Cloning, Expression and Biophysical Characterization of a Yeast-expressed Recombinant SARS-CoV-2 Receptor Binding Domain COVID-19 Vaccine Candidate

Wen-Hsiang Chen[ab], Junfei Wei[a], Rakhi Tyagi Kundu[a], Rakesh Adhikari[a], Zhuyun Liu[a], Jungsoon Lee[a], Leroy Versteeg[a], Cristina Poveda[a], Brian Keegan[a], Maria Jose Villar[a], Ana Carolina de Araujo Leao[a], Joanne Altieri Rivera[a] Portia M. Gillespie[a], Jeroen Pollet[ab], Ulrich Strych[ab], Bin Zhan[ab], Peter Hotez[abcd*], Maria Elena Bottazzi[abcd*]

[a] Texas Children's Hospital Center for Vaccine Development, Houston, TX, USA

[b] Departments of Pediatrics and Molecular Virology & Microbiology; National School of Tropical Medicine; Baylor College of Medicine, Houston, TX, USA

[c] Department of Biology, Baylor University, Waco, TX, USA

[d] James A. Baker III Institute for Public Policy, Rice University, Houston, TX, USA

correspondence:

[*]Maria Elena Bottazzi bottazzi@bcm.edu, 1102 Bates St., Ste. 550 | Houston, TX 77030

[*]Peter J Hotez hotez@bcm.edu, 1 Baylor Plaza, Houston TX 77030

ABSTRACT

Coronavirus disease 2019 (COVID-19) caused by SARS-CoV-2 has now spread worldwide to infect approximately 50 million people, with over 1 million reported deaths, and a safe and effective vaccine remains urgently needed. Based on previous experience developing vaccines against SARS and MERS, we constructed three variants of the recombinant receptor-binding domain (RBD) of the SARS-CoV-2 spike (S) protein (residues 331-549) in yeast as follows: (1) a "wild type" RBD (RBD219-WT), (2) a deglycosylated form (RBD219-N1) by deleting the first N-glycosylation site, and (3) a combined deglycosylated and cysteine (C538A-mutated variant (RBD219-N1C1)). We compared the expression yields, biophysical characteristics, and functionality of the proteins produced from these constructs. Collectively, these three recombinant protein RBDs showed similar secondary and tertiary structure thermal stability and had the same affinity for their receptor, angiotensin-converting enzyme 2 (ACE-2), suggesting that the selected deletion or mutations did not cause any significant structural changes or alteration of function. However, RBD219-N1C1 had a higher fermentation yield, was easier to purify, and had a lower tendency to form oligomers when compared to the other two proteins and was therefore selected for further vaccine development and evaluation.

KEYWORDS: coronavirus, protein characterization, *P. pastoris*

1. INTRODUCTION

Betacoronaviruses have caused major disease outbreaks in humans every decade since the beginning of the millennium. In 2002, severe acute respiratory syndrome coronavirus (SARS-CoV) was responsible for 8,098 infections and approximately 800 deaths. Beginning in 2012, the Middle East respiratory syndrome (MERS) caused by MERS-CoV has led to 2,519 infections and 866 associated deaths and the virus is still circulating in camels. First reported in China in December 2019, SARS-CoV-2, the pathogen that causes coronavirus disease 2019 (COVID-19), has now rapidly spread worldwide and infected more than 45 million people and caused approximately 1.2 million deaths up to October 2020 [1].

A coronavirus virion consists of membrane, envelope, nucleocapsid, and spike (S) proteins. Of these, the S proteins are commonly selected as vaccine candidates due to their critical functions in host cell entry [2-7]. However, even though no preclinical or human trials have yet reported that immunopathology could be triggered by the full-length S protein of SARS-CoV-2 after viral challenge [8], previous preclinical studies in mice indicated that the full-length SARS-CoV S protein could induce eosinophilic immune enhancement [9, 10] and lung immunopathology [11, 12], possibly due to epitopes outside of the receptor-binding domain (RBD). The SARS-CoV RBD vaccine was shown to induce high levels of virus-neutralizing antibodies, minimized or abrogated eosinophilic immune enhancement in mice compared to the full-length S protein, and can be easily scaled for production [12, 13]; therefore, some groups consider the RBD an attractive vaccine immunogen alternative to using the full-length S protein [14-19]. Beyond our findings with SARS-CoV, additional studies have already shown that mammalian cell-expressed SARS-CoV-2 RBD triggered high specific total IgG as well as neutralizing antibody titers in mice [7] and that the vaccination of insect cell-expressed RBD protected non-human primates (NHPs) from SARS-CoV-2 challenge *in vivo* with no significant lung histopathological changes, further supporting a rationale for using the RBD as antigen to develop an efficacious vaccine [20].

Based on previous experience in developing the SARS-CoV RBD vaccine [4, 13, 17] and its high amino acid sequence similarity between SARS-CoV and SARS-CoV S proteins [21-23], we identified a 219 amino-acid RBD sequence of SARS-CoV-2 (RBD219-WT), spanning residues 331-549 of the S protein by aligning the sequence of the S protein with the previously identified SARS-CoV RBD219 [4]. We further used the RBD219-WT sequence as a backbone to derive (i) a deglycosylated form (RBD219-N1) with Asn331 of the S protein deleted to prevent hyperglycosylation that hinders scare-up production and purification, and (ii) a deglycosylated and cysteine mutagenized form (RBD219-N1C1), with the Asn331 deletion and the Cys538 changed to an alanine residue to prevent oligomer formation via intermolecular disulfide bridging. The immunogenicity of these vaccine candidates was recently tested, demonstrating

93 that when formulated with Alhydrogel®, these yeast expressed RBDs triggered high levels of neutralizing
94 antibody titers [24].

95 Here, we examined the production yield of three tag-free RBDs in the fermentation supernatant
96 prior to purification. Two of the RBD constructs resulted in high fermentation yields, however, for the
97 RBD219-N1 construct, we only achieved low fermentation yields with likely high host cell protein levels.
98 This impeded us from getting any purified RBD219-N1 protein, hence, a hexahistidine tagged construct
99 (RBD219-N1+His) was created to obtain purified protein for further characterization. Biophysical
100 characteristics, and *in vitro* functionality of these three purified recombinant RBD proteins were
101 evaluated. Based on the results, we identified the RBD219-N1C1 construct best suited for advancement
102 to scaled-up process development, technological transfer for manufacturing, and preclinical evaluation
103 in preparation for the vaccine to advance into clinical development.

104 2. MATERIALS AND METHODS

105 2.1. Cloning and expression of SARS-CoV-2 RBDs in yeast *Pichia pastoris*

106 The DNAs encoding RBD219-WT (residues 331–549 of the SARS-CoV-2 spike protein, GenBank:
107 QHD43416.1), RBD219-N1 (residues 332-549) and RBD219-N1C1 (residues 332-549, C538A) were codon-
108 optimized based on yeast codon usage preference and synthesized by GenScript (Piscataway, NJ, USA),
109 followed by subcloning into the *Pichia* secretory expression vector pPICZαA (Invitrogen) using EcoRI/XbaI
110 restriction sites. A hexahistidine tag version for RBD219-WT and RBD219-N1 (namely, RBD219-WT+His
111 and RBD219-N1+His, respectively) was also generated by adding additional DNA encoding six histidine
112 residues at the C-terminus to facilitate purification as a backup. The recombinant plasmid DNAs were
113 then transformed into *Pichia pastoris* X-33 by electroporation. The expression of the recombinant RBDs
114 was confirmed by induction with 0.5% methanol at 30 °C for 72 h. The seed stock in 20% glycerol of each
115 recombinant construct was then generated as described previously [13].

116 2.2. Fermentation and purification of SARS-CoV-2 RBDs

117 RBD219-WT, RBD219-N1, and RBD219-N1C1 in pPICZαA/*P. pastoris X33* clones were fermented in 5 L
118 vessels as described previously with minor modifications [13]. Briefly, the seed stock of each construct
119 was used to inoculate 0.5 L Buffered Minimal Glycerol (BMG) medium until the OD_{600} reached 10±5.
120 Depending on the OD_{600} of overnight culture, 86-270 mL of the culture was then used to inoculate 2.5 L
121 sterile low salt medium (LS) in the fermenter containing 3.5 mL/L PTM1 trace elements and 3.5 mL/L
122 0.02% d-Biotin to reach to initial OD of 0.5. Fermentation was initiated at 30 °C and pH 5.0, while the gas

123 and agitation were adjusted to maintain dissolved oxygen (DO) at 30%. Upon DO spike, the pH was
124 ramped up to 6.5 using 14% ammonium hydroxide, and the temperature was lowered to 25°C over 1 h
125 and methanol was then pumped in from 0.8 mL/L/h to 11 mL/L/h and the pH was adjusted to 6.0 using
126 14% ammonium hydroxide over 6–8 h. Induction was maintained at 25 °C with minor methanol feed
127 adjustments, as needed, for 70 h. After fermentation, the culture was harvested by centrifugation. The
128 fermentation supernatant (FS) was then evaluated by SDS-PAGE and Western blot.

129 To purify RBD219-WT, the FS was first filtered through a 0.45 µm filter followed by a negative
130 capture step with a Q Sepharose XL (QXL) column to remove some host cell protein. The flow-through
131 from the QXL column was then further purified by a Butyl Sepharose HP column and a Superdex 75 size
132 exclusion column (SEC). Due to the low target protein yield and large amounts of impurities present in
133 RBD219-N1 fermentation, we were unable to successfully purify the tag-free RBD219-N1 by the same
134 approach as RBD219-WT. Instead, we purified the using the hexahistidine tagged version (RBD219-
135 N1+His, where six additional histidine residues were expressed at the C-terminus of RBD219-N1); to
136 purify RBD219-N1+His, HisTrap immobilized metal affinity column was used followed by Superdex 75
137 chromatography. Finally, to purify RBD219-N1C1, the FS was filtered through a 0.45 µm filter before a
138 Butyl Sepharose HP column followed by a Superdex 75 column. The final buffer for these three proteins
139 was TBS (20 mM Tris, 150 mM NaCl, pH 7.5).

140 **2.3. SDS-PAGE and Western Blot**

141 RBD219-WT, RBD219-N1+His, and RBD219-N1C1 were loaded on 4-20% Tris-glycine gels, and stained
142 with Coomassie Blue or transferred to a polyvinylidene difluoride membrane and probed with a
143 monoclonal anti-SARS-CoV-2 Spike rabbit antibody recognizing the RBD region (Sino Biological, Beijing,
144 China; Cat # 40150-R007) to evaluate the size and confirm the identity. These three RBDs were also
145 treated with PNGase-F (New England Biolabs, Ipswich, MA, USA; Cat# P0704S) following the
146 manufacturer's instruction and loaded on to SDS-PAGE gels to evaluate the impact of size caused by
147 glycosylation. Western blotting was also used to evaluate the fermentation yield; in short, serially
148 diluted purified RBD protein corresponding to the construct in the fermentation run was loaded on the
149 Tris-glycine gels with a fixed volume of undiluted fermentation supernatant of different RBD constructs.
150 A log-log plot of RBD intensity versus the known amount of loaded RBD was graphed and the linear
151 regression was calculated from the plot.

152 **2.4. Size and Protein Aggregation Assessment by Dynamic Light Scattering**

153 Purified RBDs were adjusted to 1 mg/mL in TBS in three to four replicates to evaluate the hydrodynamic
154 radius and molecular weight using a DynaPro Plate Reader II (Wyatt Technology) based on a globular
155 protein model. The sizes of these RBDs at room temperature were monitored for approximately 30 days.
156 Additionally, to evaluate the tendency of protein oligomerization among different RBDs, these purified
157 proteins were concentrated to approximately 7.5 mg/mL and serially diluted to approximately 0.66
158 mg/mL to calculate the diffusion interaction parameter (k_D) for each RBD using the following equation
159 [13, 25]:

160 $D=D_0(1+k_D \times c)$

161 where D is the measured diffusion coefficient, D_0 is the coefficient of the RBDs at an infinite dilution, and
162 k_D is the diffusion interaction parameter.

163 **2.5. Hydrophobicity Assessment by Extrinsic Fluorescence**

164 Purified RBDs and two controls (BSA and lysozyme) at concentrations of 7.8 ng/mL - 1.0 mg/mL in TBS
165 were probed with 12.5 nM Nile Red (Sigma-Aldrich, St. Louis, MO, USA; Cat # 3013) to generate extrinsic
166 fluorescence for surface hydrophobicity measurement, as described previously [13, 26]. The emission
167 spectra were recorded from 550 to 700 nm with the gain set at 130 after excitation at 520 nm using a
168 BioTek Synergy H4 plate reader. The surface hydrophobicity of RBD219-WT, RBD219-N1+His, RBD219-
169 N1C1, BSA, and lysozyme was determined using the slope of the emission peak intensity versus protein
170 concentration plot, as described previously [13, 27].

171 **2.6. Structural Assessment by Circular Dichroism (CD)**

172 Purified RBDs were diluted with deionized water to a final concentration of 0.2 mg/mL and CD spectra
173 were obtained from 250 to 190 nm with a Jasco J-1500 spectrophotometer set at 100 nm/min and a
174 response time of 1 s at 25°C. The obtained CD data were analyzed using a CD Analysis and Plotting Tool
175 (https://capito.uni-jena.de/index.php). In addition, the RBDs (0.5 mg/mL) were heated from 25 °C to 95
176 °C for a denaturation profile analysis.

177 **2.7. Structural Assessment by Thermal shift**

178 To evaluate the thermal stability of the tertiary structures for these three RBDs, a 1 mg/mL solution of
179 these proteins was mixed with Protein Thermal Shift™ Dye (Thermo Fisher, Waltham, MA, USA; Cat #
180 4461146) in triplicate as per the manufacturer's instructions. The samples were further heated from 25
181 °C to 95 °C to monitor the fluorescence intensity change using a ViiA™ 7 Real-Time PCR system.

182 **2.8. *in vitro* Functionality Assay by ELISA (ACE-2 binding)**

183 Ninety-six-well ELISA plates were coated with 100 µL 2 µg/mL RBD219-WT, RBD219-N1+His, RBD219-
184 N1C1 or BSA overnight in triplicate at 4°C followed by blocking with PBST/0.1% BSA. Once the plates
185 were blocked, 100 µL serially-diluted ACE-2-hFc (LakePharma, San Carlos, CA, USA; Cat # 46672) with an
186 initial concentration of 50 µg/mL was added to the wells. The plates were incubated at room
187 temperature for 2 h to allow ACE-2 to interact with each RBD. After this binding step, the plates were
188 washed with PBST four times followed by adding 100 µL 1:10,000 diluted HRP conjugated anti-human
189 IgG antibodies (GenScript, Piscataway, NJ, USA; Cat # A00166) and incubating for 1 h at room
190 temperature. Finally, 100 µL TMB substrate was added and incubated for 4 minutes in the dark to react
191 with HRP. The reaction was terminated with 100 µL HCl and absorption readings were taken at 450 nm
192 using a BioTek EPOCH 2 microplate reader. The results were analyzed using one-way ANOVA to evaluate
193 the statistic difference among the ELISA data generated for these three RBDs.
194

195 **3. RESULTS**

196 **3.1. Cloning, expression yield, and protein integrity**

197 The SARS-CoV-2 RBD vaccine antigen was originally designed and expressed using the same approach as
198 reported previously for the 70% homologous SARS-CoV RBD equivalent [13]. Three recombinant
199 constructs, RBD219-WT, RBD219-N1, and RBD219-N1C1 (Figure 1), were transformed into *P. pastoris* and
200 expressed at the 5 L fermentation scale.

201 The harvested fermentation supernatants (FS) of RBD219-WT, RBD219-N1, and RBD219-N1C1
202 were first analyzed by SDS-PAGE (Figure 2A) and their fermentation yields were quantified using
203 western blot (Supplementary Figure S1 and Table 1). When observing the FS profile for RBD219-WT on
204 SDS-PAGE gel (Figure 2A) and western blot membrane (Supplementary Figures S1A and S1C), large
205 amounts of protein bands were not recognized by the anti-SARS-CoV-2 Spike rabbit antibody, possibly
206 indicating the presence of host cell proteins (HCP), as well as yeast-derived hyperglycosylation. The
207 fermentation yield of RBD219-WT was estimated as 142 ± 8 mg/L of fermentation supernatant (Table 1).

208 To address the yeast-derived hyperglycosylation, we cloned and expressed RBD219-N1, in which
209 the N-terminal asparagine residue was removed. However, the fermentation yield of RBD219-N1 was
210 low (50 ± 13 mg/L of FS, Table 1), and the level of non-RBD proteins, likely HCP impurities, remained a
211 concern (Figure 2A and Supplementary Figures S1E and S1F). Due to the low yield and high level of non-
212 target-specific proteins, we were unable to purify tag-free RBD219-N1; thus, the hexahistidine tagged

213 construct (RBD219-N1+His) was used to obtain purified protein for later characterization. We also
214 observed that both RBD219-WT and RBD219-N1 formed dimers during fermentation. This suggested the
215 potential formation of an intermolecular disulfide bond between a potential free cysteine found in the
216 molecule, and therefore, a C538A-mutated form of RBD219-N1, RBD219-N1C1, was constructed. The
217 rationale is that there are nine cysteine residues in the RBD, where we assumed that eight residues
218 formed intramolecular bonds, leaving the last cysteine available for undesired intermolecular cross-
219 linking. The new C538A-mutated construct based on the RBD219-N1 backbone was able to express the
220 protein (RBD219-N1C1) with low yeast-derived hyperglycosylation, without the presence of extensive
221 non-RBD specific proteins or HCPs (Figure 2A), and with a fermentation yield of 280 ± 70 mg/L of FS
222 (Table 1). Even though, when fermentation yields were analyzed using western blot, one could observe a
223 dimer form of RBD219-N1C1 in the fermentation supernatant (Supplementary Figures S1I and S1K), such
224 dimer was not seen after purification in the final purified protein (Figure 2B). After treating the three
225 purified RBDs with PNGase-F to remove N-glycans, RBDs with similar size were observed (Figure 2C;
226 Please note that the band at 37 kDa was PNGase-F enzyme), which confirmed that the size differences
227 were likely due to the yeast-derived glycosylation.

228 **3.2. Size evaluation and protein-interaction assessment by DLS**

229 Dynamic light scattering (DLS) was used to evaluate the size of the purified RBDs in solution. The results
230 indicated that SARS-CoV-2 RBD219-WT has a slightly larger hydrodynamic radius (2.79 nm) than
231 RBD219-N1C1 (2.56 nm), and thus a higher calculated molecular weight (37.5 kDa) than RBD219-N1C1
232 (30 kDa) (Figure 3A), presumably due to the additional N-glycans. Interestingly, even though RBD219-
233 N1+His was less glycosylated due to the removal of the first glycosylation site, its hydrodynamic radius
234 (2.74 nm, 35.7 kDa) was similar to RBD219-WT; thus, it was suspected that potential oligomer formation
235 might have occurred. We further determined the diffusion interaction parameter (K_D) of three RBDs
236 (Figure 3B and Table 2) to evaluate the level of protein-protein interaction. The results indicated that all
237 three RBDs in TBS showed negative K_D values, implying protein attractions. However, it was noticed that
238 the K_D of RBD219-N1C1 (-16.3 mL/g) was similar to that of RBD219-WT (-14.9 mL/g), while that of
239 RBD219-N1+His was almost double (-29.7 mL/g), suggesting that RBD219-N1+His was even more prone
240 to form oligomers. Additionally, when we monitored the changes of the molecular weight for RBDs
241 stored at room temperature, it was observed that both the molecular weight of RBD219-WT and
242 RBD219-N1+His increased over time while the size of RBD219-N1C1 remained unchanged at ~30kDa
243 (Figure 3C).

244 **3.3. Hydrophobicity Assessment Using Extrinsic Fluorescence**

245 In this study, we used Nile red dye to probe these three purified RBDs to further evaluate their surface
246 hydrophobicity in TBS (Figures 4A and 4B). In general, a blue shift in emission peak wavelength λmax, or
247 a lower λmax value, indicates an increase in hydrophobicity due to more binding of Nile red to the
248 protein. The emission peaks λmax of RBD219-WT, RBD219-N1+His and RBD219-N1C1 in TBS were
249 determined as 658, 642 and 658 nm, respectively, indicating that RBD219-N1+His was more hydrophobic
250 than the other two RBD molecules, and these three RBDs were less hydrophobic than BSA (λmax = 634
251 nm) and more hydrophobic than lysozyme (λmax = 662 nm) (Figure 4A and Table 2). Additionally, the
252 surface hydrophobicity (S_0) calculated by the plot of fluorescence intensity vs concentration (Figure 4B
253 and Table 2) indicated that RBD219-WT and RBD219-N1C1 shared similar surface hydrophobicity ($S_{0;}$
254 $_{RBD219\text{-}WT}$ = 3.14 RFU*L/mg and $S_{0;\ RBD219\text{-}N1C1}$= 2.78 RFU*L/mg, respectively), while RBD219-N1+His showed
255 higher surface hydrophobicity ($S_{0;\ RBD219\text{-}N1+His}$ = 9.82 RFU*L/mg). Overall, the results indicated $S_{0;\ BSA}$ > $S_{0;}$
256 $_{RBD219\text{-}N1+His}$ > $S_{0;\ RBD219\text{-}WT}$ ~ $S_{0;\ RBD219\text{-}N1C1}$> $S_{0;\ Lysozyme}$) which were consistent with the observation based on
257 the wavelength of the extrinsic fluorescence emission peak.

258 **3.4. Secondary structure thermal stability assessment**

259 When far-UV CD spectrometry was performed to investigate the secondary structure of the three
260 purified RBDs, we observed a very similar secondary structure with 8% alpha-helix and 33% beta-sheet
261 (Fig. 5A). The thermal stability of the secondary structures was evaluated by heating the samples from
262 25 °C to 95 °C (Figs. 5B-5D) and CD melting curves and their derivatives were further examined at 231
263 nm (Figures. 5E-G). Based on the derivative, the average melting temperatures (Tm) were 52.3 °C, 53.3 °C,
264 and 51.9 °C for RBD219-WT, RBD219-N1+His, and RBD219-N1C1 respectively, suggesting similar thermal
265 stability.

266 **3.5. Tertiary structure thermal stability assessment**

267 Thermal shift assays were used for evaluating the thermal stability of the tertiary structure for the three
268 purified RBD proteins. The melting curve (Figure 6A) indicated that the onset temperature (T_{on}) for these
269 RBDs was approximately 38 °C, and the higher initial fluorescence of RBD219-N1+His observed in Figure
270 6A implied that more thermal shift dye bound to the molecule, which suggested higher hydrophobicity
271 of this molecule than for the other two RBDs. This finding is consistent with the observation in the
272 extrinsic fluorescence study. The derivatives indicated the melting temperatures (T_m) for RBD219-WT,
273 RBD219-N1+His, and RBD219-N1C1 as 50.6 ± 0.5 °C, 49.2 ± 0.5 °C and 50.8 ± 0.4 °C, respectively (Figure

274 6B). The similar thermal shift profile and thermal stability observed in the thermal shift assays further
275 suggested that the three RBDs likely shared similar tertiary structures.

276 **3.6. *in vitro* Functionality comparison of the RBD proteins in an ACE-2 binding study.**

277 To examine the *in vitro* functionality of the purified RBDs, we evaluated their ability to bind to a
278 recombinant ACE-2 protein reflecting the human receptor, ACE-2 (Figure 7). No significant binding was
279 found between ACE-2 and BSA, indicating that RBD binding to ACE-2 in this assay was specific. No
280 statistical difference (p=0.935) among the binding of ACE-2 to RBD219-WT, RBD219-N1+His, or RBD219-
281 N1C1 proteins was seen, indicating that hyperglycosylation on RBD219-WT, the N1 deletion, or the
282 cysteine mutation on RBD219-N1C1 did not impact the in vitro ability of these proteins to bind to the
283 ACE-2 receptor.
284
285

286 **4. DISCUSSION**

287 As COVID-19 spreads worldwide, it has now become predominant among populations living in poverty,
288 especially in crowded urban areas and mega-cities located in low- and middle-income countries (LMICs)
289 [28]; take India and Brazil as examples, the majority of the cases, over 1.5 million and 1.1 million, are
290 reported in Maharashtra state, where Mumbai city is located, and State of Sao Paulo, where Sao Paulo
291 city is located, respectively [1]. In the race to develop a vaccine against COVID-19, many candidates have
292 moved forward to clinical trials. Among those, recombinant protein-based vaccine candidates are mainly
293 expressed using mammalian cells or insect cells [20, 29-32] that are not always considered the most
294 cost-effective and globally accessible. To develop a safe and effective vaccine that is also affordable and
295 easily accessible, we focus our vaccine development efforts using a yeast-based expression system. For
296 example, a low-cost recombinant hepatitis B vaccine expressed in yeast is produced in several LMICs,
297 including Brazil, Cuba, India, Indonesia, and elsewhere [26]. In this study, we evaluated three different
298 yeast-expressed antigen constructs, RBD219-WT, RBD219-N1, and RBD219-N1C1 by comparing their
299 fermentation yield, biophysical characteristics, and *in vitro* functionality.

300 During the initial expression of RBD219-WT, it was found that the protein was hyperglycosylated,
301 which could pose challenges during the production process and impact yields, reproducibility, and
302 stability. Based on the lessons learned from the development of the process used for the yeast-
303 expressed SARS-CoV RBD [4], we developed the RBD219-N1 construct to remove the first amino acid

304 (corresponded to the glycosylation site N-331 on S protein). However, we discovered intermolecular
305 disulfide bridging was occurring in the expression of the RBD219-N1 protein. When investigating the
306 RBD structure by extracting residues 331-549 from PDB ID 6XEY [33], one could see that the RBD
307 contained a total of nine cysteines, eight of which were conserved and form disulfide bonds
308 (Supplementary Figure S2) while one, C538, was free to form an intermolecular disulfide bond, which
309 could lead to a potential stability issue. To alleviate this, an additional construct (RBD219-N1C1) with
310 C538 mutated to alanine was created.

311 Densitometry of Coomassie blue-stained SDS-PAGE gels is generally used to evaluate the protein
312 yield during the fermentation runs [13, 34, 35]; however, in this case, due to the high levels of likely HCP
313 impurities observed in the FS for both RBD219-WT and RBD219-N1, the method could not be applied,
314 and hence, we had to use western blots to assess the yields of the unpurified proteins after
315 fermentation. By using this immunoblot-based method, we were able to quantify the protein production
316 in the fermentation supernatants without the interference of HCPs. The protein yield quantified using
317 western blot indicated that in the FS, RBD219-N1C1 had the highest yield among the three constructs.
318 Due to the high HCP background found in RBD219-WT FS, QXL was used to first capture most of the
319 HCPs leaving the target protein in the flow-through, this flow-through was further purified using Butyl
320 HP followed by SEC. As for the RBD219-N1 construct, the target protein was barely visible in the FS when
321 analyzed by SDS-PAGE, and thus, we were unable to purify tag-free RBD219-N1 protein; instead, a
322 hexahistidine-tagged protein was expressed and purified by metal affinity column. RBD219-N1C1,
323 however, was the dominant protein found in the FS and could be easily purified by a simple two-step
324 purification scheme similar to the yeast expressed SARS-CoV RBD219-N1 [13]. Therefore, RBD219-N1C1
325 was selected for further process development and scale-up.

326 While further evaluating the purified RBDs by SDS-PAGE and western blot, dimers were
327 observed in purified RBD219-WT and RBD219-N1+His only under non-reduced condition, suggesting
328 potential intermolecular disulfide bond formation. Mutating the free cysteine residue successfully
329 reduced the propensity of oligomerization during fermentation, as no dimer formation was found in
330 purified RBD219-N1C1. Additionally, as part of the stability evaluation, when monitoring the size using
331 DLS, purified RBD219-N1C1 remained at 30 kDa for approximately 30 days while RBD219-WT and
332 RBD219-N1+His continued to form oligomers. However, when determining the diffusion interaction
333 parameter (K_D) using DLS and assessing surface hydrophobicity (S_0) using extrinsic fluorescence, we
334 observed that RBD219-WT and RBD219-N1C1 shared similar K_D and S_0 values while RBD219-N1+His

335 showed a lower K_D and higher S_0 values, implying a higher tendency to form oligomers. The high
336 oligomerization tendency for RBD219-N1+His is likely due to the additional hexahistidine as we also
337 performed an extrinsic fluorescence study on a hexahistidine-tagged version of RBD219-WT (RBD219-
338 WT+His), and the results indicated a much higher surface hydrophobicity value (12.7 RFU*L/mg;
339 supplementary Figure S3) than RBD219-WT. The pKa of the imidazole ring in histidine is approximately
340 6.0, which suggested that at pH 7.5, this aromatic ring is likely non-protonated, which makes this
341 additional hexahistidine more hydrophobic.

342 When assessing the secondary and tertiary structures, far-UV circular dichroism spectra
343 revealed that RBD219-WT, RBD219-N1+His, and RBD219-N1C1 purified proteins had similar secondary
344 structures. The melting temperatures evaluated by CD and thermal shift assays also indicated that these
345 three RBDs shared similar thermal stability for both secondary and tertiary structures. Additionally, Li *et*
346 *al*. have derived neutralizing monoclonal antibodies from COVID-19 patients, and the most potent
347 neutralizing antibodies that recognized the RBD blocked ACE-2 binding [33]. Thus, confirming the ACE-2
348 binding is crucial to ensure the epitopes within the ACE-2 binding site are still intact. A similar binding
349 affinity to ACE-2 among all three RBDs was observed, suggesting that removing the first amino acid
350 and/or mutating the free cysteine to alanine did not impact protein structure or functionality.

351 In summary, the RBD219-N1C1 construct provided higher protein yields in the FS (~280mg
352 RBD219-N1C1/L of FS), and the fermentation process was less impacted based on the detection of fewer
353 impurities. This facilitated the use of a straight forward and relatively efficient initial purification scheme,
354 with an approximate recovery yield of ~189 mg of purified RBD219-N1C1/L of FS. With respect to
355 biophysical characteristics, unlike RBD219-WT and RBD219-N1 proteins, the purified RBD219-N1C1
356 protein does not appear to form dimers via intermolecular disulfide bridging. Additionally, RBD219-WT,
357 RBD219-N1+His, and RBD219-N1C1 showed similar thermal stability and the same binding affinity to
358 their receptor, ACE-2, further suggesting that the deletion of the first amino acid and mutation of the
359 free cysteine did not cause any significant structural changes. Based on all the data, we conclude that
360 the RBD219-N1C1 construct is the superior candidate to move forward for process optimization and
361 further development and scale-up, in preparation for technology transfer and manufacturing and for its
362 advancement towards clinical development.

363

364 **Author contributions**

Wen-Hsiang Chen conceived the study, designed and performed experiments, interpreted data, and wrote the manuscript; Junfei Wei performed experiments and interpreted data; Rakhi Tyagi Kundu, Rakesh Adhikari, Zhuyun Liu, Jungsoon Lee, Leroy Versteeg, Cristina Poveda, Brian Keegan, Maria Jose Villar, Ana C. de Araujo Leao and Joanne Altieri Rivera performed experiments; Jeroen Pollet participated in interpreting data, reviewed/edited the manuscript; Ulrich Strych reviewed/edited the manuscript; Bin Zhan reviewed/edited the manuscript; Peter Hotez reviewed/edited the manuscript; Maria Elena Bottazzi, reviewed/edited the manuscript Ulrich Strych, Bin Zhan, Peter Hotez and Maria Elena Bottazzi also provided scientific guidance on the project.

Conflict of interest

The authors declare that Baylor College of Medicine recently licensed the RBD219-N1C1 technology to an Indian manufacturer for further development. The research conducted in this paper was performed in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

This work was supported by the Robert J. Kleberg Jr. and Helen C. Kleberg Foundation; Fifth Generation, Inc. (Tito's Handmade Vodka); JPB Foundation, NIH-NIAID (AI14087201); and Texas Children's Hospital Center for Vaccine Development Intramural Funds. We also would like to thank PATH Center for Vaccine Innovation and Access (Seattle, WA, USA) for their guidance as well as technical and intellectual support.

References

[1] E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time, Lancet Infect Dis 20(5) (2020) 533-534.
[2] S. Jiang, M.E. Bottazzi, L. Du, S. Lustigman, C.T. Tseng, E. Curti, K. Jones, B. Zhan, P.J. Hotez, Roadmap to developing a recombinant coronavirus S protein receptor-binding domain vaccine for severe acute respiratory syndrome, Expert Rev Vaccines 11(12) (2012) 1405-13.
[3] L. Du, G. Zhao, Y. He, Y. Guo, B.J. Zheng, S. Jiang, Y. Zhou, Receptor-binding domain of SARS-CoV spike protein induces long-term protective immunity in an animal model, Vaccine 25(15) (2007) 2832-8.
[4] W.H. Chen, L. Du, S.M. Chag, C. Ma, N. Tricoche, X. Tao, C.A. Seid, E.M. Hudspeth, S. Lustigman, C.T. Tseng, M.E. Bottazzi, P.J. Hotez, B. Zhan, S. Jiang, Yeast-expressed recombinant protein of the receptor-binding domain in SARS-CoV spike protein with deglycosylated forms as a SARS vaccine candidate, Hum Vaccin Immunother 10(3) (2014) 648-58.
[5] W.H. Chen, U. Strych, P.J. Hotez, M.E. Bottazzi, The SARS-CoV-2 Vaccine Pipeline: an Overview, Curr Trop Med Rep (2020) 1-4.

398 [6] M.P. Nyon, L. Du, C.K. Tseng, C.A. Seid, J. Pollet, K.S. Naceanceno, A. Agrawal, A. Algaissi, B.H. Peng,
399 W. Tai, S. Jiang, M.E. Bottazzi, U. Strych, P.J. Hotez, Engineering a stable CHO cell line for the expression
400 of a MERS-coronavirus vaccine antigen, Vaccine 36(14) (2018) 1853-1862.
401 [7] W. Tai, L. He, X. Zhang, J. Pu, D. Voronin, S. Jiang, Y. Zhou, L. Du, Characterization of the receptor-
402 binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral
403 attachment inhibitor and vaccine, Cell Mol Immunol (2020).
404 [8] B.F. Haynes, TBD - activ group, SCIENCE TRANSLATIONAL MEDICINE in press (2020).
405 [9] M. Jaume, M.S. Yip, Y.W. Kam, C.Y. Cheung, F. Kien, A. Roberts, P.H. Li, I. Dutry, N. Escriou, M. Daeron,
406 R. Bruzzone, K. Subbarao, J.S. Peiris, B. Nal, R. Altmeyer, SARS CoV subunit vaccine: antibody-mediated
407 neutralisation and enhancement, Hong Kong Med J 18 Suppl 2 (2012) 31-6.
408 [10] Q. Wang, L. Zhang, K. Kuwahara, L. Li, Z. Liu, T. Li, H. Zhu, J. Liu, Y. Xu, J. Xie, H. Morioka, N.
409 Sakaguchi, C. Qin, G. Liu, Immunodominant SARS Coronavirus Epitopes in Humans Elicited both
410 Enhancing and Neutralizing Effects on Infection in Non-human Primates, ACS Infect Dis 2(5) (2016) 361-
411 76.
412 [11] C.T. Tseng, E. Sbrana, N. Iwata-Yoshikawa, P.C. Newman, T. Garron, R.L. Atmar, C.J. Peters, R.B.
413 Couch, Immunization with SARS coronavirus vaccines leads to pulmonary immunopathology on
414 challenge with the SARS virus, PLoS One 7(4) (2012) e35421.
415 [12] W.H. Chen, X. Tao, A.S. Agrawal, A. Algaissi, B.H. Peng, J. Pollet, U. Strych, M.E. Bottazzi, P.J. Hotez, S.
416 Lustigman, L. Du, S. Jiang, C.K. Tseng, Yeast-expressed SARS-CoV recombinant receptor-binding domain
417 (RBD219-N1) formulated with aluminum hydroxide induces protective immunity and reduces immune
418 enhancement, Vaccine 38(47) (2020) 7533-7541.
419 [13] W.H. Chen, S.M. Chag, M.V. Poongavanam, A.B. Biter, E.A. Ewere, W. Rezende, C.A. Seid, E.M.
420 Hudspeth, J. Pollet, C.P. McAtee, U. Strych, M.E. Bottazzi, P.J. Hotez, Optimization of the Production
421 Process and Characterization of the Yeast-Expressed SARS-CoV Recombinant Receptor-Binding Domain
422 (RBD219-N1), a SARS Vaccine Candidate, J Pharm Sci 106(8) (2017) 1961-1970.
423 [14] S.K. Wong, W. Li, M.J. Moore, H. Choe, M. Farzan, A 193-amino acid fragment of the SARS
424 coronavirus S protein efficiently binds angiotensin-converting enzyme 2, J Biol Chem 279(5) (2004) 3197-
425 201.
426 [15] S. Jiang, M.E. Bottazzi, L. Du, S. Lustigman, C.-T.K. Tseng, E. Curti, K. Jones, B. Zhan, P.J. Hotez,
427 Roadmap to developing a recombinant coronavirus S protein receptor-binding domain vaccine for
428 severe acute respiratory syndrome, Expert Review of Vaccines 11(12) (2012) 1405-1413.
429 [16] L. Du, G. Zhao, C.C. Chan, S. Sun, M. Chen, Z. Liu, H. Guo, Y. He, Y. Zhou, B.J. Zheng, S. Jiang,
430 Recombinant receptor-binding domain of SARS-CoV spike protein expressed in mammalian, insect and E.
431 coli cells elicits potent neutralizing antibody and protective immunity, Virology 393(1) (2009) 144-50.
432 [17] L. Du, G. Zhao, C.C. Chan, L. Li, Y. He, Y. Zhou, B.J. Zheng, S. Jiang, A 219-mer CHO-expressing
433 receptor-binding domain of SARS-CoV S protein induces potent immune responses and protective
434 immunity, Viral Immunol 23(2) (2010) 211-9.
435 [18] L. Du, G. Zhao, L. Li, Y. He, Y. Zhou, B.J. Zheng, S. Jiang, Antigenicity and immunogenicity of SARS-
436 CoV S protein receptor-binding domain stably expressed in CHO cells, Biochem Biophys Res Commun
437 384(4) (2009) 486-90.
438 [19] L. Du, G. Zhao, Y. He, Y. Guo, B.J. Zheng, S. Jiang, Y. Zhou, Receptor-binding domain of SARS-CoV
439 spike protein induces long-term protective immunity in an animal model, Vaccine 25 (2007).
440 [20] J. Yang, W. Wang, Z. Chen, S. Lu, F. Yang, Z. Bi, L. Bao, F. Mo, X. Li, Y. Huang, W. Hong, Y. Yang, Y.
441 Zhao, F. Ye, S. Lin, W. Deng, H. Chen, H. Lei, Z. Zhang, M. Luo, H. Gao, Y. Zheng, Y. Gong, X. Jiang, Y. Xu, Q.
442 Lv, D. Li, M. Wang, F. Li, S. Wang, G. Wang, P. Yu, Y. Qu, L. Yang, H. Deng, A. Tong, J. Li, Z. Wang, J. Yang,
443 G. Shen, Z. Zhao, Y. Li, J. Luo, H. Liu, W. Yu, M. Yang, J. Xu, J. Wang, H. Li, H. Wang, D. Kuang, P. Lin, Z. Hu,
444 W. Guo, W. Cheng, Y. He, X. Song, C. Chen, Z. Xue, S. Yao, L. Chen, X. Ma, S. Chen, M. Gou, W. Huang, Y.
445 Wang, C. Fan, Z. Tian, M. Shi, F.S. Wang, L. Dai, M. Wu, G. Li, G. Wang, Y. Peng, Z. Qian, C. Huang, J.Y. Lau,

446 Z. Yang, Y. Wei, X. Cen, X. Peng, C. Qin, K. Zhang, G. Lu, X. Wei, A vaccine targeting the RBD of the S
447 protein of SARS-CoV-2 induces protective immunity, Nature (2020).
448 [21] W.H. Chen, P.J. Hotez, M.E. Bottazzi, Potential for developing a SARS-CoV receptor-binding domain
449 (RBD) recombinant protein as a heterologous human vaccine against coronavirus infectious disease
450 (COVID)-19, Hum Vaccin Immunother 16(6) (2020) 1239-1242.
451 [22] M. Hoffmann, H. Kleine-Weber, N. Krueger, M.A. Mueller, C. Drosten, S. Poehlmann, The novel
452 coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease
453 TMPRSS2 for entry into target cells, bioRxiv (2020).
454 [23] P. Zhou, X.L. Yang, X.G. Wang, B. Hu, L. Zhang, W. Zhang, H.R. Si, Y. Zhu, B. Li, C.L. Huang, H.D. Chen,
455 J. Chen, Y. Luo, H. Guo, R.D. Jiang, M.Q. Liu, Y. Chen, X.R. Shen, X. Wang, X.S. Zheng, K. Zhao, Q.J. Chen, F.
456 Deng, L.L. Liu, B. Yan, F.X. Zhan, Y.Y. Wang, G.F. Xiao, Z.L. Shi, A pneumonia outbreak associated with a
457 new coronavirus of probable bat origin, Nature 579(7798) (2020) 270-273.
458 [24] J. Pollet, W.-H. Chen, L. Versteeg, B. Keegan, B. Zhan, J. Wei, Z. Liu, J. Lee, R. Kundu, R. Adhikari, C.
459 Poveda, M.J. Villar Mondragon, A.C. de Araujo Leao, J. Altieri Rivera, P.M. Gillespie, U. Strych, P.J. Hotez,
460 M.E. Bottazzi, SARS-CoV-2 RBD219-N1C1: A Yeast-Expressed SARS-CoV-2 Recombinant Receptor-
461 Binding Domain Candidate Vaccine Stimulates Virus Neutralizing Antibodies and T-cell Immunity in Mice,
462 bioRxiv (2020) 2020.11.04.367359.
463 [25] A. Saluja, R.M. Fesinmeyer, S. Hogan, D.N. Brems, Y.R. Gokarn, Diffusion and sedimentation
464 interaction parameters for measuring the second virial coefficient and their utility as predictors of
465 protein aggregation, Biophys J 99(8) (2010) 2657-65.
466 [26] D.L. Sackett, J. Wolff, Nile red as a polarity-sensitive fluorescent probe of hydrophobic protein
467 surfaces, Anal Biochem 167(2) (1987) 228-34.
468 [27] A. Kato, S. Nakai, Hydrophobicity determined by a fluorescence probe method and its correlation
469 with surface properties of proteins, Biochim Biophys Acta 624(1) (1980) 13-20.
470 [28] P.J. Hotez, M.E. Bottazzi, S.K. Singh, P.J. Brindley, S. Kamhawi, Will COVID-19 become the next
471 neglected tropical disease?, PLoS neglected tropical diseases 14(4) (2020) e0008271.
472 [29] CHICTR, Randomized double blind, placebo controlled phase I trial for anti novel coronavirus
473 pneumonia (COVID-19) recombinant vaccine (Sf9), 2020.
474 http://www.chictr.org.cn/showprojen.aspx?proj=60581. (Accessed 10/15/2020.
475 [30] Clinicaltrials_gov, Clinical Study of Recombinant Novel Coronavirus Vaccine, 2020.
476 https://clinicaltrials.gov/ct2/show/NCT04466085. (Accessed 10/9/2020.
477 [31] Clinicaltrials_gov, A Study to Evaluate the Safety and Immunogenicity of COVID-19 (AdimrSC-2f)
478 Vaccine, 2020. https://clinicaltrials.gov/ct2/show/NCT04522089. (Accessed 10/9/2020.
479 [32] C. Keech, G. Albert, I. Cho, A. Robertson, P. Reed, S. Neal, J.S. Plested, M. Zhu, S. Cloney-Clark, H.
480 Zhou, G. Smith, N. Patel, M.B. Frieman, R.E. Haupt, J. Logue, M. McGrath, S. Weston, P.A. Piedra, C.
481 Desai, K. Callahan, M. Lewis, P. Price-Abbott, N. Formica, V. Shinde, L. Fries, J.D. Lickliter, P. Griffin, B.
482 Wilkinson, G.M. Glenn, Phase 1-2 Trial of a SARS-CoV-2 Recombinant Spike Protein Nanoparticle Vaccine,
483 N Engl J Med (2020).
484 [33] L. Liu, P. Wang, M.S. Nair, J. Yu, M. Rapp, Q. Wang, Y. Luo, J.F. Chan, V. Sahi, A. Figueroa, X.V. Guo, G.
485 Cerutti, J. Bimela, J. Gorman, T. Zhou, Z. Chen, K.Y. Yuen, P.D. Kwong, J.G. Sodroski, M.T. Yin, Z. Sheng, Y.
486 Huang, L. Shapiro, D.D. Ho, Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike,
487 Nature 584(7821) (2020) 450-456.
488 [34] W.H. Chen, M.P. Nyon, M.V. Poongavanam, Z. Liu, A.B. Biter, R.T. Kundu, U. Strych, P.J. Hotez, M.E.
489 Bottazzi, Process Characterization and Biophysical Analysis for a Yeast-Expressed Phlebotomus papatasi
490 Salivary Protein (PpSP15) as a Leishmania Vaccine Candidate, J Pharm Sci (2020).
491 [35] E. Curti, C.A. Seid, E. Hudspeth, L. Center, W. Rezende, J. Pollet, C. Kwityn, M. Hammond, R.K.
492 Matsunami, D.A. Engler, P.J. Hotez, M. Elena Bottazzi, Optimization and revision of the production

process of the Necator americanus glutathione S-transferase 1 (Na-GST-1), the lead hookworm vaccine recombinant protein candidate, Hum Vaccin Immunother 10(7) (2014) 1914-25.
[36] Y. Wan, J. Shang, R. Graham, R.S. Baric, F. Li, Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus, J Virol 94(7) (2020).

498 **Table**

499

500 **Table 1.** Assessment of the fermentation supernatants for clones RBD219-219, RBD219-N1, and RBD219-
501 N1C1. *Fermentation yield was measured by probing the RBD with the same specific antibody followed
502 by densitometry (Supplementary Figure S1). FS: fermentation supernatant; WB: western blot.

Molecule	Fermentation yield* (mg RBD/L of FS)	Level of Impurity on SDS-PAGE	Hyperglycosylation on WB	Dimer formation on WB
RBD219-WT	142 ± 8	High	Yes	Yes
RBD219-N1	50 ± 13	Mid	No	Yes
RBD219-N1C1	280 ± 70	Low	No	Yes

503

504

505

506

507 **Table 2.** Aggregation assessment by diffusion interaction parameter (K_D) using DLS and hydrophobicity
508 assessment by extrinsic fluorescence.

Protein	Dynamic light scattering	Extrinsic fluorescence	
	Diffusion interaction parameter K_D (mL/g)	Peak wavelength λ_{max} (nm)	Surface hydrophobicity S_0 (RFU*L/mg)
RBD219-WT	-14.9	658	3.1
RBD219-N1+His	-29.7	642	9.8
RBD219-N1C1	-16.3	658	2.8
BSA	N/A	634	53.4
Lysozyme	N/A	662	1.1

509

510

511 **Figure Legends**
512
513 Figure 1. Sequence comparison among RBD219-WT, RBD219-N1, and RBD219-N1C1. The deleted
514 asparagine in the first glycosylation site is highlighted in red while the mutated cysteine (to alanine) is
515 highlighted in green. The receptor-binding motif is highlighted in blue [21, 36].
516

517 **Figure 2.** Coomassie Blue stained SDS-PAGE and western blot probed with anti-SARS-CoV-2 Spike rabbit antibody.
518 (A) SDS-PAGE gel of 10 µL fermentation supernatant for tagged-free RBD219-WT, RBD219-N1, and RBD219-N1C1;
519 (B) Coomassie Blue stained SDS-PAGE gel of 3 µg purified RBDs or western blot of 1.5 µg of the purified RBDs under
520 non-reduced and reduced conditions. (C) SDS-PAGE of 3 µg PNGase-F treated purified RBDs; please note that the
521 37 kDa band observed on the PNGase-F treated gel is the N-glycosidase PNGase-F enzyme.

522

523 **Figure 3.** Dynamic light scattering results for RBD219-WT, RBD219-N1+His, and RBD219-N1C1. (A)
524 Measured Stokes radii and molecular weights. (B) Diffusion coefficient vs. concentration plot to evaluate
525 the diffusion interaction parameter (C) Stability study to monitor the changes of the molecular weight.

526

527 **Figure 4.** Extrinsic fluorescence results for RBD219-WT, RBD219-N1+His, and RBD219-N1C1. (A)
528 Excitation wavelength scan to obtain the peak emission wavelength λmax; (B) Fluorescence intensity vs
529 concentration plot to evaluate surface hydrophobicity.
530

531 **Figure 5.** Circular dichroism data for RBD219-WT, RBD219-N1+His, and RBD219-N1C1. (A) Circular
532 dichroism spectra; Thermal map of circular dichroism spectra for (B) RBD219-WT, (C) RBD219-N1+His,
533 and (D) RBD219-N1C1. Melting profile of (E) RBD219-WT, (F) RBD219-N1+His, and (G) RBD219-N1C1.
534

535 **Figure 6.** Thermal shift assay for RBD219-WT, RBD219-N1+His, and RBD219-N1C1. (A) the fluorescence-
536 temperature plot and (B) the derivative fluorescence-temperature plot.
537

538 **Figure 7.** ACE-2 binding study of RBD219-WT, RBD219-N1+His, and RBD219-N1C1.
539

Figure 1

```
RBD219-WT    NITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDL 60
RBD219-N1    -ITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDL 59
RBD219-N1C1  -ITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDL 59
              ************************************************************

RBD219-WT    CFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYN 120
RBD219-N1    CFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYN 119
RBD219-N1C1  CFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYN 119
             ************************************************************

RBD219-WT    YLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRV 180
RBD219-N1    YLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRV 179
RBD219-N1C1  YLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRV 179
             ************************************************************

RBD219-WT    VVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT 219
RBD219-N1    VVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGT 218
RBD219-N1C1  VVLSFELLHAPATVCGPKKSTNLVKNKAVNFNFNGLTGT 218
             ***************************.************
```

Figure 2

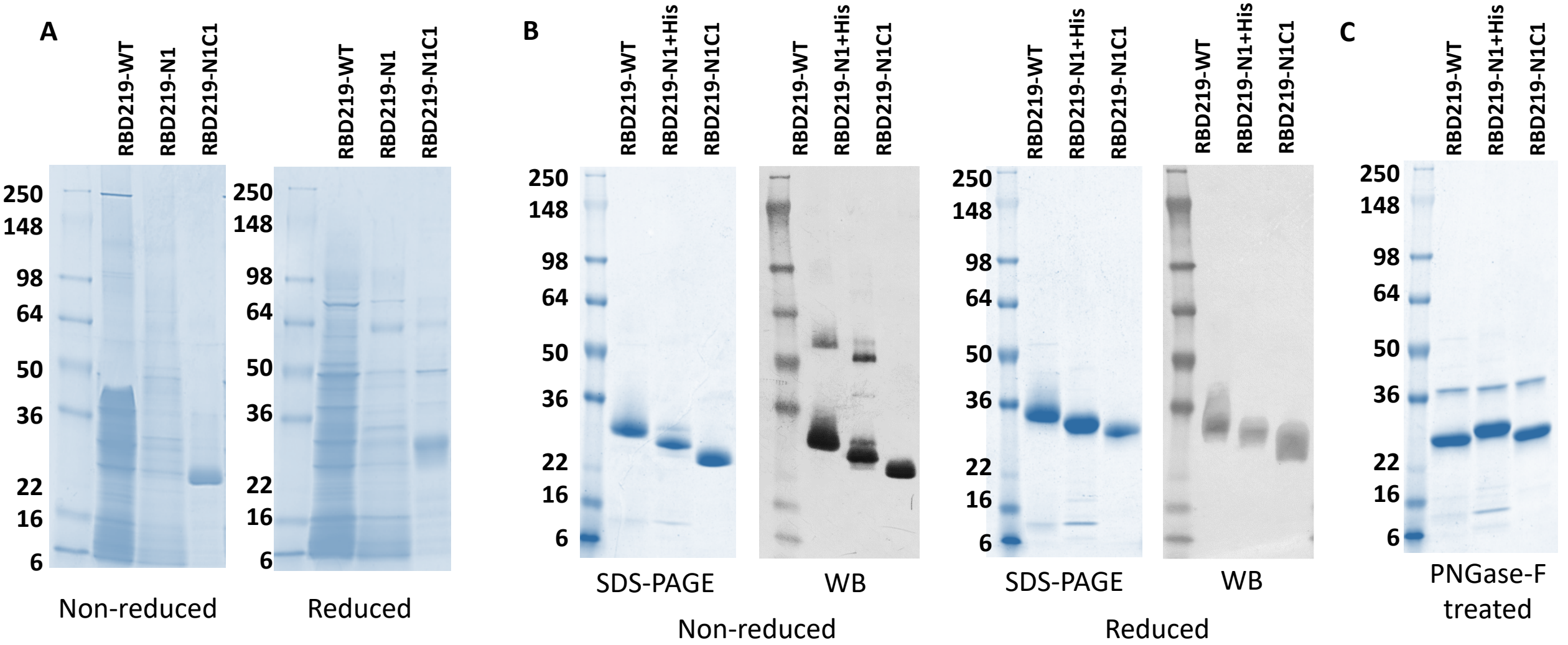

A
RBD219-WT
RBD219-N1
RBD219-N1C1
250
148
98
64
50
36
22
16
6
Non-reduced
Reduced
B
RBD219-N1+His
SDS-PAGE
WB
C
PNGase-F treated

Figure 3

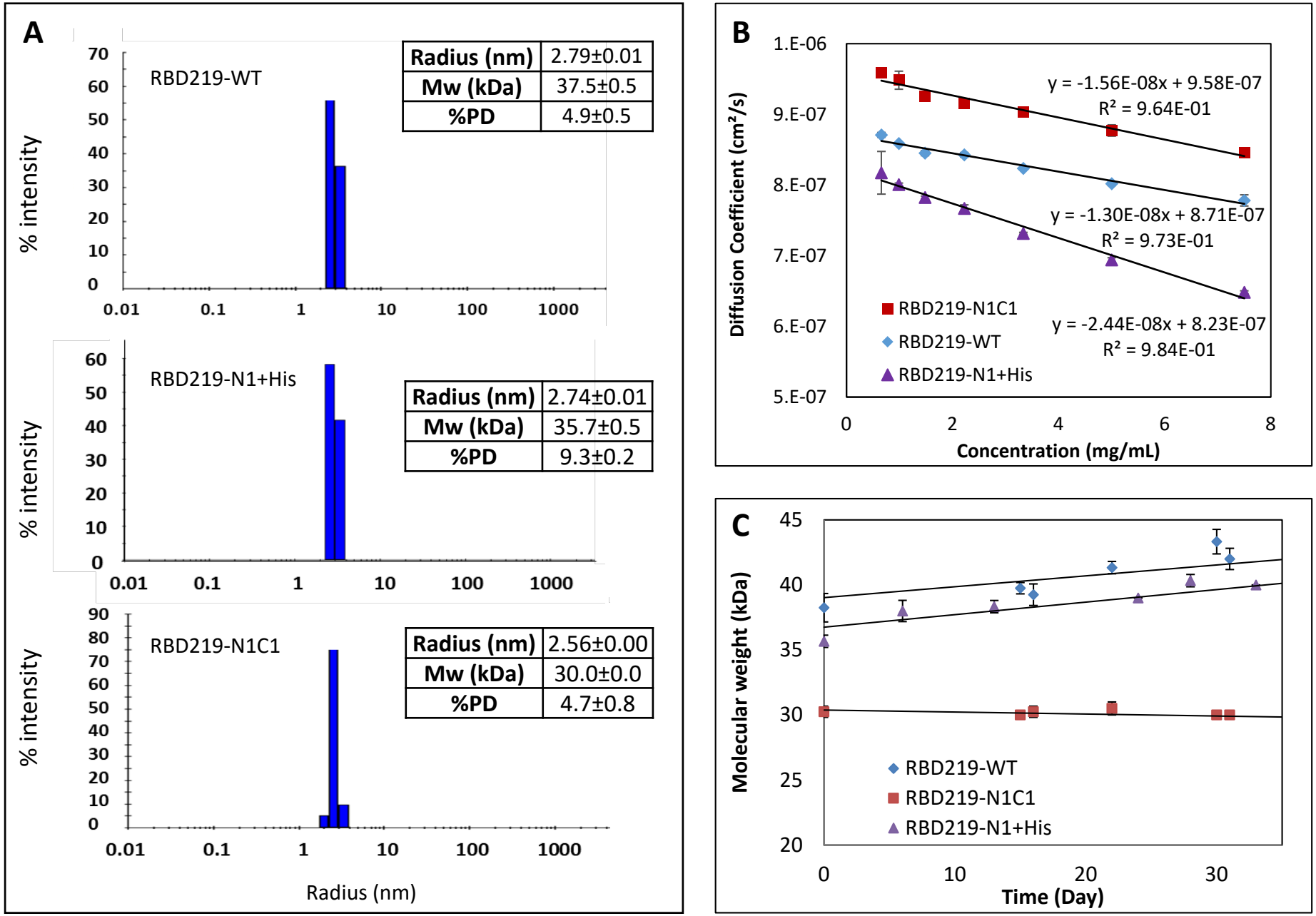

A
RBD219-WT
Radius (nm) 2.79±0.01
Mw (kDa) 37.5±0.5
%PD 4.9±0.5
RBD219-N1+His
Radius (nm) 2.74±0.01
Mw (kDa) 35.7±0.5
%PD 9.3±0.2
RBD219-N1C1
Radius (nm) 2.56±0.00
Mw (kDa) 30.0±0.0
%PD 4.7±0.8
% intensity
Radius (nm)
B
Diffusion Coefficient (cm²/s)
Concentration (mg/mL)
y = -1.56E-08x + 9.58E-07
R² = 9.64E-01
y = -1.30E-08x + 8.71E-07
R² = 9.73E-01
y = -2.44E-08x + 8.23E-07
R² = 9.84E-01
RBD219-N1C1
RBD219-WT
RBD219-N1+His
C
Molecular weight (kDa)
Time (Day)
RBD219-WT
RBD219-N1C1
RBD219-N1+His

Figure 4

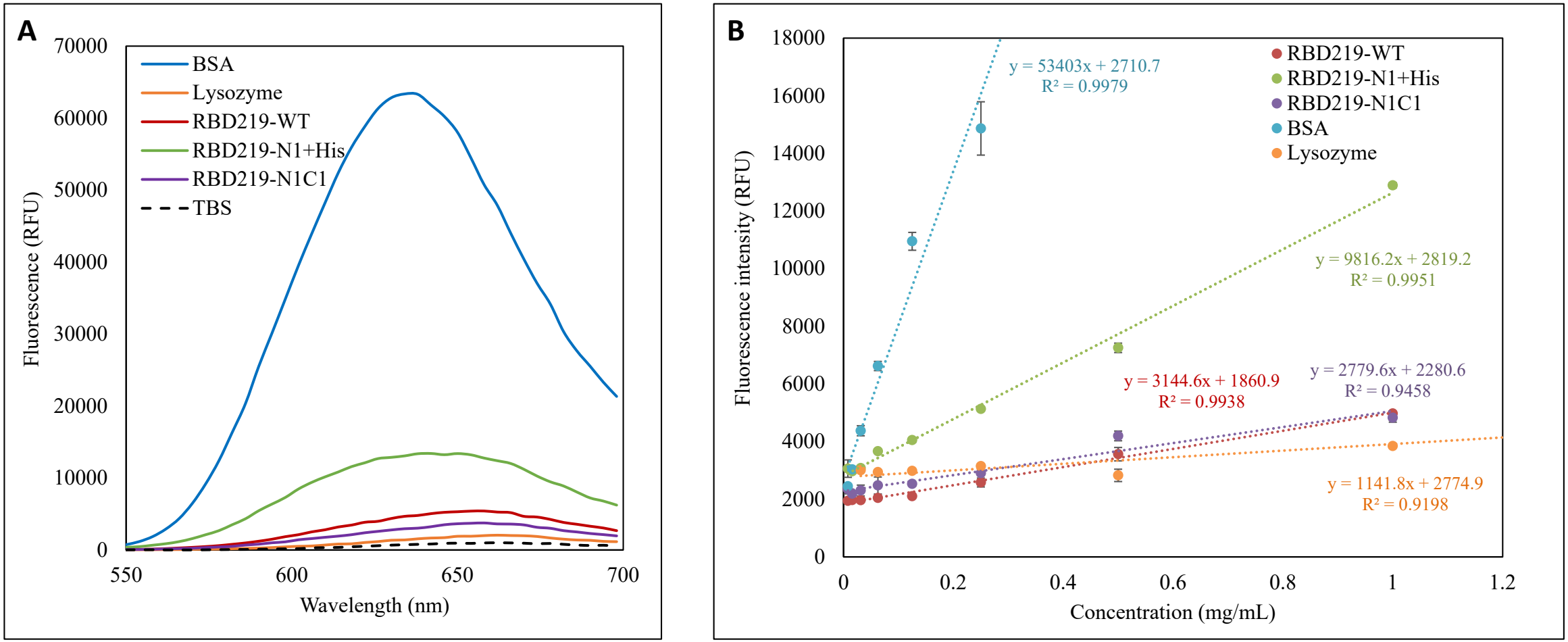

A
70000
60000
50000
40000
30000
20000
10000
0
550
600
650
700
Fluorescence (RFU)
Wavelength (nm)
BSA
Lysozyme
RBD219-WT
RBD219-N1+His
RBD219-N1C1
TBS
B
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
0
0.2
0.4
0.6
0.8
1
1.2
Fluorescence intensity (RFU)
Concentration (mg/mL)
y = 53403x + 2710.7
R² = 0.9979
y = 9816.2x + 2819.2
R² = 0.9951
y = 3144.6x + 1860.9
R² = 0.9938
y = 2779.6x + 2280.6
R² = 0.9458
y = 1141.8x + 2774.9
R² = 0.9198
RBD219-WT
RBD219-N1+His
RBD219-N1C1
BSA
Lysozyme

Figure 5

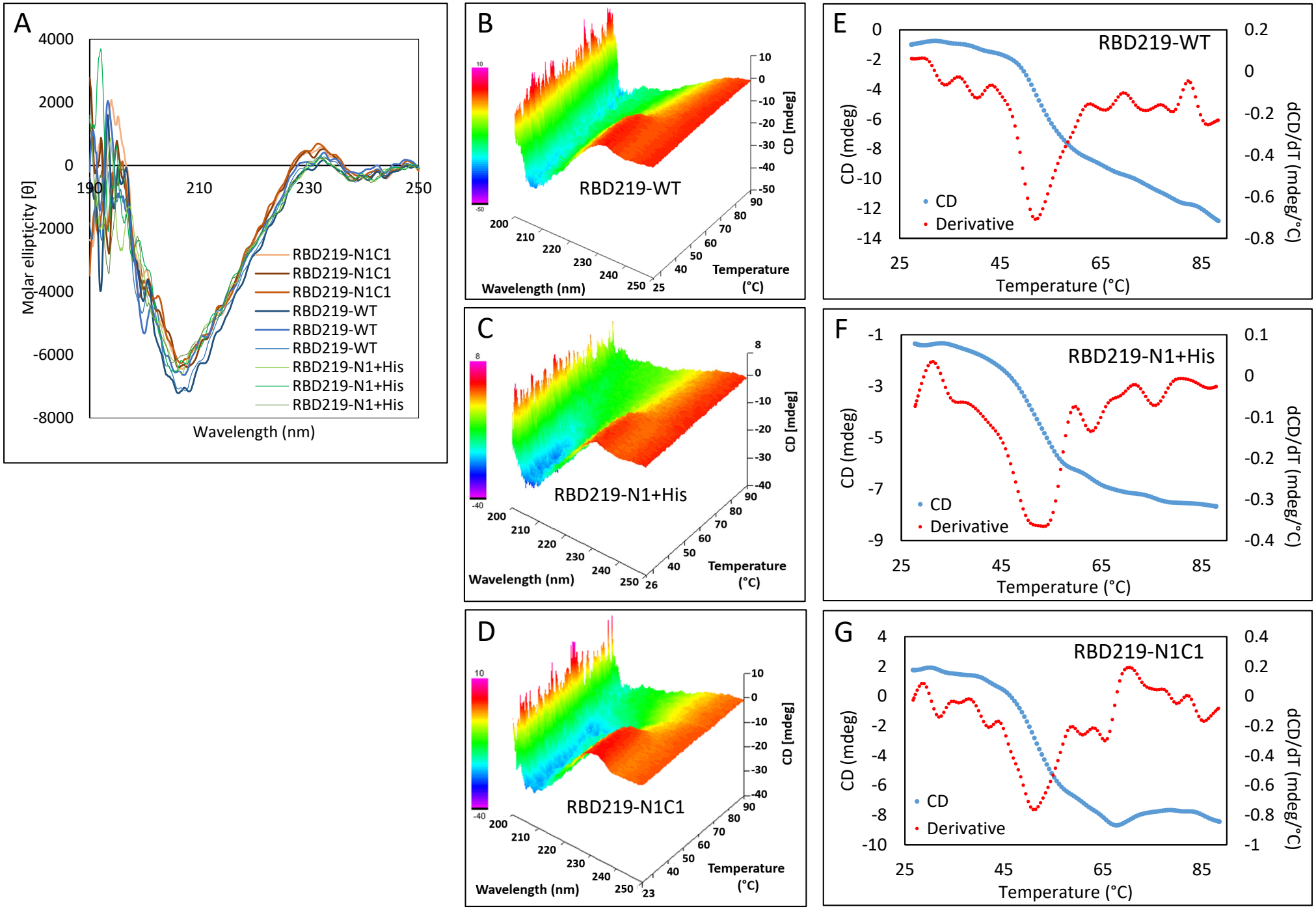
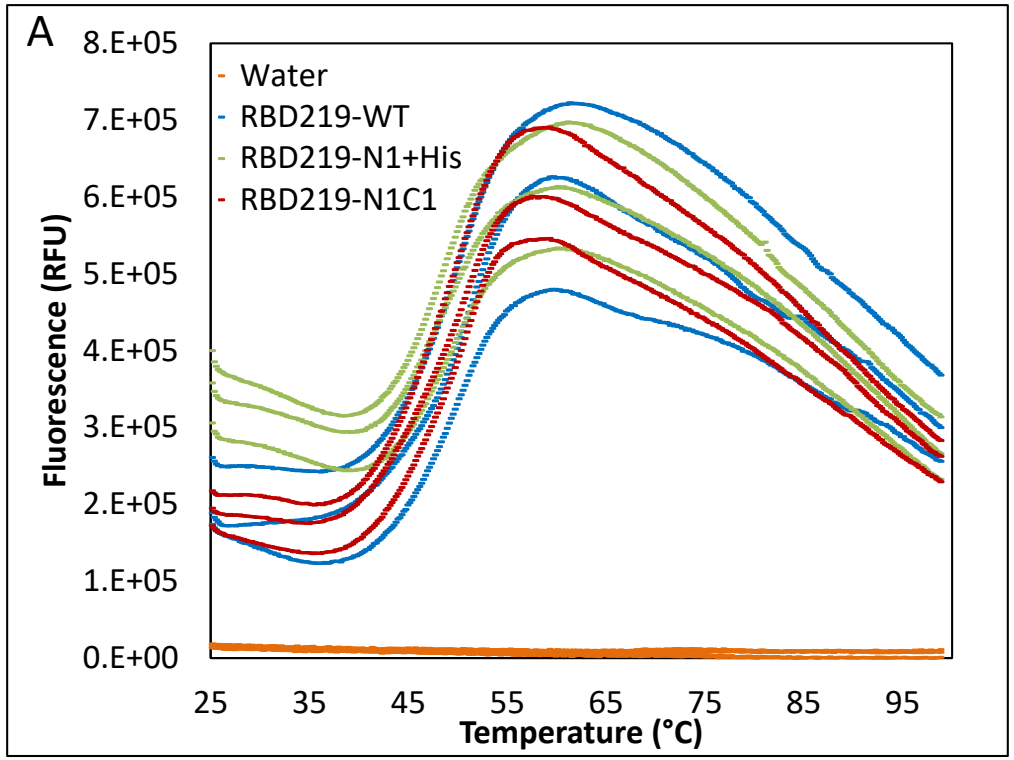

A
Molar ellipticity [θ]
Wavelength (nm)
RBD219-N1C1
RBD219-N1C1
RBD219-N1C1
RBD219-WT
RBD219-WT
RBD219-WT
RBD219-N1+His
RBD219-N1+His
RBD219-N1+His
B
RBD219-WT
Wavelength (nm)
Temperature (°C)
CD [mdeg]
C
RBD219-N1+His
Wavelength (nm)
Temperature (°C)
CD [mdeg]
D
RBD219-N1C1
Wavelength (nm)
Temperature (°C)
CD [mdeg]
E
RBD219-WT
CD
Derivative
CD (mdeg)
dCD/dT (mdeg/°C)
Temperature (°C)
F
RBD219-N1+His
CD
Derivative
CD (mdeg)
dCD/dT (mdeg/°C)
Temperature (°C)
G
RBD219-N1C1
CD
Derivative
CD (mdeg)
dCD/dT (mdeg/°C)
Temperature (°C)

Figure 6


A
8.E+05
7.E+05
6.E+05
5.E+05
4.E+05
3.E+05
2.E+05
1.E+05
0.E+00
Fluorescence (RFU)
Water
RBD219-WT
RBD219-N1+His
RBD219-N1C1
25
35
45
55
65
75
85
95
Temperature (°C)

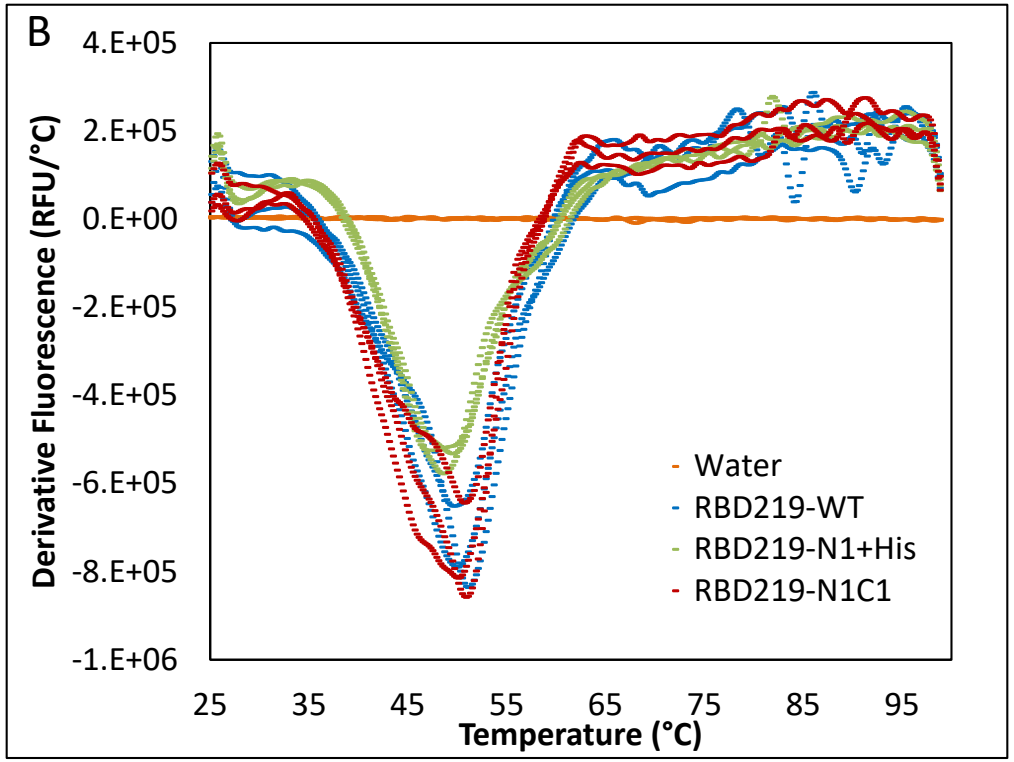
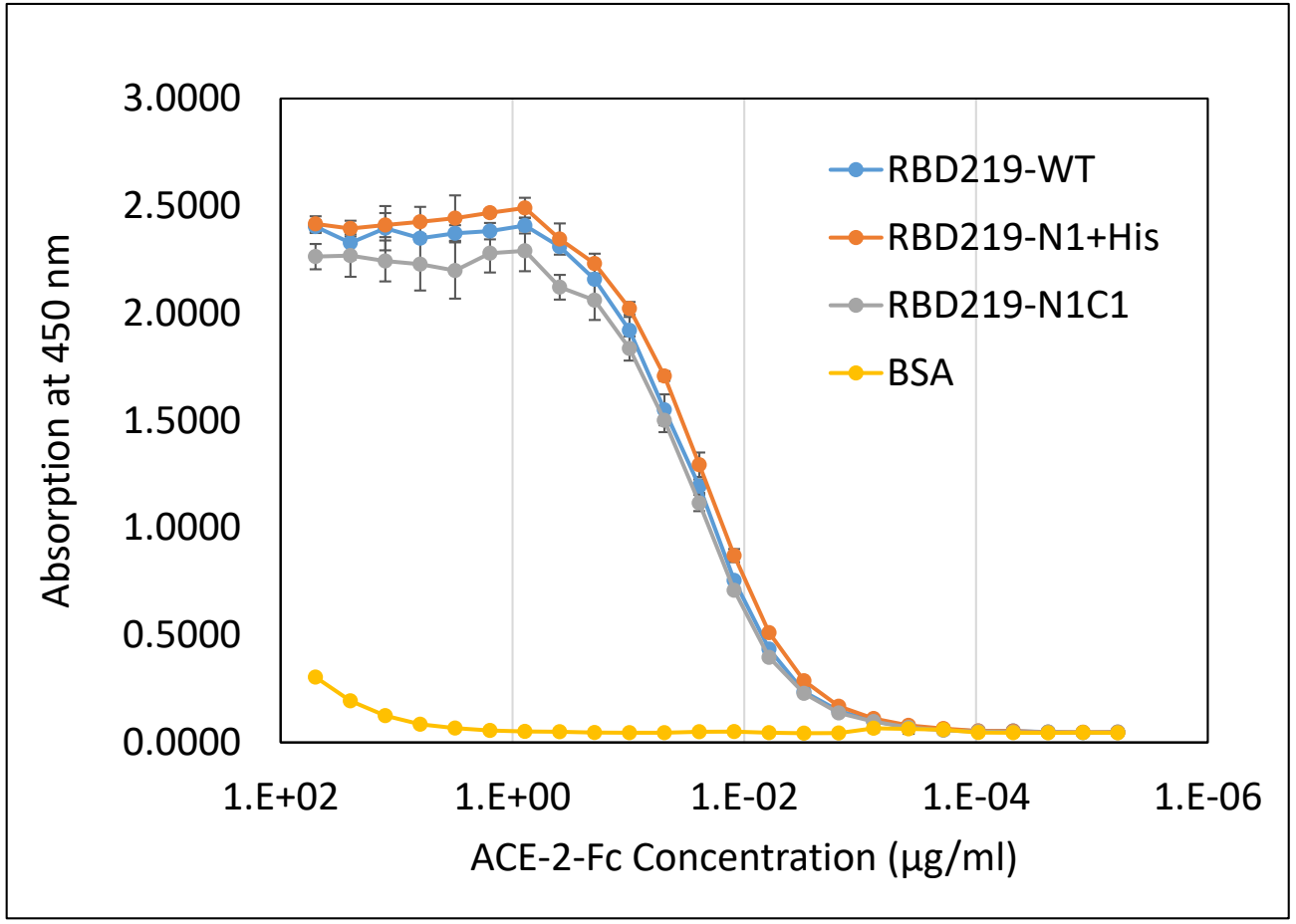

B
4.E+05
2.E+05
0.E+00
-2.E+05
-4.E+05
-6.E+05
-8.E+05
-1.E+06
Derivative Fluorescence (RFU/°C)
Water
RBD219-WT
RBD219-N1+His
RBD219-N1C1
25
35
45
55
65
75
85
95
Temperature (°C)

Figure 7


3.0000
2.5000
2.0000
1.5000
1.0000
0.5000
0.0000
Absorption at 450 nm
1.E+02
1.E+00
1.E-02
1.E-04
1.E-06
ACE-2-Fc Concentration (µg/ml)
RBD219-WT
RBD219-N1+His
RBD219-N1C1
BSA