The nuts and bolts of SARS-CoV-2 Spike Receptor Binding Domain heterologous expression

Mariano Maffei[1*], Linda Celeste Montemiglio[4], Grazia Vitagliano[2], Luigi Fedele[2†], Shaila Sellathurai[2], Federica Bucci[2], Mirco Compagnone[3], Valerio Chiarini[2], Cécile Exertier[5], Alessia Muzi[2], Giuseppe Roscilli[1,2], Beatrice Vallone[5], Emanuele Marra[1,2*]

1. Evvivax Biotech, Rome, Italy

2. Takis Biotech, Rome, Italy

3. Neomatrix Biotech, Rome, Italy

4. Institute of Molecular Biology and Pathology (IBPM), National Research Council, Rome, Italy

5. Department of Biochemical Sciences "A. Rossi Fanelli" University of Rome, Sapienza, Rome, Italy

†*Present address: Luigi Fedele, InnovaVector s.r.l., Pozzuoli, Italy*

***Correspondence to:**

maffei@takisbiotech.it

marra@takisbiotech.it

Abstract

COVID-19 is a highly infectious disease caused by a newly emerged coronavirus (SARS-CoV-2) that has rapidly progressed into a pandemic. This unprecedent emergency has stressed the significance of developing effective therapeutics to fight current and future outbreaks. The receptor-binding domain (RBD) of the SARS-CoV-2 surface Spike protein is the main target for vaccines and represents a helpful “tool” to produce neutralizing antibodies or diagnostic kits. In this work, we provide a detailed characterization of the native RBD produced in three major model systems: *Escherichia coli*, insect and HEK-293 cells. Circular dichroism, gel filtration chromatography and thermal denaturation experiments indicated that recombinant SARS-CoV-2 RBD proteins are stable and correctly folded. In addition, their functionality and receptor-binding ability were further evaluated through ELISA, flow cytometry assays and bio-layer interferometry.

Keywords

SARS-CoV-2; Receptor-Binding Domain; COVID-19; Spike protein; heterologous expression; protein production.

Introduction

At the end of 2019, a novel respiratory pathogen responsible for the COVID-19 disease, namely severe acute respiratory syndrome-related coronavirus (SARS-CoV-2), emerged in Wuhan, China[1]. Only three months later, the virus spread worldwide causing one of largest outbreak of the century that rapidly progressed into pandemic with more than 222 million of confirmed cases and 4,59 million deaths by September 9th [2]. In response to this exceptional situation, an enormous effort has been made by the scientific community to study and characterize the pathogen and to quickly develop safe and effective prophylactic and therapeutic drugs.

The SARS-CoV-2 is an enveloped virus whose surface is decorated with an integral membrane protein (M), an envelope protein (E), a surface spike protein (S), and an additional unexposed structural nucleocapsid protein (N)[3,4]. Among those, the Spike protein is critical to recognize the host-cell receptors and for mediating viral entry, therefore it represents the most studied viral component and the best candidate for drug target [5,6]. The 140 kDa SARS-CoV-2 S protein is organized into two major subunits (S1 and S2) connected by a furin-cleavage site[7]. The S1 subunit contains the receptor-binding domain (RBD; aa 319-541), a 25 kDa domain that is directly involved in the interaction with the Angiotensin-converting enzyme 2 (ACE2)[8,9]. RBD contains nine cysteines, including eight that form disulfide-bridges involved in the RBD fold. In addition, the domain displays two N-glycosylation sites (Asn_{331} and Asn_{343}) known to participate to folding, stability, and function[10–12]. Mutations occurring within this domain are constantly monitored to predict the emergence of novel variants that could be naturally selected and quickly spread, such as the recent isolated alpha (B.1.1.7), beta (B.1.351), gamma (P.1), and delta (B.1.617.2) variants of concern[13–16].

RBD as isolated protein is broadly used in different types of clinical and medical applications (serological tests, vaccine formulation etc.)[17–20] and therefore its *in-vitro* production is of paramount importance. Mammalian and insect cells are the model systems of election used for the heterologous expression of SARS-CoV-2 RBD due to its intrinsic structural complexity. Attempts have also been made using other systems such as *Pichia pastoris*[21] or *Nicotiana benthamiana*[22]. Although *Escherichia coli* (*E. coli*) represents the most common organism employed for the expression of recombinant proteins, its usage is not recommended for challenging targets that require complex folding and/or post-translational modifications such as RBD. Nevertheless, *E. coli* gathers many technical and practical advantages (low-cost, easy handling) compared to other model systems that could be beneficial both for research-scale and large industrial production[23].

In this study, we present a structural and functional comparison of the native RBD of SARS-CoV-2, recombinantly produced in the three major and most frequently used expression systems (*E. coli*, insect and mammalian HEK-293 cells). The characterization of recombinant RBD proteins is of the utmost relevance in drug design to tackle the Covid-19 pandemic.

Abbreviations

ACE2, angiotensin-converting enzyme 2; BLI, bio-layer interferometry; CD, circular dichroism; COVID-19, coronavirus disease-2019; ELISA, enzyme-linked immunosorbent assay; FACS, fluorescence-activated cell sorting; HEK, human embryonic kidney; IgG, immunoglobulin G; IMAC, immobilized metal affinity chromatography; MALDI-MS, matrix-assisted laser desorption/ionization- mass spectrometry; MFI, mean fluorescence intensity/ geometric mean; MS, mass spectrometry; MW, molecular weight; MWCO, molecular weight cut-off; OD, optical density; PMF, peptide mass fingerprint; PMTV, photomultiplier voltage;

POI, protein of interest; RBD, receptor binding domain; *E. coli*-RBD, receptor binding domain produced in *E. coli*; HEK-293-RBD, receptor binding domain produced in HEK293 cells; insect-RBD, receptor binding domain produced in insect cells; RP-UHPLC-MS, reversed-phase - ultra-high performance liquid chromatography - mass spectrometry; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; SEC, size exclusion chromatography; Tm, melting temperature; UHPLC, ultra-high performance liquid chromatography; WHO, World Health Organization;

Results

Design, expression, and purification of SARS-CoV-2 RBD in E. coli

The RBD protein (Figure 1(a)) was recombinantly expressed with a C-terminal 6x-His purification tag both in BL-21 *Star* and Lemo21 cells. The Lemo21 bacterial strain allows to express challenging targets such as toxic, highly insoluble, and membrane proteins by reducing inclusion body formation and potential inhibitory effects on cells growth, thus resulting in an increased level of properly folded products. However, only negligible amount of RBD was found in the soluble fraction, even exploring alternative growing conditions including lower temperature, distinct induction times, and increasing concentrations of L-Rhamnose (data not shown). The target protein was totally recovered from inclusion bodies with yields representing 5.2% (*Star*) and 8.1% (Lemo21) of the total protein extract (Figure 1(b), Suppl. Figure 1 (a)). Protein purification was carried out in the presence of denaturing agents (6 M urea) followed by a slow refolding-process through an over-night dialysis against buffer containing the redox pair of oxidized and reduced glutathione to induce proper disulfide bond formation (Figure 1(b), Suppl. Figure 1(b)). As shown in Figure 1(c), the purified *E. coli*-RBD protein shows a high degree of purity (>90%) and migrated as a single-smeared band at the expected height on 4-12% SDS-PAGE (theoretical mass, 26052 Da).

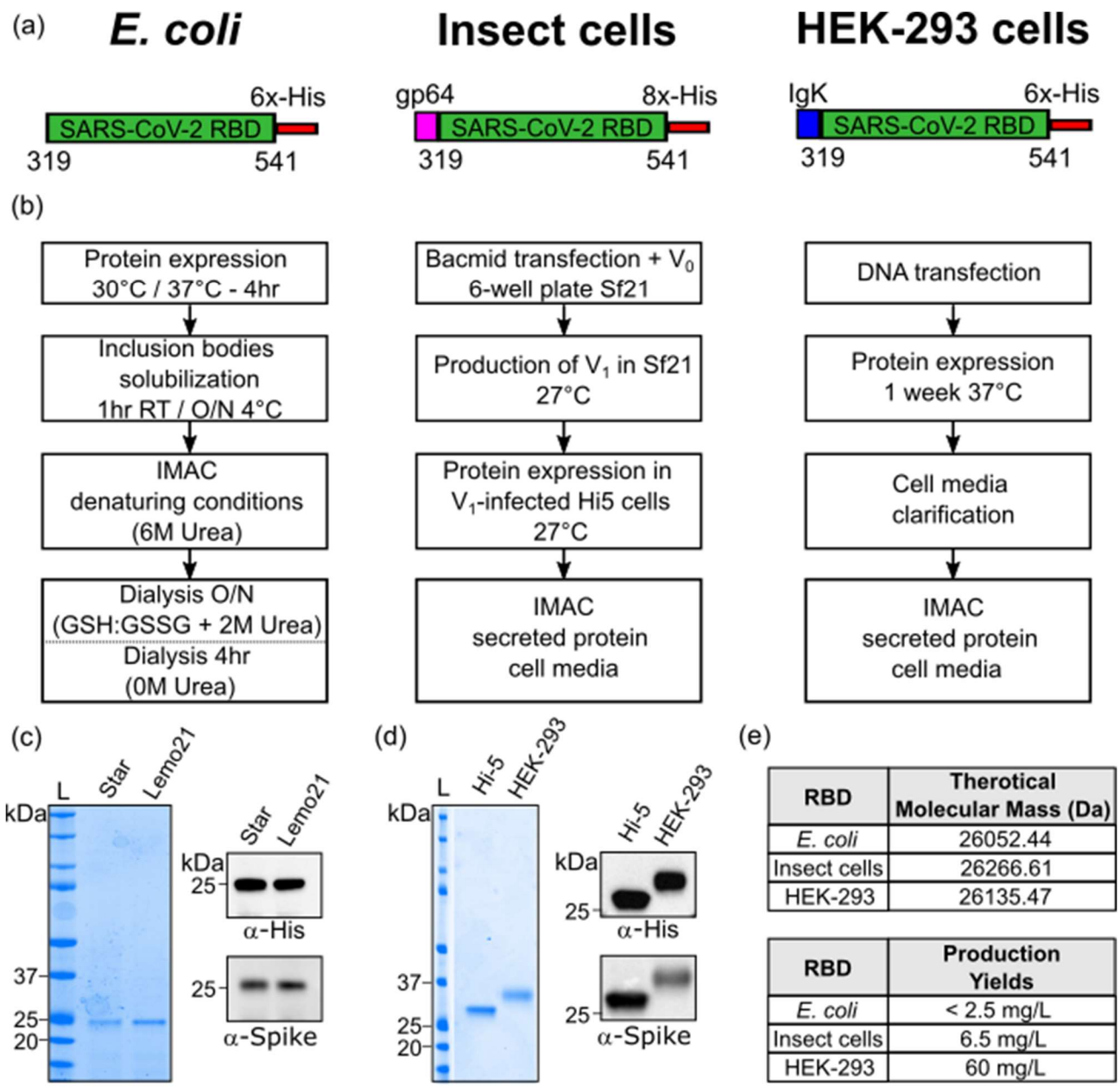
Moreover, Western blot analysis indicated that the protein was efficiently recognized by anti-His and anti-SARS-CoV-2 Spike S1 subunit antibodies (Figure 1(c)). Approximately 1.25 mg of purified RBD was obtained starting from 0.5 L of bacterial culture (final yield ~2.5 mg/L). Among distinct batches, concentration ranged from 0.1 mg/mL (3.8 μM) to 0.3 mg/mL (11.5 μM). Concentrations higher than 0.3 mg/mL led to protein precipitation. Finally, the molecular weight and primary aminoacidic sequence of SARS-CoV-2 RBD purified from *E. coli* were further validated by Mass-spec analysis (Suppl. Figure 1(c),(d)).

Design, expression, and purification of SARS-CoV-2 RBD in insect and mammalian cells

RBD fragment with a C-terminal 8x-His purification tag and was cloned downstream of the gp64 for expression in insect cells (Figure 1(a)). Generation and amplification of recombinant baculovirus were carried out in Sf21 cells, while protein expression was performed in Hi-5 infected insect cells (Figure 1(b)). The soluble protein of interest (POI) was secreted into culture media and purified through Immobilized Metal Affinity Chromatography (IMAC) using a Ni-Nta resin (Figure 1(b), Suppl. Figure 2(a)). Isolated insect-RBD migrated as a single band slightly higher than 25 kDa on 4-12% SDS-PAGE (theoretical mass, 26266 Da) exhibiting a high level of purity (>95%), and it was clearly detected by immunoblotting (anti-His and anti- SARS-CoV-2 Spike S1 subunit) (Figure 1(d)). At a laboratory-scale, final yields were around 6.5 mg RBD per Liter of insect cells with batch concentrations ranging from 0.25 mg/mL (9.5 μM) to 0.5 mg/mL (19 μM). The experimental mass (28936 Da) of the recombinant insect-RBD determined by MALDI mass-spectrometry analysis (Suppl. Figure 2(b),(c)) was higher than the theoretical one based on the amino acid composition, thus suggesting the presence of glycosylations[24].

Regarding RBD expression in mammalian cells, SARS-CoV-2 RBD flanked by a C-terminal 6x-His tag was cloned downstream the Ig Kappa chain-signal peptide responsible

for protein secretion (Figure 1(a)). Cells were transfected with DNA and left under stirring and controlled CO_2 atmosphere for 1 week expressing the POI. The RBD-containing medium was filtrated and the POI was purified by affinity chromatography (Figure 1(b), Suppl. Figure 2(d)). The eluted protein migrates as a single-slightly diffuse band below 37 kDa, indicating that RBD (theoretical mass, 26135 Da) contains glycosylations. Indeed, experimental mass obtained from MALDI-MS analysis was 31453 Da, confirming the presence of post-translation modifications as previously reported[21,25,26] (Suppl. Figure 2(e),(f)). Additionally, the eluted protein was efficiently detected by anti-His and anti-S1 subunit of SARS-CoV-2 Spike antibodies (Figure 1(d)). Around 800 mL of transfected cells yielded 58.8 mg of pure purified protein, with batch concentrations reaching up to 1.8 mg/mL (69 μM).



RBD	Therotical Molecular Mass (Da)
E. coli	26052.44
Insect cells	26266.61
HEK-293	26135.47

RBD	Production Yields
E. coli	< 2.5 mg/L
Insect cells	6.5 mg/L
HEK-293	60 mg/L

Figure 1. SARS-CoV-2 RBD production in *E. coli*, insect and mammalian cells. (a) Schematic representation of the RBD protein constructs expressed in *E. coli* (left), insect cells (middle), and mammalian HEK-293 cells (right). (b) Diagram summarizing the RBD recombinant expression from *E. coli* (left), insect cells (middle), and mammalian HEK-293 cells (right) and the subsequent purification. (c) SDS-PAGE (left panel) and Western blot analysis (right panels) of *E. coli*-purified RBD protein. (d) SDS-PAGE (left panel) and Western blot analysis (right panels) of RBD fragment produced in Hi-5 insect cells and mammalian HEK-293. L = molecular weight ladder (e) Theoretical molecular masses

157 calculated according to RBD amino acid composition (above) and RBD production yields
158 (below).

159

160 *Biochemical characterization of RBD*

161 RBD produced in HEK-293, insect cells and *E. coli* were analyzed by size exclusion
162 chromatography (SEC) (Figure 2(a)). *E. coli*-RBD elutes as a single and narrow peak
163 centered at 18.7 mL whereas RBD proteins produced in HEK-293 and insect cells display
164 elution peaks shifted to lower retention volumes owing to the presence of glycosylations. In
165 fact, HEK-293-RBD elutes as a main peak centered at 15.3 mL, while RBD produced in
166 insect cells elutes as a major one centered at 16 mL and a minor one at 14.6 mL. The
167 presence of two peaks in the insect-RBD elution profile suggests the existence of at least
168 two populations of the protein showing alternative glycosylation patterns and differing from
169 the one of HEK-293-RBD. The lack of glycosylation of *E. coli*-RBD shifts the retention
170 volumes to higher values. Altogether, all the elution peaks observed are all consistent with
171 a ~30 kDa protein.

172 RBD proteins were analyzed by far-UV CD spectroscopy. The spectral profiles of
173 HEK-293-RBD and insect-RBD reported in Figure 2(b) are by-an-large identical, both
174 displaying a single minimum at ~206 nm and a maximum at ~230 nm. Conversely, the far-
175 UV CD spectrum of *E. coli*-RBD differs from those of the eukaryotic counterparts, as also
176 observed by Mycroft-West et al[27]. However, the analysis of the secondary structure
177 composition returned an overall similar distribution (Figure 2(c)).

178 The conformational stability of RBDs was investigated by means of temperature
179 denaturation experiments: we followed the variations in far-UV CD ellipticity at 222 nm upon
180 increase of temperature from 290 K to 360 K (Figure 2(d)). The change in ellipticity,
181 monitored for each construct, followed a sigmoidal dependence upon temperature increase,

suggesting that the RBDs reversibly unfold. Differently to what expected for a typical folded-to-unfolded transition followed by far-UV CD, the ellipticity values decrease with raising temperature (no loss of CD signal).

The observed denaturation curves could be well fitted to a 2-state transition according to equation (1). The resulting T_m values are identical within experimental error, yielding a mean value of 323.1 ± 0.6 K for HEK-293-RBD, 323.1 ± 1.4 K for insect-RBD, and 322.7 ± 1.3 K for *E. coli*-RBD over three independent experiments. The estimated T_m values are consistent to what previously reported for RBD produced in eukaryotic cells in similar ionic strength conditions[21]. It is worth mentioning that the denaturation curve of *E. coli*-RBD shows a biphasic behaviour, indicating the existence of an initial unfolding event preceding the main one and taking place around 305 K (data not shown). We surmise that this initial phase is likely due to a minor portion of the protein that failed refolding during sample purification from inclusion bodies, probably owing to the absence of glycosylations.

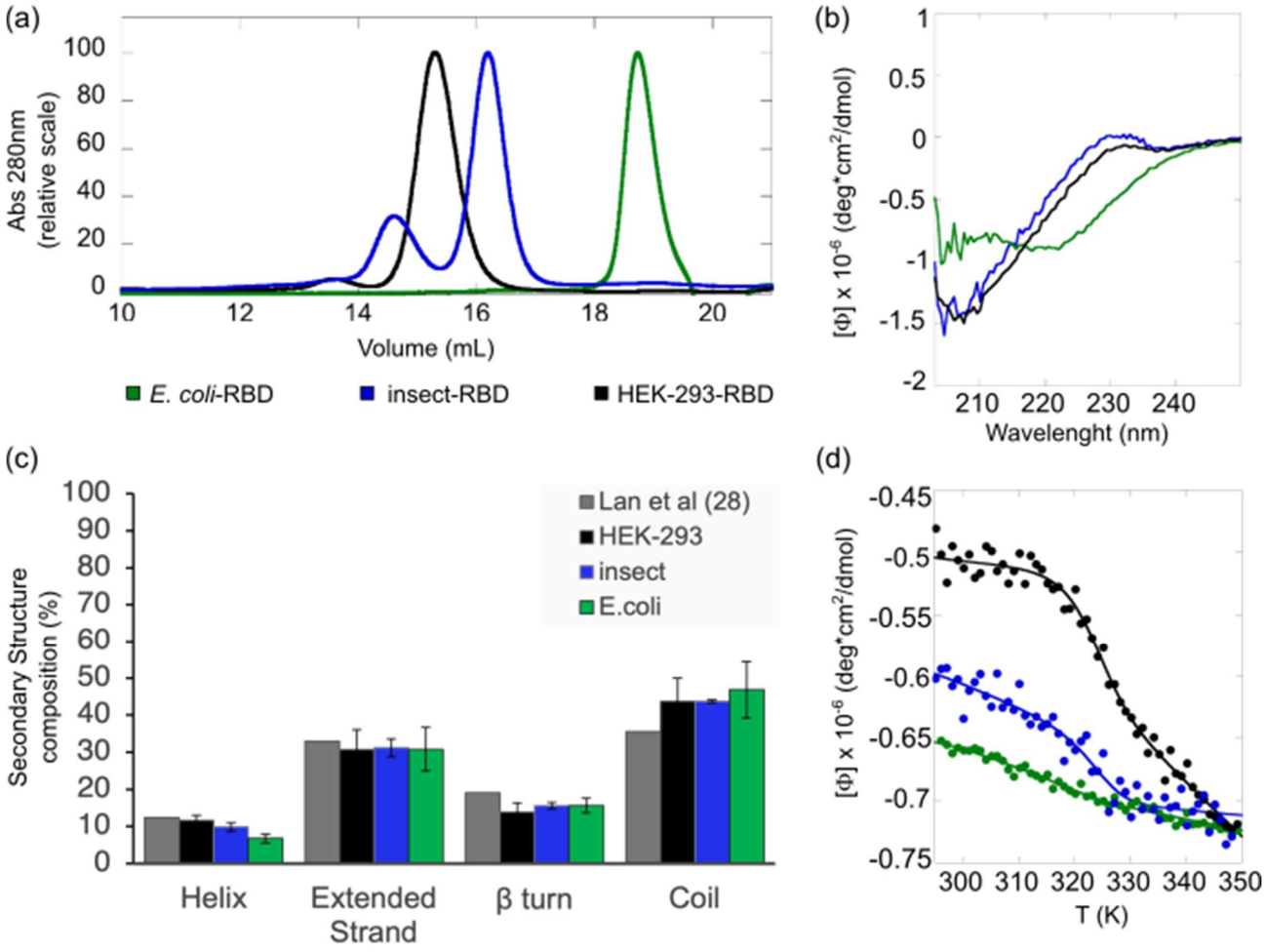


Figure 2. Biochemical characterization of recombinant RBD. (a) Gel filtration chromatographic profiles. Protein separation was performed at room temperature using a Superdex 200 Increase 10/300 GL with 40 µg of RBD produced in HEK-293 (black), 47 µg of RBD produced in insect (blue), 40 µg of RBD produced in *E. coli* (green), each in 50 mM Tris·HCl, 150 mM NaCl, pH 8.3. (b) Far-UV CD spectra of RBD produced in HEK-293 (black), insect (blue) and *E. coli* (green) cells. All spectra were collected at 20°C, using a 0.1 cm path length quartz cuvette. (c) The histogram reports the distribution of the secondary structure content determined for the RBD proteins (at least three independent CD experiments (mean ± standard deviation)), in comparison with the secondary structure composition of RBD reported by Lan et al. (dark grey bars)[28]. (d) Thermal denaturation profiles of RBD E. coli (green), insect (blue), and HEK-293 (black) continuously monitored by far-UV CD at 222 nm over the range 290–370 K. Data were fitted using a two-state model.

209

210 *ELISA assays*

211 The functionality of the RBD protein was determined through ELISA assays using
212 plates coated with RBD produced either in *E. coli*, insect or in HEK-293 cells. First, to test
213 coating conditions, 50 ng/well of RBD proteins were used for coating in 50 µL of phosphate-
214 buffered saline (PBS) or carbonate buffer. Serially diluted (1:1000, 1:10000, 1:50000) rat
215 sera (immunized with COVID-*e*Vax vaccine[29]) were used to detect the optical density (OD)
216 associated with antibody-RBD interaction under distinct conditions. Significant differences
217 were observed between plates, suggesting that PBS buffer is the most efficient buffer for
218 coating (data not shown).

219 Subsequently, plates were coated with increasing RBD protein concentrations
220 (ranging between 1 and 5 µg/mL) and serum from rats previously immunized with COVID-
221 *e*Vax[29] vaccine was applied to each plate for RBD protein binding. Of note, independently
222 from protein concentration (1, 3 and 5 µg/mL) both insect-RBD and HEK-293-RBD were
223 efficiently recognized by rat IgG, whereas rat IgG-*E. coli*-RBD interaction was much lower
224 (Figure 3(a)). The observed differences between RBD produced in *E. coli* and in insect or
225 mammalian counterparts are probably due to the major affinity of the latter ones to the IgG
226 produced in rats.

227 We also monitored the interaction between RBD and a commercial antibody against
228 the S1 subunit of SARS-CoV-2 Spike. As shown in Figure 3(b), the observed OD signal of
229 insect-RBD was not markedly different from that of HEK-293-RBD although at
230 concentrations < 1 µg/mL, RBD from insect showed a slightly higher binding ability
231 compared to HEK-293-RBD (Suppl. Figure 3(a),(b)). In contrast, *E. coli*-RBD showed a lower
232 binding compared to HEK-293 and insect RBDs. We hypothesized that this lower binding

ability of RBD produced in *E. coli* may, again, be due to the presence of a sub-population of the protein that failed refolding.

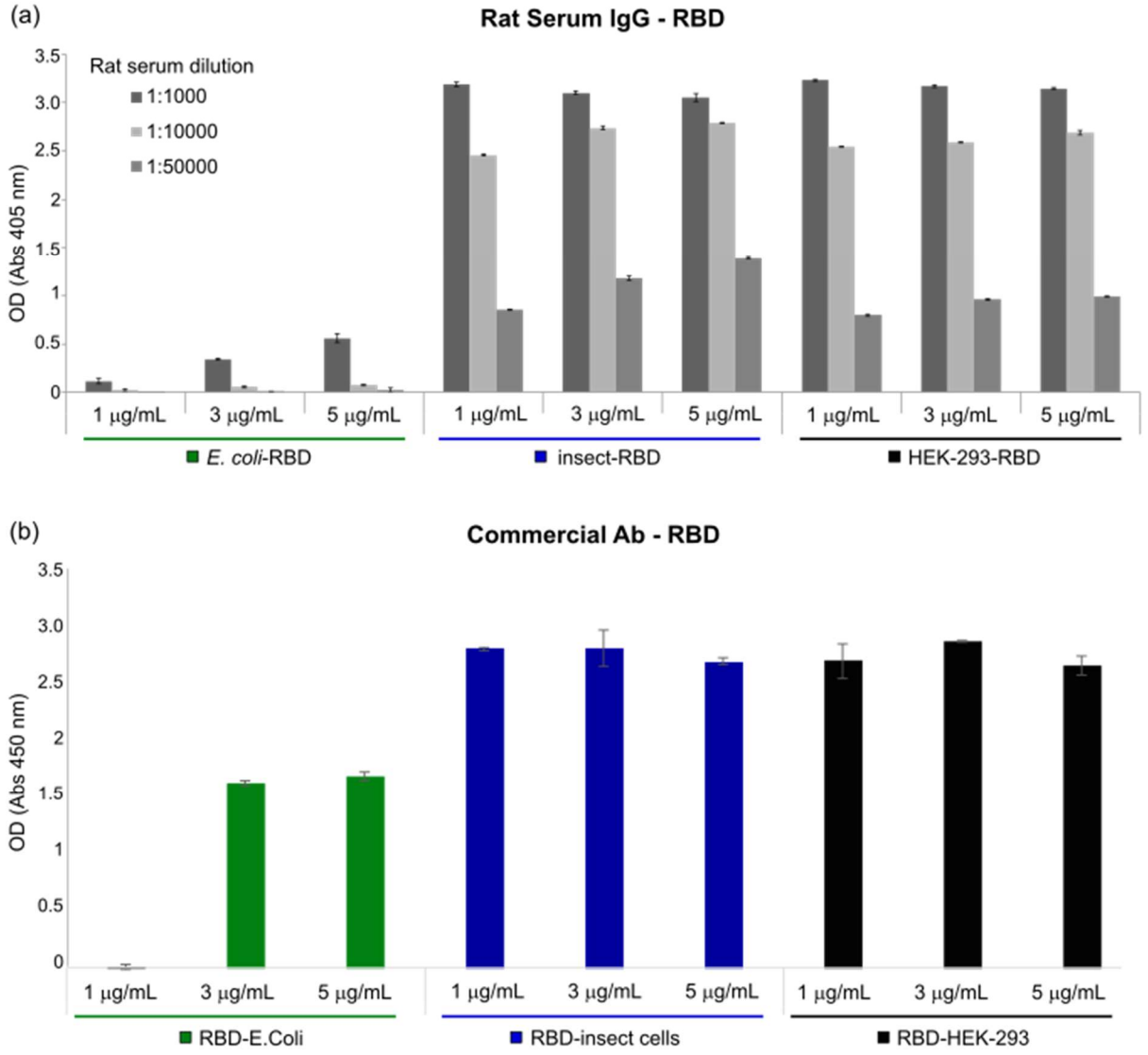
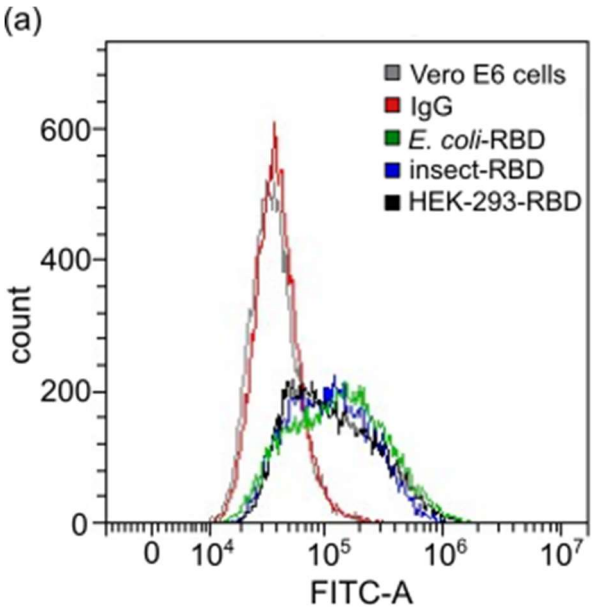


Figure 3. Indirect ELISA assays (a) Serum from immunized rat with COVID-*e*Vax was used to compare different concentrations (1, 3, 5 µg/mL) of RBD expressed in *E. coli* (green), insect (blue) and HEK-293 cells (black). The y-axis represents the optical density (OD) measured at 405 nm while the x-axis accounts for RBD concentrations and serum dilution factors (1:1000, 1:10000, 1:50000). Bars indicate standard deviations. (b) Commercial antibody against the S1 subunit of SARS-CoV-2 Spike was used to compare different concentrations (1, 3, 5 µg/mL) of RBD produced in *E. coli* (green), insect (blue) and HEK-

293 cells (black). Optical density (OD) was measured at 450 nm and bars indicate standard deviations.

Flow cytometry assay

The receptor binding ability and functionality of RBDs produced in the three presented model systems were further investigated through flow cytometry. Vero E6 cells have been shown to express ACE2 receptor on their apical membrane and to be susceptible to SARS-CoV-2 infection[30,31]. Thus, we tested RBD-ACE2 binding by incubating recombinant RBD proteins with cultured Vero E6 cells. Figure 4 shows that all the three studied RBDs were able to efficiently bind Vero E6 cells, while no signal was observed when cells were incubated only with antibodies (Suppl. Figure 4). This result suggests that recombinant RBD proteins are efficient in recognizing ACE2.

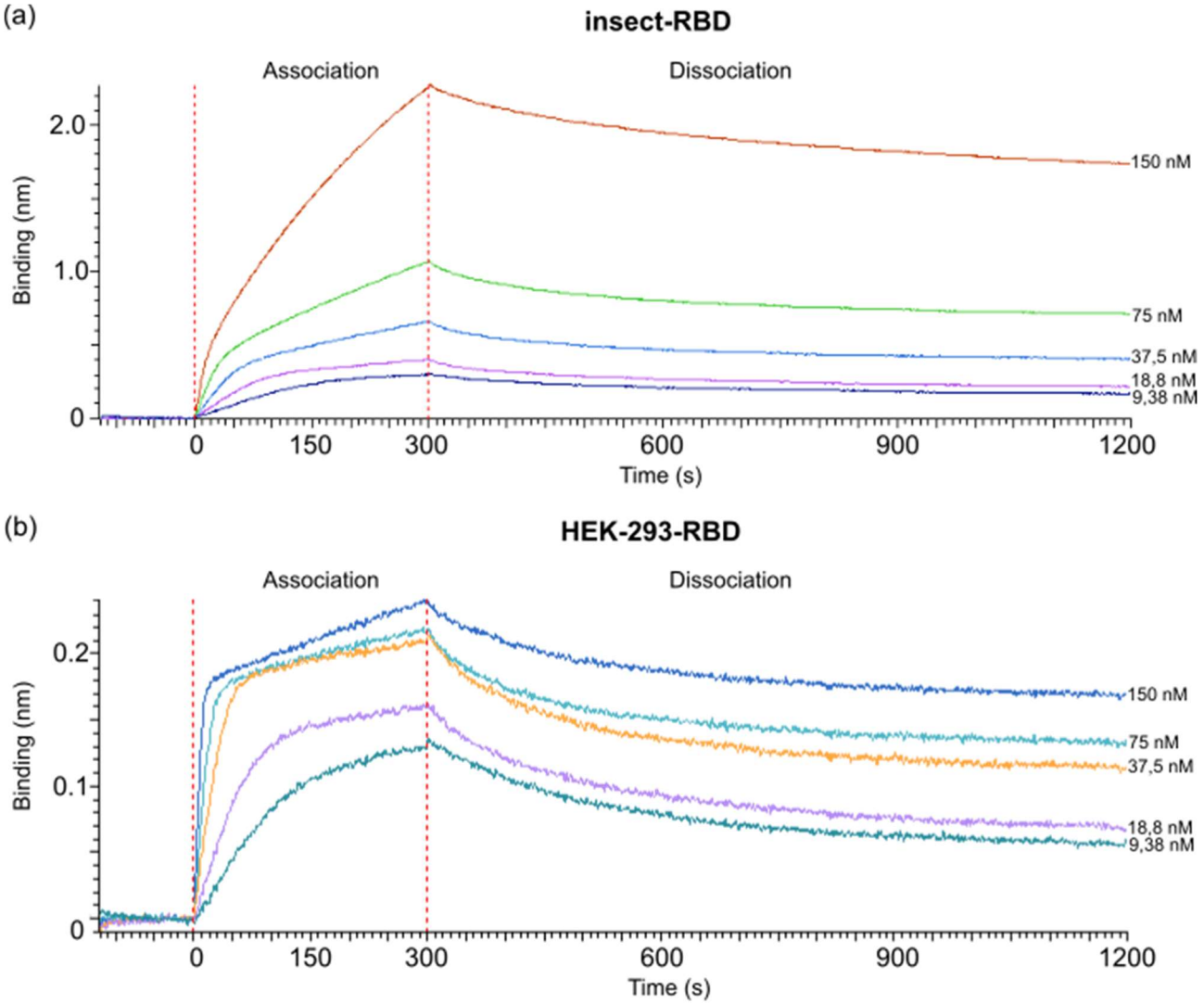


	Geometric Mean (MFI)
Vero E6	37617.1
IgG	40232
E. coli	115218
Insect	109547.3
HEK-293	115657.9

Figure 4. Flow cytometry assays. RBD-Vero E6 cells binding experiment. (a) Gray curve: Vero E6 cells alone; Red curve: Vero E6 cells incubated with only secondary antibody; Green curve: Vero E6 cells incubated with *E. coli*-RBD; Blue curve: Vero E6 cells incubated with Insect-RBD; Black Curve: Vero E6 cells incubated with HEK-293-RBD. Incubation with RBD was followed by anti-RBD primary antibody and secondary antibody. Table represents the intensity of the staining measured as geometric mean (median fluorescence intensity, MFI) value.

Bio-Layer interferometry binding assay

Finally, the binding affinity to ACE-2 receptor of the RBD produced in *E. coli*, insect and HEK-293 cells was evaluated using bio-layer interferometry (BLI). ACE2-hFc fusion protein was immobilized onto anti-human Fc biosensor and different concentrations of RBD proteins (range 150 nM – 9.8 nM) were tested to obtain association curves. After fitting, the dissociation constant (K_d) of ACE2-hFc to insect-RBD and to HEK-293-RBD was determined to be $7.49 \cdot 10^{-9}$ M and $5.34 \cdot 10^{-10}$ M, respectively, while much lower binding affinity was observed for *E. coli*-RBD ($K_d = 1.21 \cdot 10^{-6}$ M) (Figure 5(a),(b) and Suppl. Figure 5).



Model System	K_d (M)
Insect cells	$7{,}49 \times 10^{-9}$
HEK-293	$5{,}34 \times 10^{-10}$

Figure 5. BLI measurements. (a) BLI profiles accounting for the binding of insect-RBD and (b) HEK-293-RBD to ACE2-hFc. After a baseline, the sensorgram starts with the association (0 - 300 s) of RBD to the ACE2 loaded sensor, followed by the dissociation phase (900 s).

Discussion

The emergence of the novel SARS-CoV-2 pathogen at the end of 2019, which has quickly degenerated into a pandemic, has underlined the importance of immediate and responsive actions from the local governments, health authorities, and the world scientific community in order to tackle this situation that probably represents the biggest challenge that modern society has faced. As a result, during the last two years, several vaccines have been developed and many drugs are currently under screening or evaluation in clinical trials[32–34]. Most of those therapeutics targets the Spike protein and more specifically its receptor-binding domain, which is exposed on the viral envelope and that is directly involved in receptor-binding and cell-entry. Moreover, both full-length Spike and RBD are widely used as viral antigens for diagnostic tests representing a critical tool for a fast response to the pandemic.

In this study, we provide technical insights into the heterologous expression, purification, and characterization of the native SARS-CoV-2 RBD produced in *E. coli*, insect and HEK-293 model systems. Bacterial RBD production was achieved by recovering the protein from the insoluble fraction and through a careful process of refolding. Efforts to increase its solubility by using Lemo21 *E. coli* strain failed in agreement with previous attempts to produce this protein, or its ancestor (SARS-CoV RBD), in *E. coli* in its native soluble form[21,35,36]. After refolding, isolated *E. coli*-RBD showed a good degree of purity and was efficiently recognized by commercial antibodies. Considering the challenging target, final obtained yields were not high, but enough to carry out most of the lab-scale downstream applications. In contrast, production in insect and HEK-293 cells resulted in more soluble, highly glycosylated RBD proteins with yields up to 60 mg/L. The presence of post-translational modifications (glycosylations) in the latter samples was indirectly observed by SDS-PAGE, mass-spectrometry analysis, and size-exclusion chromatography. The nature

and the type of glycosylation was not further investigated. Of note, although the far-UV-CD spectrum profile of the RBD from *E. coli* appears different from the one observed for HEK-293-RBD and insect, the overall distribution of the secondary structure composition and the measured T_m were similar among the three samples, suggesting that most of the RBD expressed in bacteria recovers the proper fold, homogeneity, and conformational stability even though lacking glycosylations.

RBD functionality was demonstrated *in-vitro* using Enzyme-linked immunosorbent assay (ELISA). RBD produced in *E. coli* displayed a weak binding affinity to IgG produced in rats and to commercially available antibodies while efficient response was observed for both insect and mammalian derived RBDs. Remarkably, at lower concentration (< 1 µg/mL), insect-RBD gave a slightly better signal than HEK-293-RBD. We also investigated the capability of isolated RBDs to bind the ACE2 receptor. All the RBDs produced in this work efficiently bind to Vero E6 cells as confirmed by FACS assay. ACE2-RBD binding was further confirmed and quantified by bio-layer interferometry, with the bacterial-RBD displaying again the lowest binding efficiency. Our data suggest that the absence of glycosylations could partially affect ACE-2 binding *in-vitro* as also previously observed[10]. In addition to this, we must consider that the presence of a sub-population of protein that failed refolding, as indicated by circular dichroism and ELISA assays, might also contribute to the observed lower binding efficiency of the bacterial RBD.

To summarize, this work offers a technical and practical overview of RBD production using the three most widely used expression systems, highlighting the main advantages and drawbacks, reported in Table 1. RBD obtained from both eukaryotic systems resulted in a high-quality final bioproduct potentially eligible for diverse downstream applications (vaccine design, diagnostic kits, drug screening etc.). However, the high costs (resources), the time-consuming production, the requirement of specific equipment, and access to dedicated

facilities could be a limitation for many laboratories or for the industrial production. By contrast, the bacterial-derived RBD offers a low production cost, a broader availability, and easy handling as main advantages, which make it more accessible. However, limitations in the quality of the produced sample include the absence of glycosylation that partially affects protein stability and efficiency, the presence of heterogeneous folded populations, and the relative low production yields which may result in a final product not eligible for some clinical and medical applications. Overall, all the recombinantly produced RBDs represent valuable tools for research purposes against the pandemic. Recently, expression and purification strategies described in this article have been also proved to be successful in the production of mutants of RBD corresponding to the variants of concern.

Model System	Cost	Production time	Yields	Glycosylation	Folding	ELISA	FACS	ACE-2 receptor binding
E. Coli	Low	≤ 1 week	+	No	+/-	+	++	+
Insect	Medium	2-3 weeks	++	Yes ++	+++	+++	+++	+++
HEK-293	High	2-3 weeks	+++	Yes +++	+++	+++	+++	+++

Table 1. Overview of the main aspects of RBD produced in the three major model systems.

Material and Methods

RBD protein production in E. coli

The SARS-CoV-2 Spike Receptor Binding Domain sequence (aa 319-541, Uniprot ID P0DTC2) was cloned with a C-terminal 6x-His tag into a pET-21a(+) plasmid. *E. coli* BL21 Star™ (DE3) (genotype: F⁻*ompT hsdS*$_B$ (r_B^-, m_B^-) *galdcmrne*131) competent cells, and *E. coli* Lemo21 (DE3) (genotype: *fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS/ pLemo*(Cam^R)) competent cells were transformed with 100 ng of plasmid of interest. A single colony was incubated in 15 mL starter culture (LB) with Ampicillin (BL21 Star strain) or

Ampicillin/Chloramphenicol (Lemo21 strain), grown at 37°C on agitation over-night. Starter culture was successively inoculated in 500 mL (LB) with antibiotics incubated at 37°C until mid-log phase (OD_{600} = 0.6 to 0.8). Protein expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 30°C or 37°C for 4 hours on agitation. Cells were harvested at 6'000 rpm for 10 minutes, washed once with 50 mM Tris·HCl (pH 8.0) and then further centrifuged. Pellet was resuspended in a solution containing 50 mM Tris·HCl, 500 mM NaCl (pH 8.0) also containing a protease inhibitor cocktail (Roche 11836170001) prior sonication. The suspension was then centrifuged at 11'000 rpm for 45 minutes to separate soluble and insoluble fractions. The pellet containing the RBD target protein was resolubilized in extraction buffer containing 50 mM Tris·HCl, 500 mM NaCl, 20% glycerol, 10 mM β-mercaptoethanol, 8 M urea (pH 8.0). The washed inclusion bodies were shortly sonicated and left 1 h at room temperature (RT) or O/N at 4°C on agitation. Protein was purified using IMAC (His-Trap, Cytiva) under denaturing conditions (Elution buffer: 50 mM Tris·HCl, 500 mM NaCl, 20% Glycerol, 10 mM β-Mercaptoethanol, 6 M urea, 300 mM imidazole pH 8.0). Eluted fractions were analysed by SDS-PAGE and firstly dialyzed over-night against buffer containing 50 mM Tris·HCl, 500 mM NaCl, 20% Glycerol, GSH-GSSG (3 mM : 1 mM), 2 M urea (pH 8.0) with slow agitation. The day after, protein solution was dialyzed against 50 mM Tris·HCl, 500 mM NaCl, 1 mM TCEP (pH 8.0) or PBS 1x (pH 7.4) for 4 hours. The purified *E. coli*-RBD sample protein was quantified by UV-visible spectroscopy, aliquoted, and stored at -80°C.

RBD protein production in insect cells

The SARS-CoV-2 Spike Receptor Binding Domain sequence (aa 319-541, Uniprot ID P0DTC2) was cloned into a pFAST-bac1 plasmid downstream of the gp64 signal sequence to promote secretion, along with a C-terminal 8x-His tag for affinity purification.

100 ng of plasmid was transformed into DH10Bac competent cells (MAX Efficiency™ DH10Bac Competent Cells, Gibco #10361012) for bacmid DNA production. Each bacmid, extracted from 3 mL of an O/N colony culture, was diluted in a final volume of 220 µL of Sf900 III medium and then combined with a mix of 10 µL of XtremeGene (Cellfectin™ II Reagent, Gibco #10362100) in 100 µL Sf900 III medium. This solution was left 15 minutes at room temperature to allow complex formation, according to the manufacturer's protocol. For transfection, the latter solution was added dropwise onto Sf21 cells (Gibco #11497013) previously plated on a 6-well plate at 1.0 x 10^6 cells/well confluency. 60 hours post-transfection, supernatant containing the first generation of recombinant baculovirus (V_0) was harvested and amplified to obtain a high titer of virus. Hi-5 cells (BTI-TN-5B1-4) (Gibco #B85502) were cultured in Express Five™ SFM (Serum-Free Media) medium (Gibco # B85502 Expression Systems) at a cell density of 0.5 x 10^6 cells/mL and infected with recombinant virus. Cells were kept at 27°C and 130 rpm for protein expression. After 72 hours, supernatant containing secreted RBD was collected and subjected to IMAC (His-Trap Excel, Cytiva). RBD was eluted using 50 mM Tris·HCl, 150 mM NaCl, 300 mM imidazole, pH 8.0). Eluted fractions were analysed on 4-12% SDS-PAGE and dialyzed over-night against 50 mM Tris·HCl, 150 mM NaCl, pH 8.0 with slow agitation. Purified RBD protein was quantified by UV-visible spectroscopy, aliquoted, and stored at -80°C.

RBD protein production in HEK-293 cells

C-terminal 6x-His tagged SARS-CoV-2 RBD fragment (aa 319-541, Uniprot ID P0DTC2) was cloned downstream of the Ig Kappa chain-signal peptide for expression as secreted protein in mammalian cells (Expi293). Cells were transfected at a concentration of ~3x10^6/mL with 1 µg of DNA per milliliter of cell culture. Feed enhancers and PEN-STREP were added to the cells after 20 and 24 hours, respectively. Cells were left in agitation at

37°C for 1 week before clarification by centrifugation at 12700 rcf. After filtration, the RBD-containing supernatant was purified by affinity chromatography on a 1 mL INDIGO column (Cube Biotech). Sample was diluted with binding buffer (20 mM NaPi, pH 7.4, 500 mM NaCl) and loaded at 1 mL/min flowrate. Elution was carried out in the same conditions, with a single step of 250 mM imidazole (20 mM NaPi, pH 7.4, 500 mM NaCl, 250 mM Imidazole). Eluted protein was readily dialyzed against DPBS 1x. Isolated RBD protein was quantified by UV-visible spectroscopy, aliquoted, and stored at -80°C.

Mass-spectrometry analysis

SARS-CoV-2 RBD recombinant protein(s) molecular weight and primary aminoacidic sequence were determined by MALDI-MS and by peptide mass fingerprint (PMF), respectively. Determination of the molecular weight was achieved by MALDI mass spectrometry analysis on a MALDI Ultraextreme (Bruker, GmbH) in positive linear mode. 30 µL of sample were desalted by diafiltration using Amicon filters with 3.5 kDa MWCO or by Zip Tip C18 (Millipore) and 2 µL of sample were mixed with a solution of the matrix superDHB. A volume of 2 µL of the resulting solution were deposited on the target plate and left dry in the air.

In order to acquire information on the primary aminoacidic sequence, an aliquot of each sample was reduced, alkylated, digested with trypsin, and analysed by RP-UHPLC-MS/MS. RP-UHPLC-MS analysis was performed on a Q-Exactive HF-X (ThermoFisher Scientific) mass spectrometer coupled with an UHPLC Ultimate 3000 RSLCnano System (ThermoFisher Scientific). A volume of 1 µL of the resulting peptide mixtures were injected on a column EasySpray PepMap RSLC C18 100 Å 2 µm, 75 µm x 15 cm (Thermo Fisher Scientific). The column oven was maintained at 35°C, the analysis was carried using a gradient elution (phase A: 0.1% formic acid in water; phase B: 0.1% formic acid in

acetonitrile). The flow rate was maintained at 300 nL/min. The mass spectra were acquired using a “data dependent scan”, able to acquire both the full mass spectra in high resolution and to “isolate and fragment” the twelve ions with highest intensity present in the mass spectrum. Raw data were analyzed using the Biopharma Finder 2.1 software from ThermoFisher Scientific.

SDS-PAGE and Western blot

Purified proteins (500 ng) were analyzed on 4-12% NuPAGE Bis-Tris gels (Life Technologies) under reducing conditions followed by Coomassie Brilliant Blue staining (Invitrogen LC6060). For Western blot analysis, gels were electroblotted onto nitrocellulose membranes (Bio Rad). Blots were incubated with primary antibodies in 5% non-fat dry milk in PBS plus 0.1% Tween20 overnight at 4°C. Detection was achieved using horseradish peroxidase-conjugate secondary antibody anti-rabbit and anti-mouse (Bio Rad #1706516, #1706515) and visualized with ECL (Cytiva RPN2232). Images were acquired by using a ChemiDoc™ Touch Imaging System (Bio Rad) and analyzed by Image Lab software (Bio Rad).

Antibodies

The primary antibodies used in this study are: rabbit anti-SARS-CoV-2 Spike S1 Subunit (Sino Biological, 40150-T62), mouse anti-His Tag (Invitrogen MA1-21315). Secondary antibody used are: horseradish peroxidase-conjugate anti-rabbit and anti-mouse (Bio Rad #1706516, #1706515).

447 *Densitometric analysis*

448 Intensities of bands corresponding to RBD proteins were measured using Gel Doc
449 2000 and Image Lab software (Bio-Rad, Hercules, CA, USA) in order to measure protein
450 expression levels. Briefly, blots were acquired using the Gel Doc 2000 apparatus; images
451 were imported into the Image Lab software; contrast was adjusted such that the bands were
452 clearly visible on the blot image; area around each band was selected; background intensity
453 was subtracted from the blot image; bands were then selected by drawing a tight boundary
454 around them; intensities of the selected bands were exported in excel format file which was
455 used to perform further analysis.

456

457 *ELISA assay*

458 ELISA plates were coated with different concentrations of *E. coli*-RBD, Insect-RBD
459 and HEK-293-RBD proteins. After washing and blockade of free protein-binding sites with
460 PBS – 0.05% Tween20 - 3% BSA, different concentrations of rat serum (immunized with
461 COVID-*e*Vax vaccine) or anti-SARS-CoV-2 Spike S1 Subunit antibody (Sino Biological,
462 40150-T62) were added to each well and incubated overnight at 4°C in PBS – 0.05%
463 Tween20 - 1% BSA. After washing, AP-conjugated goat anti-rat IgG antibody (SIGMA
464 A8438) or AP-conjugated goat anti-rabbit IgG antibody (SIGMA A8025) was added, and the
465 plates were further incubated for 1 hour at RT. Finally, 3,3′,5,5′-Tetramethylbenzidine (TMB)
466 Liquid Substrate System (Sigma T8665) or alkaline Phosphatase Yellow (pNPP) Liquid
467 Substrate System for Elisa (Sigma P7998) were added as a substrate. After 30 minutes the
468 TMB reaction was stopped with the stop reagent for TMB substrate (Sigma S5814) and the
469 absorbance was measured at 450 nm, while the pNPP reaction was measured at 405nm at
470 different time points.

FACS

Vero E6 cells were incubated with RBD protein (0.45 µg/mL, final concentration) followed by incubation with human anti-RBD antibody (primary antibody) (40150-D003, Sino Biological) and goat anti-human IgG AF488-conjugated antibody (secondary antibody) (A-11013, Thermo Fisher Scientific). Staining with only secondary antibody was used to determine the level of background due to non-specific antibody binding. Each staining step was performed at 4°C for 20 minutes in FACS buffer. Samples were run on a CytoFlex flow cytometer (Beckman Coulter). Analysis was performed using the CytExpert software (Beckman Coulter).

CD static spectra and thermal denaturation

Far-UV (200-250 nm) circular dichroism (CD) spectra of the SARS-CoV-2 HEK-293-RBD and (insect) were recorded using 0.6 mg/mL protein solution in PBS pH 7.4 and 0.3 mg/mL protein solution in 50 mM Tris·HCl, 150 mM NaCl, respectively. CD spectra of the *E. coli*-RBD were monitored at 0.5 mg/mL protein concentration in 50 mM Tris·HCl, 150 mM NaCl, 1 mM TCEP, 20% glycerol, pH 8. All CD static spectra were collected at 20°C, scanning at 50 nm/min, using a 0.1 cm path length quartz cuvette (Hellma, Plainview, NY) and a JASCO-815 spectropolarimeter equipped with a Jasco programmable Peltier element (Jasco, Easton, MD, USA). For each sample five scans were averaged and scans corresponding to buffer solution were averaged and subtracted from the sample spectra. Results are expressed as the molar ellipticity ([Θ]). The formula used to calculate the molar ellipticity is: $[\theta] = (\theta \times MW)/(C \times L \times 10)$, where $[\theta]$ is the molar ellipticity, θ the experimental ellipticity in mdeg, MW the molecular weight of the protein in Daltons, C the protein

concentration in mg/mL, and L the path length of the cuvette in cm. Secondary structure composition was assessed using the BeStSel analysis server[37,38].

CD thermal denaturation experiments were followed at 222 nm, heating from 20 to 90°C at a rate of 1 °C min^{-1} controlled by a Jasco programmable Peltier element (Jasco, Easton, MD, USA). The dichroic activity at 222 nm and the photomultiplier voltage (PMTV) were continuously monitored in parallel every 1.0 °C[39]. Data were fitted to a standard two-state denaturation[40], according to the equation (1):

$$\Delta G_{D-N} = \Delta H_m \left(1 - \frac{T}{T_m}\right) + \Delta C_p \left[T - T_m - \left(T \, ln \frac{T}{T_m}\right)\right]$$

where $\Delta G_{D\text{-}N}$ is the free energy of the unfolding process, T_m is the melting temperature that corresponds to midpoint of the thermal denaturation, ΔH_m is the enthalpy of denaturation at the transition midpoint, and ΔC_p is the change of heat capacity of denaturation. The latter parameter is related to the amount of hydrophobic area that becomes exposed to solvent upon unfolding and it is constant for a given protein. To a first approximation, the melting temperature (T_m) of unfolding has been estimated using the ΔC_p value reported for a α-chymotrypsin (241 amino acids)[41]. All denaturation experiments were performed in triplicate.

Size exclusion chromatography

Analytical gel filtration chromatography was performed using a Superdex 200 Increase 10/300 GL SEC column (Cytiva, USA) coupled to an HPLC system (Azura System, Knauer- Berlin, Ge) equipped with a UV-vis absorbance detector (Smartline 2520, Knauer-Berlin, Ge). The column was equilibrated with 50 mM Tris·HCl pH 8.0, containing 150 mM NaCl. A total of 40 µg of HEK-293-RBD, 47 µg of Insect-RBD and 40 µg of *E. coli*-RBD were injected into the column and eluted at a flow rate of 0.75 mL/min in isocratic mode. Elution profile was followed at 280 nm at room temperature. The shape of the elution profiles and

519 the difference between HEK-293-RBD, Insect-RBD, and RBD (E. c oli) were observed
520 reproducibly in three independent experiments.

521

522 *Bio-layer interferometry (BLI)*

523 Binding studies were carried out using the Octet Red system (Forte Bio). All steps
524 were performed at 25°C with shaking at 600 rpm in a 96-well plate (microplate 96 well, F-
525 bottom, black, 655209, from Greiner bio-one) containing 200 µL of solution in each well.
526 Kinetics buffer 1x (cat. No.18-1105, Forte Bio) was used throughout this study for samples
527 dilution and for sensors washing.

528 Kinetic assays were performed by first capturing ACE2-hFc using anti-human Fc
529 Octet biosensors (Anti-human IgG Fc Capture Biosensors, cat. No. 18-5060, FORTEBIO).
530 Biosensors were soaked for 10 min in 1x kinetic buffer followed by a baseline signal
531 measurement for 60 s and then loaded with ACE2-hFc recombinant protein (10 µg/mL) for
532 300 s (until the biosensor was fully saturated). After a wash step in 1× kinetic buffer for 120
533 s, the ACE2-Fc-captured biosensor tips were then submerged for 300 s in wells containing
534 different concentrations of antigen (RBD E. coli, insect, and HEK-293) to evaluate
535 association curves, followed by 900 s of dissociation time in kinetic buffer. The ACE2-hFc
536 captured biosensor tips were also dipped in wells containing kinetic buffer to allow single
537 reference subtraction to compensate for the natural dissociation of captured ACE2-hFc.
538 Biosensor tips were used without regeneration.

539 The binding curve data were collected and then analysed using data analysis
540 software version 12.0 (FORTEBIO). Binding sensorgrams were first aligned at the last 5
541 seconds of the baseline step average. The single reference subtraction binding
542 sensorgrams were globally fit to a 1:1 Langmuir binding model to calculate K_d values.

Conflict of interest

Mariano Maffei, Grazia Vitagliano, Shaila Sellathurai, Federica Bucci, Alessia Muzi and Valerio Chiarini are employees of Takis and Evvivax companies. Mirco Compagnone is employee of NeoMatrix. Luigi Fedele is a former employee of Takis company. Giuseppe Roscilli and Emanuele Marra are co-founders of Takis and Evvivax companies.

The other authors declare no conflict of interest.

Author Contributions

Mariano Maffei: Conceptualization; Investigation; Supervision; writing-original draft; writing-review and editing. **Linda Celeste Montemiglio:** Investigation; writing-review and editing. **Grazia Vitagliano:** Investigation; editing. **Luigi Fedele:** Investigation. **Shaila Sellathurai:** Investigation. **Federica Bucci:** Investigation. **Mirco Compagnone:** Investigation. **Valerio Chiarini:** Investigation. **Cécile Exertier:** Investigation. **Alessia Muzi:** Investigation. **Giuseppe Roscilli:** Resources; writing-review and editing. **Beatrice Vallone:** Resources, writing-review and editing. **Emanuele Marra:** Conceptualization; Supervision; resources, writing-review and editing.

Acknowledgments

We acknowledge the assistance of L. Salvini and L.Tinti of the Toscana Life Science facilities for the support with mass spectrometry measurements and analysis of proteins. This work was supported by “Regione Lazio” (POR FESR 2014-2020, Avviso Pubblico “EMERGENZA CORONAVIRUS E OLTRE”, CUP F84E21000030006).

References

1. Lai CC, Shih TP, Ko WC, Tang HJ, Hsueh PR (2020) Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. Int. J. Antimicrob. Agents [Internet] 55:105924. Available from: https://doi.org/10.1016/j.ijantimicag.2020.105924

2. Anon World Health Organization (WHO) coronavirus (2021). Available from: https://covid19.who.int

3. Tai W, He L, Zhang X, Pu J, Voronin D, Jiang S, Zhou Y, Du L (2020) Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. Cell. Mol. Immunol. [Internet] 17:613–620. Available from: http://dx.doi.org/10.1038/s41423-020-0400-4

4. Wang MY, Zhao R, Gao LJ, Gao XF, Wang DP, Cao JM (2020) SARS-CoV-2: Structure, Biology, and Structure-Based Therapeutics Development. Front. Cell. Infect. Microbiol. 10:1–17.

5. Yang Y, Du L (2021) SARS-CoV-2 spike protein: a key target for eliciting persistent neutralizing antibodies. Signal Transduct. Target. Ther. [Internet] 6:2020–2022. Available from: http://dx.doi.org/10.1038/s41392-021-00523-5

6. Shamsi A, Mohammad T, Anwar S, Amani S, Khan MS, Husain FM, Rehman MT, Islam A, Hassan MI (2021) Potential drug targets of SARS-CoV-2: From genomics to therapeutics. Int. J. Biol. Macromol. [Internet] 177:1–9. Available from: https://doi.org/10.1016/j.ijbiomac.2021.02.071

7. Johnson BA, Xie X, Kalveram B, Lokugamage KG, Muruato A, Zou J, Zhang X, Juelich T, Smith JK, Zhang L, et al. (2020) Furin Cleavage Site Is Key to SARS-CoV-2 Pathogenesis. bioRxiv Prepr. Serv. Biol. [Internet]. Available from:

590 http://www.ncbi.nlm.nih.gov/pubmed/32869021%0Ahttp://www.pubmedcentral.nih.gov/artic
591 lerender.fcgi?artid=PMC7457603

592 8. Cao W, Dong C, Kim S, Hou D, Tai W, Du L, Im W, Zhang XF (2021) Biomechanical
593 characterization of SARS-CoV-2 spike RBD and human ACE2 protein-protein interaction.
594 Biophys. J. [Internet] 120:1011–1019. Available from:
595 https://doi.org/10.1016/j.bpj.2021.02.007

596 9. Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, Geng Q, Auerbach A, Li F (2020)
597 Structural basis of receptor recognition by SARS-CoV-2. Nature [Internet] 581:221–224.
598 Available from: http://dx.doi.org/10.1038/s41586-020-2179-y

599 10. Azad T, Singaravelu R, Taha Z, Jamieson TR, Boulton S, Crupi MJF, Martin NT, Brown
600 EEF, Poutou J, Ghahremani M, et al. (2021) Nanoluciferase complementation-based
601 bioreporter reveals the importance of N-linked glycosylation of SARS-CoV-2 S for viral entry.
602 Mol. Ther. [Internet] 29. Available from: https://doi.org/10.1016/j.ymthe.2021.02.007

603 11. Teng S, Sobitan A, Rhoades R, Liu D, Tang Q (2021) Systemic effects of missense
604 mutations on SARS-CoV-2 spike glycoprotein stability and receptor-binding affinity. Brief.
605 Bioinform. [Internet] 22:1239–1253. Available from: /pmc/articles/PMC7665319/

606 12. Casalino L, Gaieb Z, Goldsmith JA, Hjorth CK, Dommer AC, Harbison AM, Fogarty CA,
607 Barros EP, Taylor BC, Mclellan JS, et al. (2020) Beyond shielding: The roles of glycans in
608 the SARS-CoV-2 spike protein. ACS Cent. Sci. [Internet] 6:1722–1734. Available from:
609 https://dx.doi.org/10.1021/acscentsci.0c01056

610 13. Lopez-Rincon A, Perez-Romero CA, Tonda A, Mendoza-Maldonado L, Claassen E,
611 Garssen J, Kraneveld AD (2021) Design of Specific Primer Sets for the Detection of B.1.1.7,
612 B.1.351 and P.1 SARS-CoV-2 Variants using Deep Learning. bioRxiv [Internet]
613 70:2021.01.20.427043. Available from: https://doi.org/10.1101/2021.01.20.427043

614 14. Lai S, Floyd J, Tatem A (2021) WorldPop : Preliminary risk analysis of the international
615 spread of new COVID-19 variants. Available from:
616 https://www.worldpop.org/events/covid_variants

617 15. Singh J, Samal J, Kumar V, Sharma J, Agrawal U, Ehtesham NZ, Sundar D, Rahman
618 SA, Hira S, Hasnain SE (2021) Structure-function analyses of new sars-cov-2 variants
619 b.1.1.7, b.1.351 and b.1.1.28.1: Clinical, diagnostic, therapeutic and public health
620 implications. Viruses [Internet] 13:439. Available from: https://doi.org/10.3390/v13030439

621 16. Anon Tracking SARS-CoV-2 variants. Available from:
622 https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/

623 17. Soltani-Zangbar MS, Aghebati-Maleki L, Hajivalili M, Haji-Fatahaliha M, Motavalli R,
624 Mahmoodpoor A, Kafil HS, Farhang S, Pourakbari R, Jadidi-Niaragh F, et al. (2021)
625 Application of newly developed SARS-CoV2 serology test along with real-time PCR for early
626 detection in health care workers and on-time plasma donation. Gene Reports 23:101140.

627 18. Frumence E, Lebeau G, Viranaicken W, Dobi A, Vagner D, Lalarizo Rakoto M,
628 Sandenon Seteyen A-L, Giry C, Septembre-Malaterre A, Raffray L, et al. (2021) Robust and
629 low-cost ELISA based on IgG-Fc tagged recombinant proteins to screen for anti-SARS-CoV-
630 2 antibodies. J. Immunol. Methods [Internet] 495:113082. Available from:
631 https://linkinghub.elsevier.com/retrieve/pii/S0022175921001277

632 19. Salvatori G, Luberto L, Maffei M, Aurisicchio L, Roscilli G, Palombo F, Marra E (2020)
633 SARS-CoV-2 spike protein: An optimal immunological target for vaccines. J. Transl. Med.
634 [Internet] 18:222. Available from: https://doi.org/10.1186/s12967-020-02392-y

635 20. Amanat F, Stadlbauer D, Strohmeier S, Nguyen THO, Chromikova V, McMahon M, Jiang
636 K, Arunkumar GA, Jurczyszak D, Polanco J, et al. (2020) A serological assay to detect
637 SARS-CoV-2 seroconversion in humans. Nat. Med. 2020 267 [Internet] 26:1033–1036.

638 Available from: https://www.nature.com/articles/s41591-020-0913-5

639 21. Consortium AA (2020) Structural and functional comparison of SARS-CoV-2-spike
640 receptor binding domain produced in Pichia pastoris and mammalian cells. Sci. Rep.
641 10:21779.

642 22. Shin Y-J, König-Beihammer J, Vavra U, Schwestka J, Kienzl NF, Klausberger M, Laurent
643 E, Grünwald-Gruber C, Vierlinger K, Hofner M, et al. (2021) N-Glycosylation of the SARS-
644 CoV-2 Receptor Binding Domain Is Important for Functional Expression in Plants. Front.
645 Plant Sci. 0:1154.

646 23. Tripathi NK (2016) Production and purification of recombinant proteins from Escherichia
647 coli. ChemBioEng Rev. 3:116–133.

648 24. Li T, Zheng Q, Yu H, Wu D, Xue W, Xiong H, Huang X, Nie M, Yue M, Rong R, et al.
649 (2020) SARS-CoV-2 spike produced in insect cells elicits high neutralization titres in non-
650 human primates. https://doi.org/10.1080/22221751.2020.1821583 [Internet] 9:2076–2090.
651 Available from: https://www.tandfonline.com/doi/abs/10.1080/22221751.2020.1821583

652 25. Shajahan A, Supekar NT, Gleinich AS, Azadi P (2020) Deducing the N- and O-
653 glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. Glycobiology
654 [Internet] 30:981. Available from: /pmc/articles/PMC7239183/

655 26. Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M (2020) Site-specific glycan
656 analysis of the SARS-CoV-2 spike. Science (80-.). [Internet] 369:330–333. Available from:
657 https://science.sciencemag.org/content/369/6501/330

658 27. Mycroft-West C, Su D, Elli S, Li Y, Guimond S, Miller G, Turnbull J, Yates E, Guerrini M,
659 Fernig D, et al. (2020) The 2019 coronavirus (SARS-CoV-2) surface protein (Spike) S1
660 Receptor Binding Domain undergoes conformational change upon heparin binding. :1–9.

661 28. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, Zhang Q, Shi X, Wang Q, Zhang L, et al.

662 (2020) Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2
663 receptor. Nat. 2020 5817807 [Internet] 581:215–220. Available from:
664 https://www.nature.com/articles/s41586-020-2180-5

665 29. Conforti A, Marra E, Palombo F, Roscilli G, Ravà M, Fumagalli V, Muzi A, Maffei M,
666 Luberto L, Lione L, et al. (2021) COVID-eVax, an electroporated plasmid DNA vaccine
667 candidate encoding the SARS-CoV-2 Receptor Binding Domain, elicits protective immune
668 responses in animal models of COVID-19. bioRxiv [Internet]:2021.06.14.448343. Available
669 from: https://www.biorxiv.org/content/10.1101/2021.06.14.448343v1

670 30. Ogando NS, Dalebout TJ, Zevenhoven-Dobbe JC, Limpens RWAL, van der Meer Y,
671 Caly L, Druce J, de Vries JJC, Kikkert M, Barcena M, et al. (2020) SARS-coronavirus-2
672 replication in Vero E6 cells: Replication kinetics, rapid adaptation and cytopathology. J. Gen.
673 Virol. [Internet] 101:925–940. Available from:
674 https://www.microbiologyresearch.org/content/journal/jgv/10.1099/jgv.0.001453

675 31. Ren X, Glende J, Al-Falah M, de Vries V, Schwegmann-Wessels C, Qu X, Tan L,
676 Tschernig T, Deng H, Naim HY, et al. (2006) Analysis of ACE2 in polarized epithelial cells:
677 Surface expression and function as receptor for severe acute respiratory syndrome-
678 associated coronavirus. J. Gen. Virol. [Internet] 87:1691–1695. Available from:
679 https://www.microbiologyresearch.org/content/journal/jgv/10.1099/vir.0.81749-0

680 32. Creech CB, Walker SC, Samuels RJ (2021) SARS-CoV-2 Vaccines. JAMA - J. Am. Med.
681 Assoc. [Internet] 325:1318–1320. Available from: https://jamanetwork.com/

682 33. Nagy A, Alhatlani B (2021) An overview of current COVID-19 vaccine platforms. Comput.
683 Struct. Biotechnol. J. 19:2508–2517.

684 34. Anon COVID-19 Studies from the World Health Organization Database -
685 ClinicalTrials.gov. Available from: https://clinicaltrials.gov/ct2/who_table

686 35. Chen J, Miao L, Li JM, Li YY, Zhu QY, Zhou CL, Fang HQ, Chen HP (2005) Receptor-
687 binding domain of SARS-Cov spike protein: Soluble expression in E.coli, purification and
688 functional characterization. World J. Gastroenterol. 11:6159–6164.

689 36. Prahlad J, Struble L, Lutz WE, Wallin SA, Khurana S, Schnaubelt A, Broadhurst MJ,
690 Bayles K, Borgstahl GEO (2021) Bacterial expression and purification of functional
691 recombinant SARS-CoV-2 spike receptor binding domain. bioRxiv
692 [Internet]:2021.02.03.429601. Available from:
693 http://biorxiv.org/content/early/2021/02/03/2021.02.03.429601.abstract

694 37. Micsonai A, Wien F, Kernya L, Lee Y-H, Goto Y, Réfrégiers M, Kardos J (2015) Accurate
695 secondary structure prediction and fold recognition for circular dichroism spectroscopy.
696 Proc. Natl. Acad. Sci. [Internet] 112:E3095–E3103. Available from:
697 https://www.pnas.org/content/112/24/E3095

698 38. Micsonai A, Wien F, Bulyáki É, Kun J, Moussong É, Lee Y-H, Goto Y, Réfrégiers M,
699 Kardos J (2018) BeStSel: a web server for accurate protein secondary structure prediction
700 and fold recognition from the circular dichroism spectra. Nucleic Acids Res. [Internet]
701 46:W315–W322. Available from:
702 https://academic.oup.com/nar/article/46/W1/W315/5035652

703 39. Benjwal S, Verma S, Röhm K-H, Gursky O (2006) Monitoring protein aggregation during
704 thermal unfolding in circular dichroism experiments. Protein Sci. [Internet] 15:635–639.
705 Available from: https://onlinelibrary.wiley.com/doi/full/10.1110/ps.051917406

706 40. Baldwin RL (2000) Structure and mechanism in protein science. A guide to enzyme
707 catalysis and protein folding, by A. Fersht. 1999. New York: Freeman. 631 pp. $67.95
708 (hardcover). Protein Sci. [Internet] 9:207–207. Available from:
709 https://onlinelibrary.wiley.com/doi/full/10.1110/ps.9.1.207

41. Myers JK, Pace CN, Scholtz JM (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. Protein Sci. [Internet] 4:2138. Available from: /pmc/articles/PMC2142997/?report=abstract

Supporting information

Additional supporting information may be found online in the Supporting Information section.