO ultrafiltration device (Pall Corporation), and stored at 4°C.
358 The sample concentration was estimated to be 370nM using SDS-PAGE with Coomassie staining against a
359 standard of similar molecular weight with a known concentration.

360 *Binding studies*

361 Heparin HyperD™ M Affinity Chromatography Sorbent conjugated to 15kDa porcine intestinal mucosa heparin
362 was purchased from Pall Corporation. HyperD™ M was washed and equilibrated with 20mM HEPES pH 7.0,
363 250mM NaCl, 0.1mM TCEP, 2mM EDTA and allowed to settle to the bottom of the tube. Each reaction
364 contained 2µl of settled HyperD™ M and the indicated amount of VLP (see below). Samples were incubated at
365 20°C for the indicated amount of time (see below –kinetic binding assays) on a revolver to allow PCV2 to bind to
366 the sorbent. The HyperD™ M sorbent was gently pelleted at 100xg for 30s, washed two-times using the
367 equilibration buffer, suspended in water, and processed using SDS-PAGE. The PAGE was stained using
368 InstantBlue™ from Expedeon, and band intensities were quantitated using the Li-Cor Odyssey Blot Imager
369 equipped with Image Studio™ Software version 5.0. The reported Fraction Bound is the quotient of the
370 intensity for each band (**Figure S2**) and the intensity of the band for the lowest concentration of polysaccharide
371 tested. The washes did not remove VLP from the sorbent, as exhibited by SDS-PAGE (data not shown).

372 Kinetic binding assays: 4µl of VLP at concentrations of 370nM and 92.5nM were incubated with 2µl of HyperD™
373 M for 5 seconds (**s**), 10s, 30s, 1 minute (**m**), 5m, 10m, 20m, 1 hour (**h**), 3h and 24h on a revolver at 20°C. The
374 sorbent was processed as described above.

375 Elution assays: VLP bound to HyperD™ M sorbent (as described above) was incubated with increasing
376 concentrations of porcine intestinal mucosa heparin (see below –competition assays) for 3h at 20°C. The bound
377 VLP was quantitated as described above.

378 Competition assays: PCV2 was incubated with the GAG reagents at 20°C for 3h prior to incubation with 2µl of
379 settled HyperD™ M sorbent for 3h. The sorbent was processed as described above for SDS-PAGE.

380 *Glycosaminoglycans and polysaccharide source*

381 Heparin oligosaccharides (HO06, HO10, HO20), and chondroitin sulfate B (DS01) were purchased from Galen
382 laboratories, and dissolved in the VLP dialysis buffer. Porcine intestinal mucosa heparin (average molecular
383 weight 12kDa) was purchased from JT Baker, hyaluronic acid and dextran sulfate 8kDa from Alfa Aesar, and
384 Dextran 6000 (Sigma-Aldrich). All reagents were used within 48 hours after being dissolved in the VLP dialysis
385 buffer. Concentrations were calculated based on the disaccharide molecular weight of 659 g/mol for heparin,

386 551 g/mol for chondroitin sulfate B, 389 g/mol for Hyaluronic acid, and 728 g/mol for dextran sulfate 8kDa.
387 Samples were stored at 4°C.

388 *Electron microscopy*

389 Ultrathin carbon film on Lacey Carbon support film (TedPella, 01824) were used for cryo-electron microscopy.
390 Grids were glow discharged using a PELCO easiGlow™ for 30 seconds. Unliganded and heparin liganded VLP
391 (4µl) were adsorbed to the grid for 15 seconds, blotted with a force of 2 for 6 seconds, and plunge frozen into
392 liquid ethane using an FEI Mark IV vitrobot. Data for the unliganded PCV2 was collected using a side entry FEI
393 Titan Halo with the Leginon software (61). Images were collected at a calibrated magnification of 105,000x on a
394 Gatan K2 camera operating in counting mode. Cryo-EM grids for the PCV2 liganded to heparin were prepared
395 by incubating 20µl of the VLP with 7µl of 60mg/ml 12 kDa heparin for 3 hours. The mixture was then incubated
396 with 2µl of HyperD™ M sorbent for 3 hours at 20°C on a revolver to remove VLP that was not bound to heparin.
397 Cryo-EM grids were prepared as described for the unliganded PCV2. Data for the heparin liganded PCV2 was
398 collected on a Titan Krios with the Leginon software. Images were collected at a calibrated magnification of
399 105,000x on a Gatan K2 camera that follows a GIF Quantum Energy Filter with a slit width of 15eV. Data
400 collection parameters for both data sets are reported in **Table 1**.

401 *Image reconstruction*

402 Motion correction was performed using the MotionCor2 package (62). The first frame of each movie was
403 discarded during the alignment. With the exception of a patch 5 option, the default options were used for the
404 alignment and dose weighting. Particles were automatically detected using Gautomatch v0.53 and per-particle
405 contrast transfer function (**CTF**) estimation was performed on the aligned micrographs using Gctf v0.50 (63).
406 Particles were extracted from the dose-weighted micrographs using the coordinates identified by Gautomatch.
407 Particles from the unliganded dataset were extracted into boxes of 256x256 pixels, and particles from the
408 heparin-liganded data set were extracted into boxes of 300x300 pixels. Three iterations of reference free 2D
409 classification with a diameter of 250Å, 128 classes, and default Relion 2.0 parameters were used to discard bad
410 particles between each iteration (44). The PCV2 crystal structure (PDB entry 3R0R) was used to generate a 60Å
411 resolution initial model using the *molmap* function of UCSF Chimera (64). 3D classification was carried out with

412 the default Relion 2.0 parameters using C1 symmetry. A total of 39,048 particles for the unliganded dataset and
413 a total of 118,853 particles for the heparin-liganded dataset were classified using 3D classification. A total of 4
414 classes were requested for the unliganded dataset and a total of 10 classes were requested for the liganded
415 dataset. From each dataset, particles from the 3D class with the highest resolution image reconstruction were
416 selected for further full resolution refinement using I1 symmetry. The default parameters of Refine3D in Relion
417 2.0 were used to achieve the highest resolution possible. For each data set, binary masks for postprocessing
418 were generated using the *relion_mask_create* program of Relion 2.0 using the following options: 1) a low pass
419 filtered final image reconstruction at 15Å resolution, 2) a threshold value just high enough to eliminate noise
420 exterior to the PCV2 capsid, 3) mask extension by 7 pixels, and 4) additional 2 soft edge pixels expansion. Masks
421 were manually inspected with UCSF Chimera to ensure that no cavities existed in the interior of the masks.
422 Postprocessing was performed with the default parameters of Relion 2.0 (44).

423 The particle orientation parameters attained from the full resolution refinement described above were
424 expanded from I1 to C1 symmetry using the *relion_particle_symmetry_expand* program of Relion 2.0 for 13,600
425 heparin-liganded particles. The consequence of this is to generate 60x13,600 (816,000) subunits that have been
426 aligned to one another. The molecular envelope for a PCV2 subunit that was refined (see below) into the I1
427 image reconstruction was calculated to 25Å resolution using the *molmap* function of UCSF Chimera, threshold
428 masked, extended by 7 pixels, and extended by 2 soft edge pixels using the *relion_mask_create* program of
429 Relion 2.0 (44, 64). The mask was manually moved radially outward to include area peripheral to the capsid in
430 case heparin may be present on the surface of the capsid. The alignment parameters and images were
431 converted to a format for Frealign using the *relion_preprocess* function of Relion and the conversion script
432 provided by the Frealign authors. Twenty rounds of masked classification (10 classes), no alignment, were
433 performed with Frealign (46, 65). Difference maps were calculated using the *vop* function of UCSF Chimera (64).
434 The image reconstruction from each class was subtracted from an image reconstruction generated using the
435 same number of particles randomly selected from the data set. This was done so as both image reconstructions
436 have comparable signal-to-noise ratio. The image reconstructions from the randomly selected particles were
437 visually indistinguishable from the icosahedrally averaged image reconstruction.

438 The image reconstructions were rendered at 100% mass content of the capsid shell (60x27,853Da, or 1.67 MDa).
439 The threshold level to render the images was calculated using the *volume* command from EMAN with the pixel
440 size of the image reconstruction (66).

441 *Structure refinement*

442 The reported crystal structure of PCV2 (35) was manually docked into the symmetrized image reconstructions
443 using UCSF Chimera. The resulting coordinates were iteratively refined using *phenix.real_space_refine* from the
444 Phenix software package with non crystallographic symmetry (**NCS**) constraints applied, and manual fitting with
445 Coot (67, 68). Heparin molecules (PDB entries 3IRI, 3IRJ, 3IRK, and 3IRL) were treated as rigid bodies and
446 manually fitted into the difference peaks of the five classes described above. The capsid subunits pertaining to
447 these molecular envelopes were further refined using manual fitting with Coot and *phenix.real_space_refine* –
448 without NCS.

449 *Generating Composite VLP, representative models of individual VLP*

450 The focused classification identifies one of ten class averages that represents each subunit of a symmetry
451 expanded VLP, and the difference peaks represent each of these ten classes. Consequently, the class averages
452 can be combined with the symmetry elements used in the symmetry expansion process to generate a model of
453 each imaged VLP. The icosahedral image reconstructions were generated using the I1 symmetry of Relion, and
454 symmetry expansion was performed using the I1 symmetry matrices. Consequently, inverse matrices for the
455 Relion I1 matrices should position the class averages, or their corresponding difference peaks, onto the VLP from
456 where they originate. Inverse matrices were calculated for the I1 symmetry matrices using MATLAB R2018a
457 (MathWorks). These symmetry operators were then applied to difference peaks representing each subunit of a
458 VLP with UCSF Chimera (64).

459 *Mutagenesis and binding efficiency*

460 Single point mutants of the PCV2 capsid were made to the consensus sequence construct described in Khayat et
461 al. 2011 (35). Briefly, this construct possesses a hexahistidine and a thrombin cleavage site at the N-terminus in
462 lieu of the forty-two amino acids in the GeneBank ACN59889.1. The remaining amino acid sequence is identical.

463 Mutations were made using the New England Biolabs Inc. site directed mutagenesis kit. Sequences were
464 confirmed using the services of Genewiz. Proteins were purified as previously described (35). Virus like particles
465 (**VLP**s) were assembled by mixing equal volumes of the capsid protein at 37μM concentration with 12%
466 polyethylene glycol 3350, 5% isopropanol, and 0.6M Ammonium Citrate pH 5.0 at 4°C overnight. VLPs form
467 crystals in this solution. The crystals were pelleted via centrifugation at 14,000xg for 10minutes at 4°C. The
468 pellet was suspended in 0.3M HEPPS pH 9.0, 0.6M NaCl, and 5mM β-mercaptoethanol (β-Me) and processed
469 through a HiPrep 16/60 Sephacryl S-500 HR (GE Healthcare Life Sciences) equilibrated with 20mM HEPPS pH 9.0,
470 0.6M NaCl, and 5mM β-Me. The eluted peak was concentrated using a 100kDa MWCO ultrafiltration device
471 (Pall Corporation) and assessed with negative stained electron microscopy. Briefly, 3μl of the sample was
472 applied to a glow discharged carbon coated 400 mesh copper grid (TedPella, 01814-F) and stained with 2%
473 Uranyl Formate. Images were collected using a JEOL JEM-1230 microscope operating at 120kV and equipped
474 with a Gatan US-4000 charge coupled device.

475 Binding studies were performed as described above. The interaction between the sorbent and 370nM
476 concentration of VLP were measured five times to generate error bars. Statistical analysis was performed using
477 MATLAB 2018a to generate the p-values via the Wilcoxon rank sum test (MathWorks).

478

479 **Acknowledgement**

480 SD performed electron microscopy experiments and helped write the manuscript, BA performed electron
481 microscopy experiments, SF performed sample preparation, RK designed and performed experiments, analyzed
482 data, and wrote manuscript. Funds responsible for supporting these studies were provided by NIH National
483 Institute of General medical Sciences and National institute of Allergy and Infectious Diseases (5SC1AI114843)
484 and by Grant Number 5G12MD007603-30 from the National Institute on Minority Health and Health
485 Disparities. Data collection for the PCV2 heparin liganded sample was performed at the Simons Electron
486 Microscopy Center and National Resource for Automated Molecular Microscopy located at the New York
487 Structural Biology Center, supported by grants from the Simons Foundation (349247), NYSTAR, and the NIH

488 National Institute of General Medical Sciences (GM103310). Data collection for the unliganded PCV2 sample
489 was performed at the Imaging Facility of City University of New York (**CUNY**) Advanced Science Research Center
490 (**ASRC**). We would like to acknowledge the scientific and technical assistance from the NYSBC and CUNY ASRC.

491 The image reconstructions and atomic coordinates have been deposited into the EMDB and PDB respectively.
492 Accession codes (EMDB/PDB) are: Unliganded image reconstruction (8939/6DZU); symmetrized liganded image
493 reconstruction (8969/6E2R); Class 02 (8970/6E2X); Class 04 (8971/6E2Z); Class 06 (8972/6E30); Class 07
494 (8973/6E32); Class 09 (8974/6E34); Class 10 (8975/6E39).

495

496 **Figures and Tables**

497 **Table 1: Statistics of cryo-EM image reconstructions and refined atomic coordinates.**

498 **Figure 1. Cartoon depiction of the PCV2 subunit and VLP atomic coordinates.** A) Ribbon diagram of the capsid
499 subunit. The β-strands (BCDEFGHI) are labeled according to the sequence in which they occur. The N- and C-
500 termini are shown as a blue and red sphere, respectively. The Cα atoms for basic and polar residues identified
501 as heparin binding residues in this study are depicted as cyan spheres. B) Ribbon diagram of the VLP showing
502 the molecular organization of the subunits. A single subunit is shown in a dark grey color. Images produced
503 with UCSF Chimera (64).

504 **Figure 2. Interaction between PCV2 and heparin conjugated chromatography sorbent.** A) Binding kinetics
505 between PCV2 VLP and the chromatography sorbent: 370nM (black) and 92nM (red) concentration of VLP. B)
506 Elution of PCV2 from the chromatography sorbent in the presence of increasing concentration of heparin. C)
507 Competition assays testing the ability of various GAGs and polysaccharides to displace PCV2 from the
508 chromatography sorbent: (red) dp6 heparin (6-hexose), (blue) dp10 heparin (10-hexose), (dark grey) dp20
509 heparin (20-hexose), (black) dp36 heparin (36-hexose), (magenta) dextran sulfate 8kDa, (brown) dextran sulfate
510 6K, (light grey) chondroitin sulfate B 41kDa, and (green) hyaluronic acid. Error bars represent the standard
511 deviation of three measurements. Images were made with Origin 2016 (OriginLab).

512 **Figure 3. Symmetrized cryo-EM image reconstructions of unliganded and heparin liganded PCV2.** A) Full atom
513 models of an unliganded (left) and heparin liganded (right) PCV2 subunits shown as stick models. Overlaid onto
514 the stick models are mesh representations of the corresponding 3.3Å (left) and 2.8Å (right) symmetrized image
515 reconstructions. The unliganded image reconstruction is rendered at 1.5-sigma, and the liganded image
516 reconstruction is rendered at 2.6-sigma. The rendering thresholds correspond to 100% mass content –see
517 methods and materials section. B) Cross eye stereo representation comparing the atomic models derived from
518 the unliganded (light grey) and liganded (dark grey) image reconstructions. The backbone atoms are depicted as
519 a ribbon diagram, and the side chains are shown as stick models. Blue and red spheres represent the N- and C-
520 termini, respectively. Images generated with UCSF Chimera (64).

521 **Figure 4. Focused classification of the PCV2 VLP liganded to heparin image reconstruction.** Symmetry
522 expansion of 13,600 particles generated 816,000 capsid subunits for analysis. The image reconstructions are
523 shown at 2.6-sigma. A) The molecular envelope for a subunit is shown in green. The mask used for the
524 classification is shown as a transparent cyan blob. B) Ten classes were requested. The number of particles and
525 resolution reported by Frealign is written below each class. C) Difference peaks for each class average shown in
526 B. There is no difference peak for Class 1. The blue patches represent the molecular envelope of Arg/Lys side
527 chains. Class numbers referred to in the text are written at the bottom. Images made with UCSF ChimeraX (69).

528 **Figure 5. PCV2 heparin interaction.** The PCV2 capsid is shown as a cyan ribbon cartoon. The difference peak
529 (at 27-sigma) is shown as a grey transparent blob –low pass filtered to 7Å. A modeled heparin is shown as a grey
530 stick model. The dashed lines represent hydrogen bonds and/or salt bridges to the sulfates (red and yellow
531 sticks) of heparin. A) Class 2 binding site. B) Class 4 binding site. C) Class 6 binding site. D) Class 7 binding site.
532 E) Class 10 binding site. Images made with UCSF Chimera (64). F) Bar graph representing the efficiency of
533 370nM mutant PCV2 VLP for binding to heparin, with wild-type (WT) VLP having been normalized to 1. The
534 error bars represent errors from five independent measurements. Image made with Origin 2016 (OriginLab).
535 The p-value was calculated using the *ranksum* (Wilcoxon rank sum test) routine of MATLAB 2018a (MathWorks).
536 G) The surface amino acids of PCV2 are plotted on a stereographic projection and colored from blue (80Å) to red
537 (101Å) according to the distance from the center of the particle –see color bar. Overlaid are contour plots at 27-
538 sigma (black outline) for the six difference peaks. The plots are colored according to Figure 4. Residues with a
539 white outline were mutated to Ala in this study, assembled into capsids, and tested for their ability to bind to
540 heparin. Note that Lys137 and Lys227 have not been colored because they did not assemble into capsids and a
541 binding assay could not be performed. The sulfates observed in the crystal structure of PCV2 are shown as gold
542 spheres (35). The icosahedral 5-fold, 3-fold, and 2-fold symmetry elements are depicted as a black pentagon,
543 triangle, and ellipse respectively. Figure generated with RIVEM (49).

544 **Figure 6. Composite VLPs, where each model represents an imaged VLP-heparin complex.** The coloring
545 scheme used for the difference peaks is identical to Figure 4. Top, the dashed box outlines the occupied binding
546 sites that a heparin oligosaccharide is modeled into (bottom). Bottom, a heparin oligosaccharide (grey stick

547 model) occupying the difference peaks (transparent blobs) identified by the dashed box (top). The binding site
548 residues are shown as grey space filling models and the PCV2 capsid is shown as a cyan ribbon model. A red
549 icosahedral cage distinguishes the facets. A) VLP with 18 binding sites occupied by heparin. A dp36 heparin has
550 been modeled into the outlined sites. B) VLP with 31 binding sites occupied by heparin. A dp24 heparin has
551 been modeled into the outlined sites. C) VLP with 40 binding sites occupied by heparin. A dp36 heparin has
552 been modeled into the outlined sites. Images made with UCSF Chimera (64).

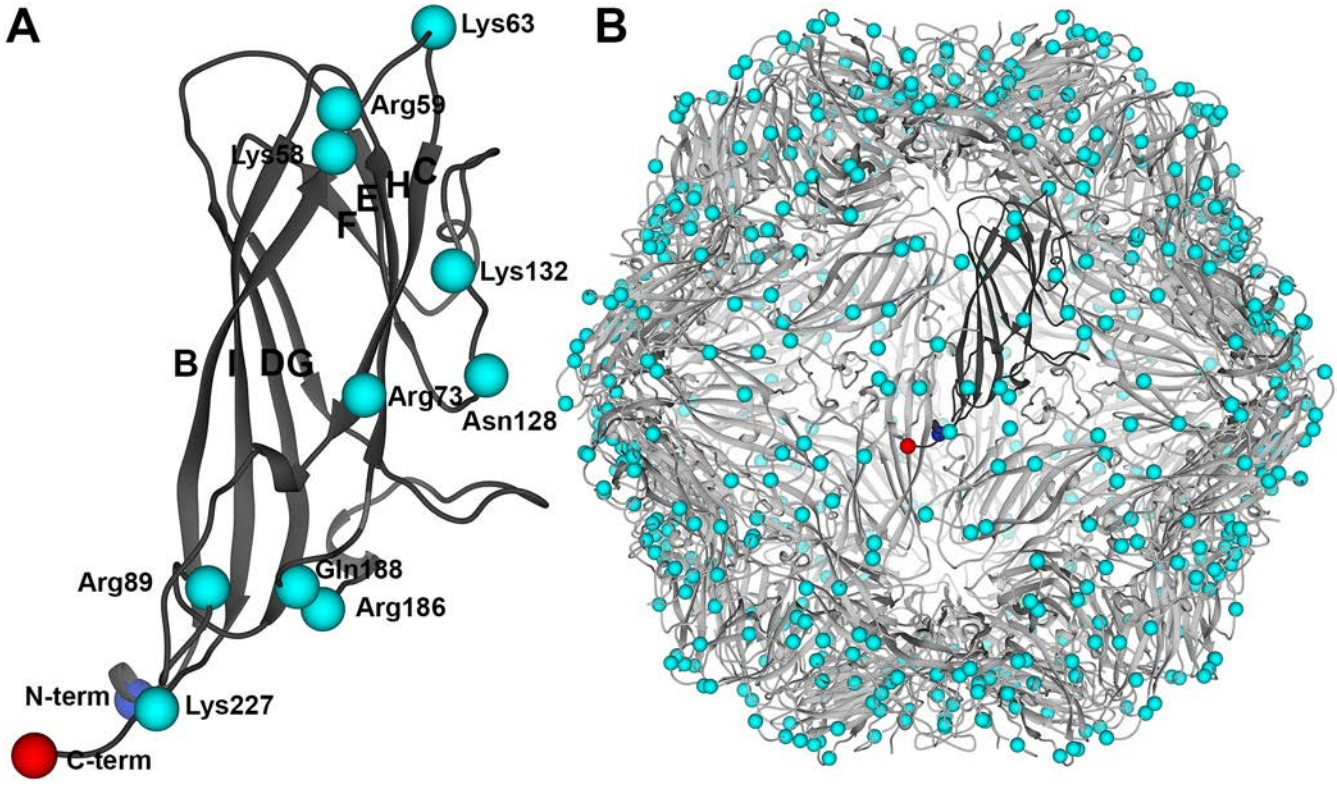
553

554 **Table 1.**

	Unliganded PCV2	PCV2 + Heparin
Image reconstruction		
Number of Micrographs	1,149	3,725
Defocus range	-0.28 to -3.2 μm	-0.25 to -3.6 μm
Total dose	35 e- $Å^{-2}$	32 e- $Å^{-2}$
Dose rate	7.0 e^{-} $Å^{-2}$ sec^{-1}	6.4 e^{-} $Å^{-2}$ sec^{-1}
Number of Frames	50	25
Particles extracted	54,409	166,369
Particles used for reconstruction	28,938	93,725
FSC resolution *	3.3 Å	2.8 Å
B-factor sharpening †	-152.2 $Å^2$	-122.92 $Å^2$
Atomic modeling		
CC	82%	83%
B-factor average	35.5 $Å^2$	38.1 $Å^2$
RMSD ‡	0.012	0.016
RMSD §	1.278	1.271
Molprobity ¶	1.76	1.62
EMRinger	4.89	5.02
Ramachandran favored (%)	76%	87%
Ramachandran allowed (%)	24%	13%
EMDB	8939	8973, 8975, 8974, 8969, 8970, 8971, 8972
PDB	6DZU	6E32, 6E39, 6E34, 6E2R, 6E2X, 6E2Z, 6E30

555 * Fourier shell correlation reported by Relion 2.0 using the gold standard method at a CC of 0.143.
556 † The B-factor sharpening reported by Relion 2.0 during the post-refinement process
557 ‡ The root-mean standard deviation for bonds reported by Phenix.real_space_refine
558 § The root-mean standard deviation for angles reported by Phenix.real_space_refine
559 ¶ Molprobity Overall score
560
561

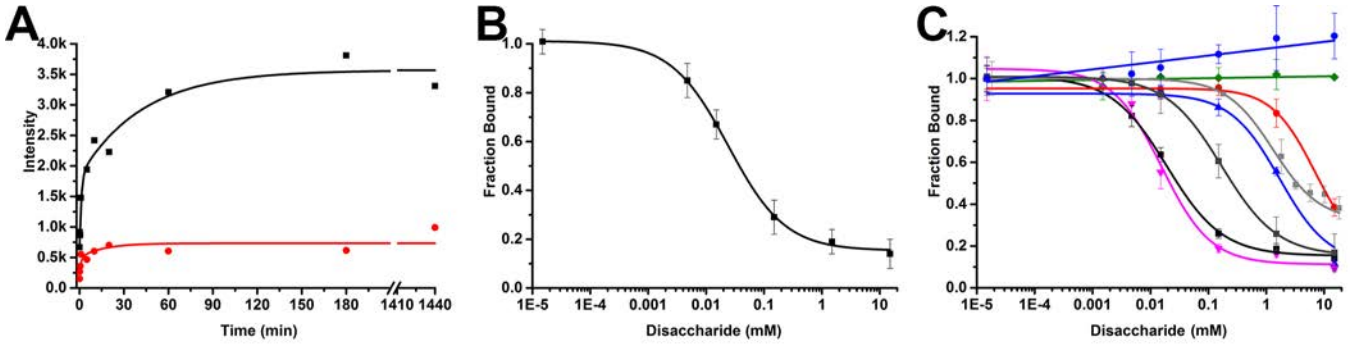
562 **Figure 1.**



563

564

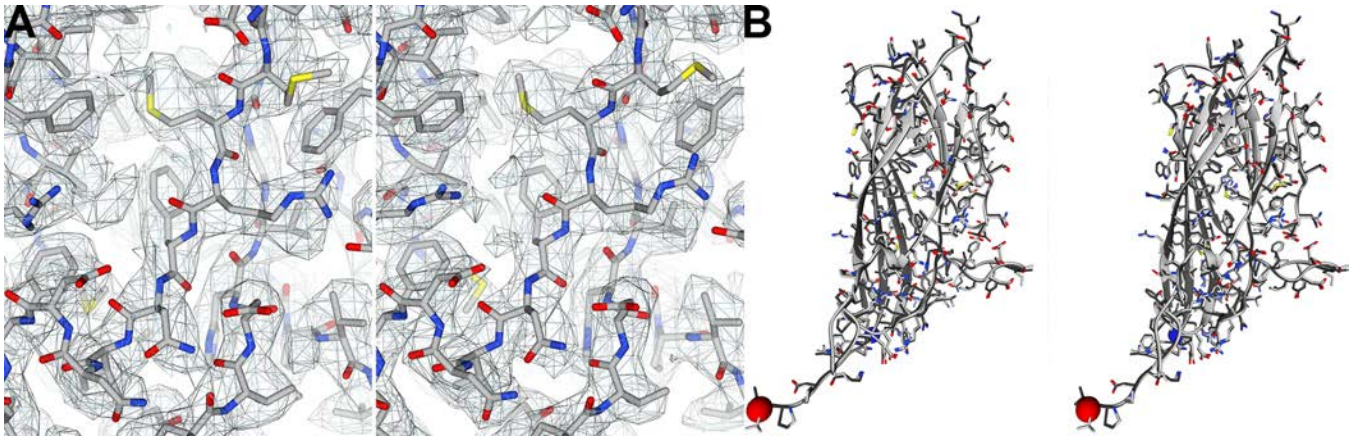
565 **Figure 2.**



566

567

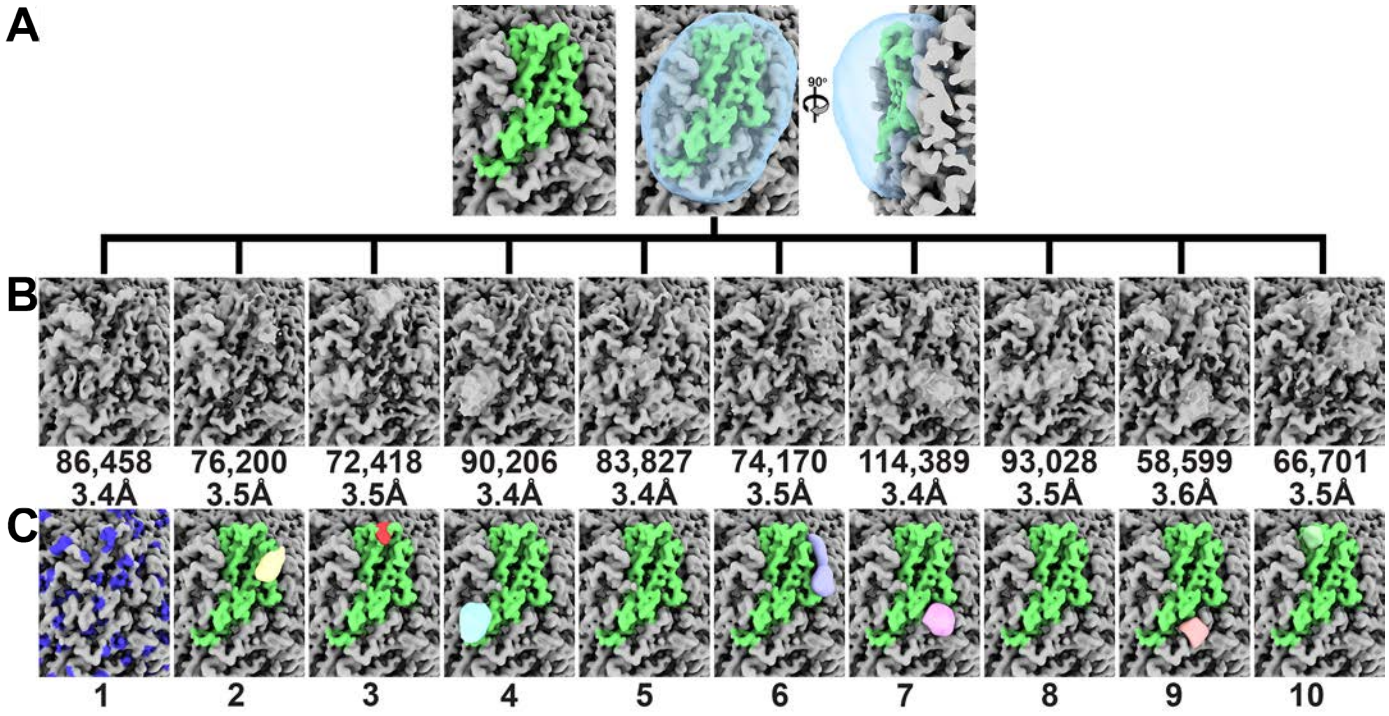
568 **Figure 3.**



569

570

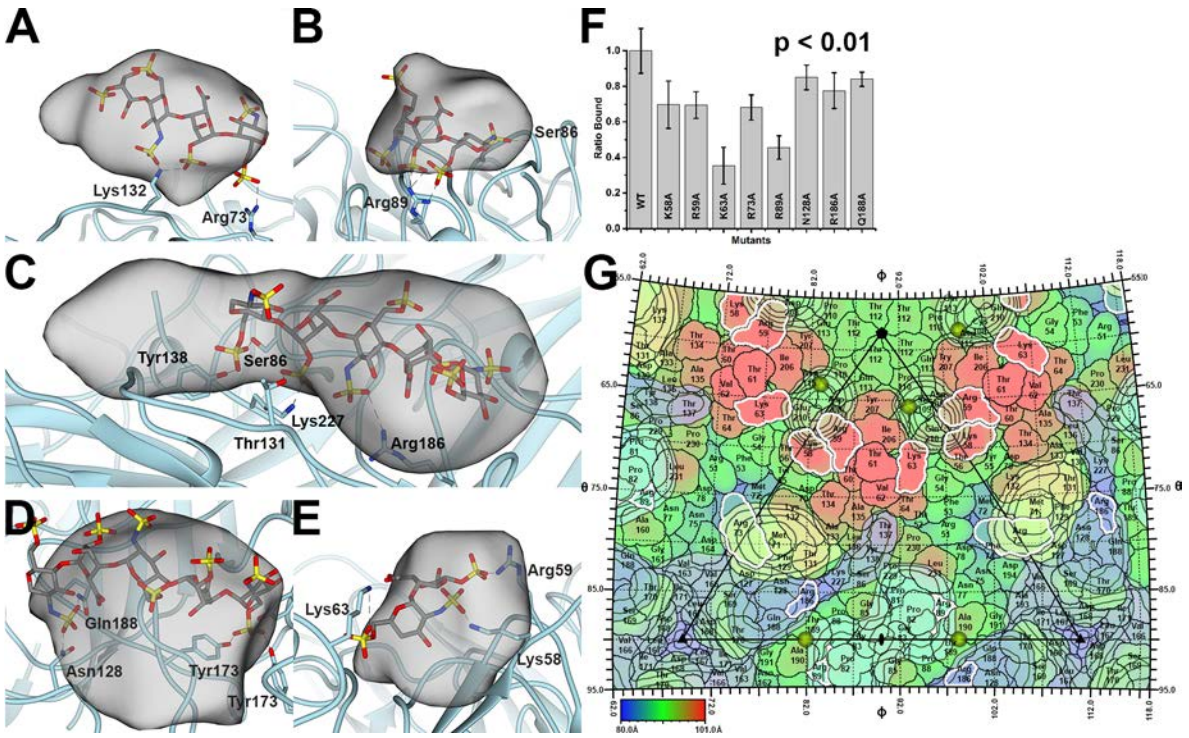
571 **Figure 4.**



572

573

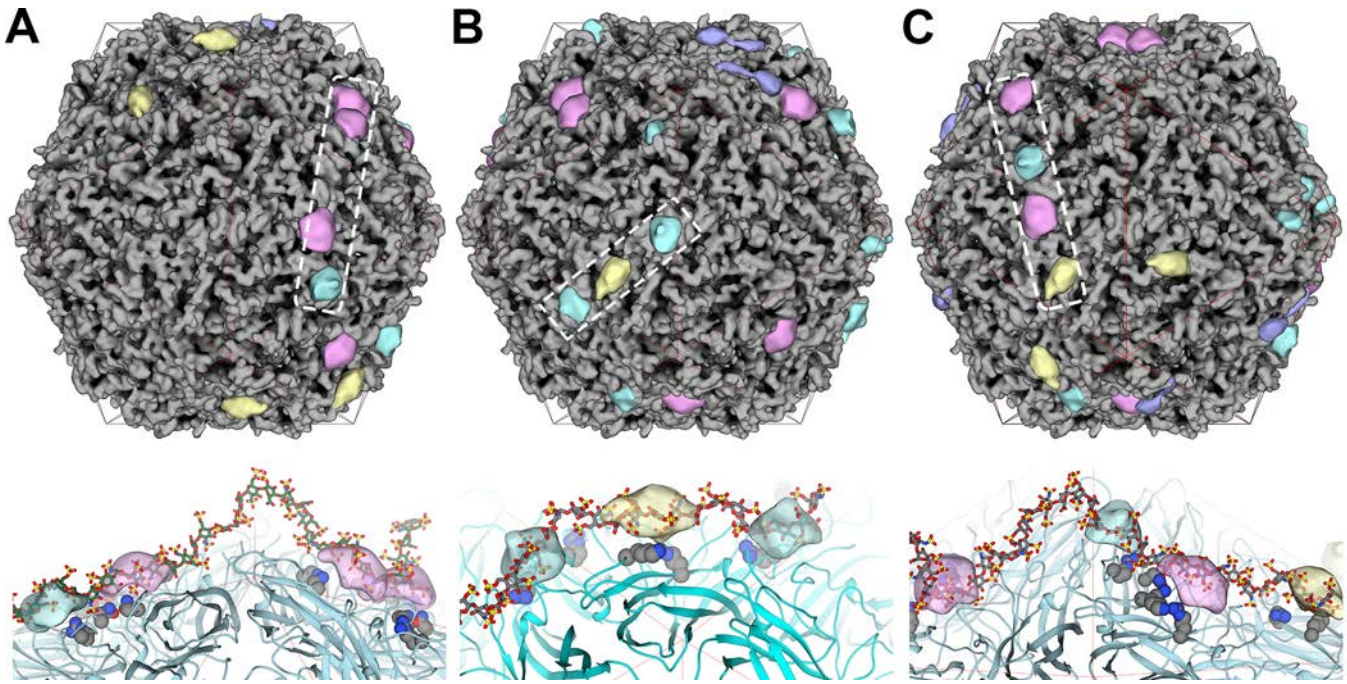
574 **Figure 5.**



575

576

577 **Figure 6.**



578

579

580

581 **References**

582 1. **Vincent IE**, **Carrasco CP**, **Guzylack-Piriou L**, **Herrmann B**, **McNeilly F**, **Allan GM**, **Summerfield A**,
583 **McCullough KC**. 2005. Subset-dependent modulation of dendritic cell activity by circovirus type 2.
584 Immunology **115**:388–398.

585 2. **Todd D**. 2000. Circoviruses: Immunosuppressive threats to avian species: A review. Avian Pathology
586 **29**:373–394.

587 3. **Nauwynck HJ**, **Sanchez R**, **Meerts P**, **Lefebvre DJ**, **Saha D**, **Huang L**, **Misinzo G**. 2012. Cell tropism and
588 entry of porcine circovirus 2. Virus Research **164**:43–45.

589 4. **Lorincz M**, **Csagola A**, **Farkas SL**, **Szekely C**, **Tuboly T**. 2011. First detection and analysis of a fish
590 circovirus. Journal of General Virology **92**:1817–1821.

591 5. **Tischer I**, **Gelderblom H**, **Vettermann W**, **Koch MA**. 1982. A very small porcine virus with circular single-
592 stranded DNA. Nature **295**:64–66.

593 6. **Tischer I**, **Rasch R**, **Tochtermann G**. 1974. Characterization of papovavirus-and picornavirus-like particles
594 in permanent pig kidney cell lines. Zentralblatt fur Bakteriologie, Parasitenkunde, Infektionskrankheiten
595 und Hygiene Erste Abteilung Originale Reihe A: Medizinische Mikrobiologie und Parasitologie **226**:153–
596 167.

597 7. **Ellis J**, **Hassard L**, **Clark E**, **Harding J**, **Allan G**, **Willson P**, **Strokappe J**, **Martin K**, **McNeilly F**, **Meehan B**,
598 **Todd D**, **Haines D**. 1998. Isolation of circovirus from lesions of pigs with postweaning multisystemic
599 wasting syndrome. Can Vet J **39**:44–51.

600 8. **Finsterbusch T**, **Mankertz A**. 2009. Porcine circoviruses--small but powerful. Virus Research **143**:177–183.

601 9. **Morozov I**, **Sirinarumitr T**, **Sorden SD**, **Halbur PG**, **Morgan MK**, **Yoon KJ**, **Paul PS**. 1998. Detection of a
602 novel strain of porcine circovirus in pigs with postweaning multisystemic wasting syndrome. J Clin
603 Microbiol **36**:2535–2541.

604 10. **Rodriguez-Cariñg C**, **Segalés J**. 2009. Ultrastructural Findings in Lymph Nodes from Pigs Suffering from
605 Naturally Occurring Postweaning Multisystemic Wasting Syndrome. Veterinary Pathology **46**:729–735.

606 11. **Opriessnig T**, **Meng X-J**, **Halbur PG**. 2007. Porcine circovirus type 2 associated disease: update on current
607 terminology, clinical manifestations, pathogenesis, diagnosis, and intervention strategies. Journal of
608 Veterinary Diagnostic Investigation **19**:591–615.

609 12. **Segalés J**. 2012. Virus Research. Virus Research **164**:10–19.

610 13. **Mankertz A**. 2012. Virus Research. Virus Research **164**:54–60.

611 14. **Ramamoorthy S**, **Meng X-J**. 2009. Porcine circoviruses: a minuscule yet mammoth paradox. Anim Health
612 Res Rev **10**:1–20.

613 15. **Meng X-J**. 2013. Porcine Circovirus Type 2 (PCV2): Pathogenesis and Interaction with the Immune
614 System. Annu Rev Anim Biosci **1**:43–64.

615 16. **Hattermann K**, **Roedner C**, **Schmitt C**, **Finsterbusch T**, **Steinfeldt T**, **Mankertz A**. 2004. Infection studies
616 on human cell lines with porcine circovirus type 1 and porcine circovirus type 2. Xenotransplantation
617 **11**:284–294.

618 17. **Pinheiro ALBC**, **Bulos LHS**, **Onofre TS**, **de Paula Gabardo M**, **de Carvalho OV**, **Fausto MC**, **Guedes RMC**,
619 **de Almeida MR**, **Júnior AS**. 2013. Research in Veterinary Science. Research in Veterinary Science **94**:764–
620 768.

621 18. **Kiupel M**, **Stevenson GW**, **Choi J**, **Latimer KS**, **Kanitz CL**, **Mittal SK**. 2001. Viral replication and lesions in
622 BALB/c mice experimentally inoculated with porcine circovirus isolated from a pig with postweaning
623 multisystemic wasting disease. Veterinary Pathology **38**:74–82.

624 19. **Halami MY**, **Freick M**, **Shehata AA**, **Müller H**, **Vahlenkamp TW**. 2014. Veterinary Microbiology. Vet
625 Microbiol **173**:125–131.

626 20. **Palinski R**, **Piñeyro P**, **Shang P**, **Yuan F**, **Guo R**, **Fang Y**, **Byers E**, **Hause BM**. 2017. A Novel Porcine
627 Circovirus Distantly Related to Known Circoviruses Is Associated with Porcine Dermatitis and
628 Nephropathy Syndrome and Reproductive Failure. J Virol **91**.

629 21. **Liu J**, **Chen I**, **Kwang J**. 2005. Characterization of a previously unidentified viral protein in porcine
630 circovirus type 2-infected cells and its role in virus-induced apoptosis. J Virol **79**:8262–8274.

631 22. **Liu C**, **Wei Y**, **Zhang C**, **Lu Y**, **Kong X**. 2007. Construction and characterization of porcine circovirus type 2
632 carrying a genetic marker strain. Virus Research **127**:95–99.

633 23. **Chaiyakul M**, **Hsu K**, **Dardari R**, **Marshall F**, **Czub M**. 2010. Cytotoxicity of ORF3 proteins from a
634 nonpathogenic and a pathogenic porcine circovirus. J Virol **84**:11440–11447.

635 24. **Lv Q**, **Guo K**, **Wang T**, **Zhang C**, **Zhang Y**. 2015. Porcine circovirus type 2 ORF4 protein binds heavy chain
636 ferritin. J Biosci **40**:477–485.

637 25. **Misinzo G**, **Delputte PL**, **Meerts P**, **Lefebvre DJ**, **Nauwynck HJ**. 2006. Porcine circovirus 2 uses heparan
638 sulfate and chondroitin sulfate B glycosaminoglycans as receptors for its attachment to host cells. J Virol
639 **80**:3487–3494.

640 26. **Sarrazin S**, **Lamanna WC**, **Esko JD**. 2011. Heparan Sulfate Proteoglycans. Cold Spring Harbor Perspectives
641 in Biology **3**:a004952–a004952.

642 27. **Mizumoto S**, **Yamada S**, **Sugahara K**. 2015. ScienceDirectMolecular interactions between chondroitin.
643 Curr Opin Struct Biol **34**:35–42.

644 28. **Xu D**, **Esko JD**. 2014. Demystifying Heparan Sulfate–Protein Interactions. Annu Rev Biochem **83**:129–157.

645 29. **Rabenstein DL**. 2002. Heparin and heparan sulfate: structure and function. Nat Prod Rep **19**:312–331.

646 30. **Zehnder JL**, **Galli SJ**. 1999. Mast-cell heparin demystified. Nature **400**:714–715.

647 31. **Misinzo G**, **Meerts P**, **Bublot M**, **Mast J**, **Weingartl HM**, **Nauwynck HJ**. 2005. Binding and entry
648 characteristics of porcine circovirus 2 in cells of the porcine monocytic line 3D4/31. J Gen Virol **86**:2057–
649 2068.

650 32. **Shuai J**, **Wei W**, **Li X**, **Chen N**, **Zhang Z**, **Chen X**, **Fang W**. 2007. Genetic characterization of porcine
651 circovirus type 2 (PCV2) from pigs in high-seroprevalence areas in southeastern China. Virus Genes
652 **35**:619–627.

653 33. **Mankertz A**, **Mankertz J**, **Wolf K**, **Buhk HJ**. 1998. Identification of a protein essential for replication of
654 porcine circovirus. J Gen Virol **79 (Pt 2)**:381–384.

655 34. **Crowther RA**, **Berriman JA**, **Curran WL**, **Allan GM**, **Todd D**. 2003. Comparison of the structures of three

656 circoviruses: chicken anemia virus, porcine circovirus type 2, and beak and feather disease virus. J Virol
657 **77**:13036–13041.

658 35. **Khayat R**, **Brunn N**, **Speir JA**, **Hardham JM**, **Ankenbauer RG**, **Schneemann A**, **johnson JE**. 2011. The 2.3-
659 angstrom structure of porcine circovirus 2. J Virol **85**:7856–7862.

660 36. **Harrison SC**, **Olson AJ**, **Schutt CE**, **Winkler FK**, **Bricogne G**. 1978. Tomato bushy stunt virus at 2.9 A
661 resolution. Nature **276**:368–373.

662 37. **Khan S**, **Gor J**, **Mulloy B**, **Perkins SJ**. 2010. Semi-Rigid Solution Structures of Heparin by Constrained X-ray
663 Scattering Modelling: New Insight into Heparin–Protein Complexes. J Mol Biol **395**:504–521.

664 38. **Laurent TC**. 1995. An early look at macromolecular crowding. Biophysical Chemistry **57**:7–14.

665 39. **Necas J**, **Bartosikova L**, **Brauner P**, **Kolar J**. 2008. Hyaluronic acid (hyaluronan): a review. Veterinarni
666 Medicina **53**:397–411.

667 40. **Thelin MA**, **Bartolini B**, **Axelsson J**, **Gustafsson R**, **Tykesson E**, **Pera E**, **Oldberg Å**, **Maccarana M**,
668 **Malmstrom A**. 2013. Biological functions of iduronic acid in chondroitin/dermatan sulfate. FEBS J
669 **280**:2431–2446.

670 41. **Maiti R**, **Van Domselaar GH**, **Zhang H**, **Wishart DS**. 2004. SuperPose: a simple server for sophisticated
671 structural superposition. Nucleic Acids Research **32**:W590–W594.

672 42. **Burra PV**, **Zhang Y**, **Godzik A**, **Stec B**. 2009. Global distribution of conformational states derived from
673 redundant models in the PDB points to non-uniqueness of the protein structure. Proc Natl Acad Sci USA
674 **106**:10505–10510.

675 43. **Kufareva I**, **Abagyan R**. 2012. Methods of Protein Structure Comparison, pp. 231–257. *In* Methods in
676 Molecular Biology. Humana Press, Totowa, NJ.

677 44. **Scheres SHW**. 2012. RELION: implementation of a Bayesian approach to cryo-EM structure
678 determination. J Struct Biol **180**:519–530.

679 45. **Ilca SL**, **Kotecha A**, **Sun X**, **Poranen MM**, **Stuart DI**, **Huiskonen JT**. 1AD. Localized reconstruction of
680 subunits from electron cryomicroscopy images of macromolecular complexes. Nature Communications
681 **6**:1–8.

682 46. **Lyumkis D**, **Brilot AF**, **Theobald DL**, **Grigorieff N**. 2013. Journal of Structural Biology. J Struct Biol
683 **183**:377–388.

684 47. **Grigorieff N**. 2016. Frealign: An Exploratory Tool for Single-Particle Cryo-EMMethods In Enzymology:
685 Computer Methods, Part C, 1st ed. Elsevier Inc.

686 48. **Vilas JL**, **Gómez-Blanco J**, **Conesa P**, **Melero R**, **la Rosa-Trevín de JM**, **Otón J**, **Cuenca J**, **Marabini R**,
687 **Carazo JM**, **Vargas J**, **Sorzano COS**. 2018. MonoRes: Automatic and Accurate Estimation of Local
688 Resolution for Electron Microscopy Maps. Structure/Folding and Design **26**:337–344.e4.

689 49. **Xiao C**, **Rossmann MG**. 2007. Interpretation of electron density with stereographic roadmap projections.
690 J Struct Biol **158**:182–187.

691 50. **Mobli M**, **Nilsson M**, **Almond A**. 2007. The structural plasticity of heparan sulfate NA-domains and hence
692 their role in mediating multivalent interactions is confirmed by high-accuracy 15N-NMR relaxation
693 studies. Glycoconj J **25**:401–414.

694 51. **Dechecchi MC**, **Tamanini A**, **Bonizzato A**, **Cabrini G**. 2000. Heparan Sulfate Glycosaminoglycans Are
695 Involved in Adenovirus Type 5 and 2-Host Cell Interactions. Virology **268**:382–390.

696 52. **O'Donnell J**, **Taylor KA**, **Chapman MS**. 2009. Adeno-associated virus-2 and its primary cellular receptor—
697 Cryo-EM structure of a heparin complex. Virology **385**:434–443.

698 53. **Fry EE**, **Lea SM**, **Jackson T**, **Newman JW**, **Ellard FM**, **Blakemore WE**, **Abu-Ghazaleh R**, **Samuel A**, **King AM**,
699 **Stuart DI**. 1999. The structure and function of a foot-and-mouth disease virus-oligosaccharide receptor
700 complex. EMBO J **18**:543–554.

701 54. **Giroglou T**, **Florin L**, **Schafer F**, **Streeck RE**, **Sapp M**. 2001. Human Papillomavirus Infection Requires Cell
702 Surface Heparan Sulfate. J Virol **75**:1565–1570.

703 55. **Vauquelin G**, **Charlton SJ**. 2013. Exploring avidity: understanding the potential gains in functional affinity
704 and target residence time of bivalent and heterobivalent ligands. British Journal of Pharmacology
705 **168**:1771–1785.

706 56. **Spillmann D**. 2001. Heparan sulfate: anchor for viral intruders? Biochimie **83**:811–817.

707 57. **Xie Q**, **Lerch TF**, **Meyer NL**, **Chapman MS**. 2011. Structure–function analysis of receptor-binding in adeno-
708 associated virus serotype 6 (AAV-6). Virology **420**:10–19.

709 58. **Zhang Z**, **Smith DL**. 1993. Determination of amide hydrogen exchange by mass spectrometry: A new tool
710 for protein structure elucidation - Zhang - 2008 - Protein Science - Wiley Online Library. Protein Science.

711 59. **Guan J**, **Bywaters SM**, **Brendle SA**, **Ashley RE**, **Makhov AM**, **Conway JF**, **Christensen ND**, **Hafenstein S**.
712 2017. Cryoelectron Microscopy Maps of Human Papillomavirus 16 Reveal L2 Densities and Heparin
713 Binding Site. Structure/Folding and Design **25**:253–263.

714 60. **Johson JE**, **Rueckert RR**. 1997. Structural Biology of Viruses. New York.

715 61. **Suloway C**, **Pulokas J**, **Fellmann D**, **Cheng A**, **Guerra F**, **Quispe J**, **Stagg S**, **Potter CS**, **Carragher B**. 2005.
716 Automated molecular microscopy: the new Leginon system. J Struct Biol **151**:41–60.

717 62. **Zheng SQ**, **Palovcak E**, **Armache J-P**, **Verba KA**, **Cheng Y**, **Agard DA**. 2017. MotionCor2: anisotropic
718 correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods **14**:331–332.

719 63. **Zhang K**. 2016. Journal of Structural Biology. J Struct Biol **193**:1–12.

720 64. **Pettersen EF**, **Goddard TD**, **Huang CC**, **Couch GS**, **Greenblatt DM**, **Meng EC**, **Ferrin TE**. 2004. UCSF
721 Chimera--a visualization system for exploratory research and analysis. J Comput Chem **25**:1605–1612.

722 65. **Grigorieff N**. 2007. FREALIGN: High-resolution refinement of single particle structures. J Struct Biol
723 **157**:117–125.

724 66. **Ludtke SJ**, **Baldwin PR**, **Chiu W**. 1999. EMAN: semiautomated software for high-resolution single-particle
725 reconstructions. J Struct Biol **128**:82–97.

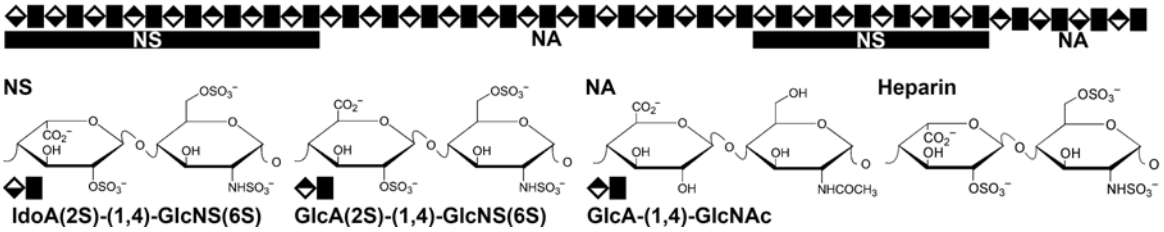
726 67. **Afonine PV**, **Headd JJ**, **Adams TCTPD**. 2013. New tool: phenix.real_space_refine.

727 68. **Emsley P**, **Cowtan K**. 2004. research papers. Acta Cryst (2004) D60, 2126-2132
728 [doi:101107/S0907444904019158] 1–7.

729 69. **Goddard TD**, **Huang CC**, **Meng EC**, **Pettersen EF**, **Couch GS**, **Morris JH**, **Ferrin TE**. 2017. UCSF ChimeraX:
730 Meeting Modern Challenges in Visualization and Analysis. Protein Sci.

733 **Supplemental Information**

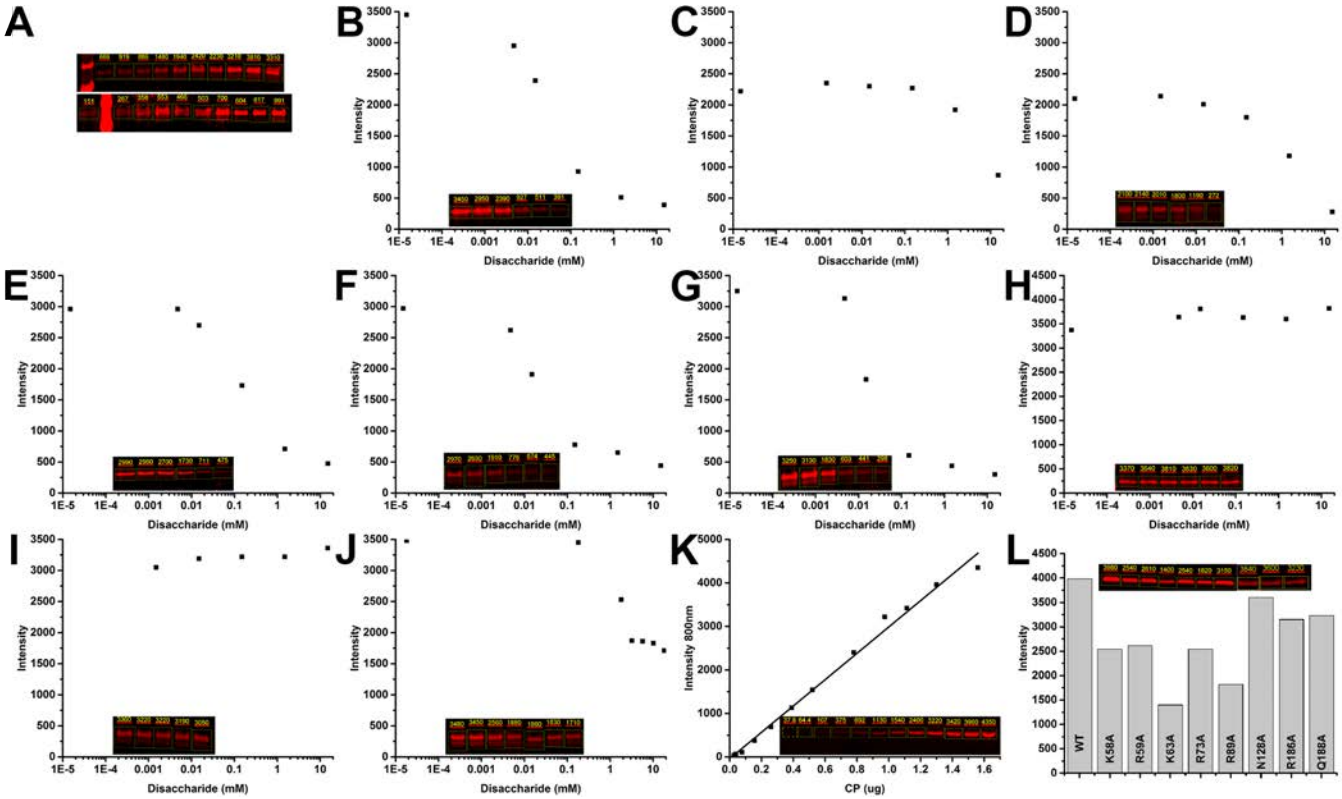
734 **Figure S1. Heparan sulfate and heparin**. Top) Cartoon representation of the NS and NA domains of a heparan
735 sulfate glycosaminoglycan chain. The NS domain is composed of 3-8 repeating disaccharides, and the NA domain
736 is composed of 2-12 repeating disaccharides. Bottom) Chemical structure of the heparan sulfate and heparin
737 repeating disaccharides. The repeating disaccharide of the heparan sulfate NS domain is similar to the repeating
738 disaccharide of heparin.



739

740

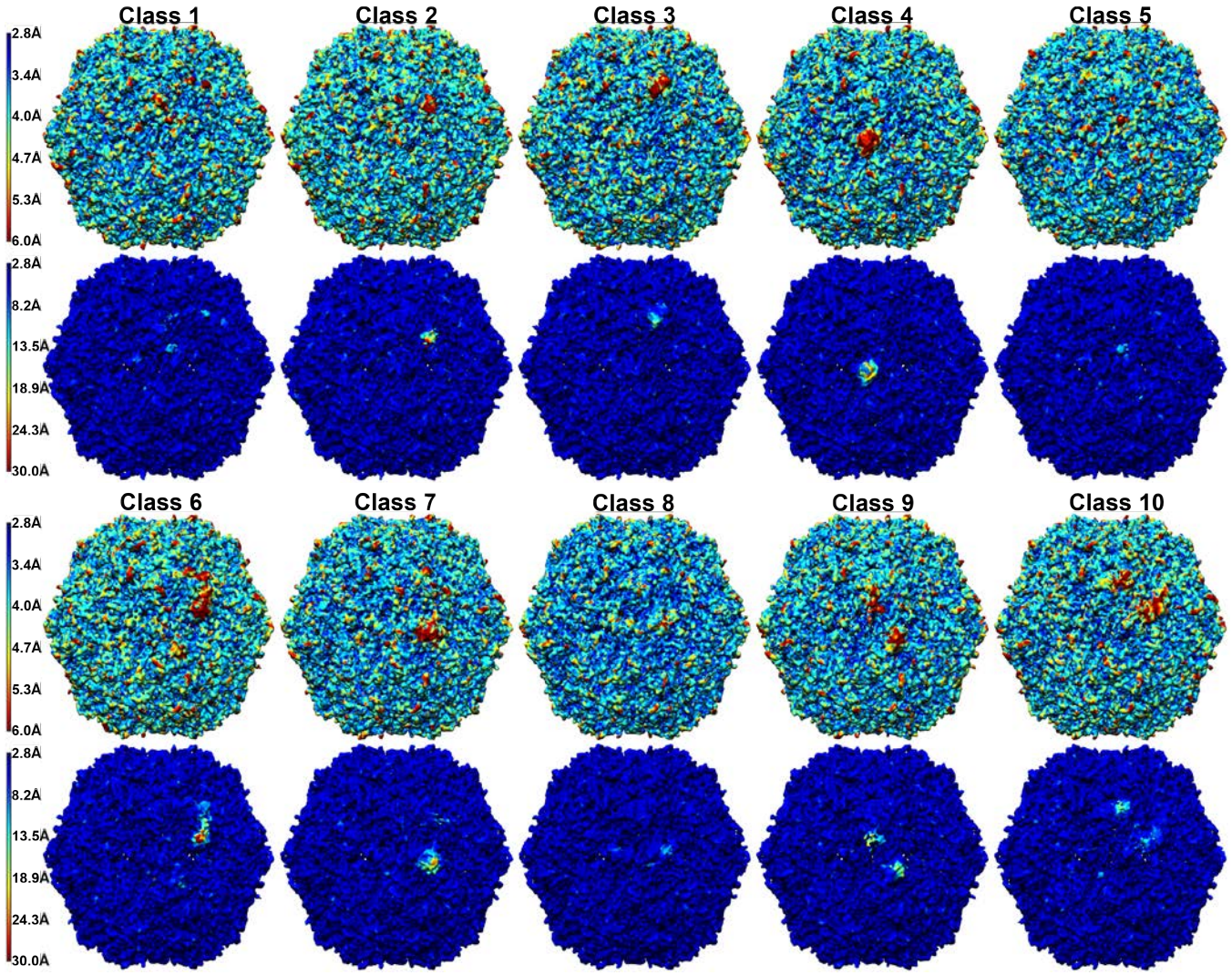
741 **Figure S2. SDS-PAGE densitometry of PCV2 VLP heparin binding interaction.** Images represent the interaction
742 between PCV2 VLP and HyperD™ chromatography sorbent. The contrast of each digitized SDS-PAGE was scaled
743 to enhance the signal for visualization. This modification does not affect the intensity reported by the Image
744 Studio™ software used for the analysis. The numbers above each band represent the intensity corrected by
745 background subtraction within each box. A) Binding kinetics of 370nM and 97nM VLP demonstrating that the
746 signal from 370nM produces a robust readout (Figure 2A). B) Elution of VLP from the sorbent with increasing
747 heparin concentration (Figure 2B). GAG competition assay representatives of: C) 6-hexose, D) 10-hexose, E) 20-
748 hexose, F) 36-hexose heparin, G) dextran sulfate 8kDa, H) dextran 6kDa, I) hyaluronic acid, J) chondroitin sulfate
749 B. K) Linear response curve of the LI-CORR Odyssey CLX to increasing amount of PCV2 capsid protein, indicating
750 that intensity is directly correlated to protein content. L) Binding efficiency of *in vitro* assembled WT, K58A,
751 R59A, K63A, R73A, R89A, N128A, R186A, and Q188A capsids (Figure 5). Inset is: WT K58A, R59A, K63A, R73A,
752 R89A, WT, N128A, R186A, and Q188A capsids. Figures made with Image Studio™ (LI-CORR Biosciences) and
753 Origin 2016 (OriginLab).



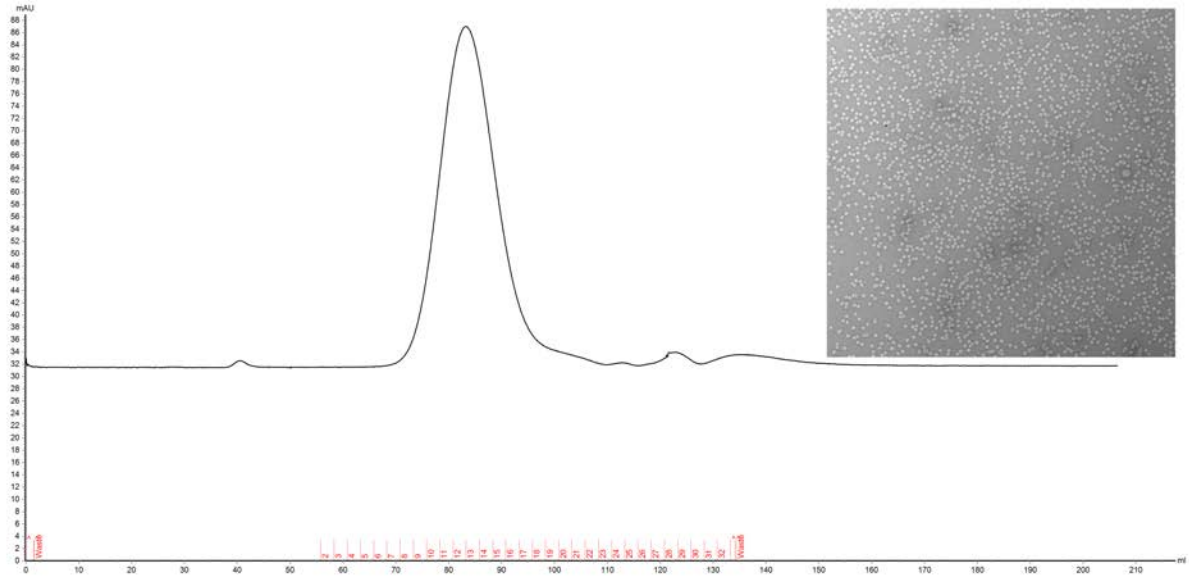
754

755

Figure S3. Local resolution. The local resolution for each of the 10 classes was determined using the program MonoRes. Two plots are shown for each class due to the large range in resolution. The color bar on the left of each row depicts the resolution for the indicated color. A) The top plot was calculated using a resolution range of 2.8Å to 6Å and the bottom plot was calculated using a resolution range of 2.8Å to 30Å for classes 1 to 5. B) Similar to A but for classes 6 to 10 (48). Image generated using UCSF Chimera (64).



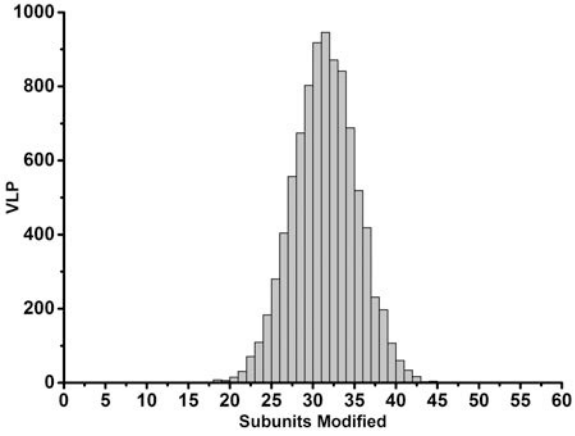
763 **Figure S4. Assembly of mutant VLP.** Size exclusion chromatography profile of an *in vitro* assembled PCV2 VLP.
764 Image made with Unicorn 6.1 (GE Health Care Lifesciences) and a corresponding electron microscope
765 micrograph of the negative stained wild-type VLP.



766

767

768 **Figure S5. Occupied heparin binding sites on PCV2 VLPs.** The distribution follows a Gaussian profile with a
769 minimum of 17, a mode of 31, and a maximum of 44 binding sites occupied on individual VLPs. Image made
770 with Origin 2016 (OriginLab).



771