Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Identification and application of self-binding zipper-like sequences in SARS-CoV spike protein

Si Min Zhang a,b, Ying Liao c, Tuan Ling Neo a, Yanning Lu c, Ding Xiang Liu a, Anders Vahlne b, James P. Tam a,⁎

a School of Biological Sciences, Nanyang Technological University, Singapore
b Division of Clinical Microbiology, Department of Laboratory Medicine, Karolinska Institute, Sweden
c Shanghai Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Shanghai, China

ARTICLE INFO

Keywords:
Steric β-zipper
SARS-CoV spike protein
Class I viral fusion glycoprotein
self-binding peptides
viral detection

ABSTRACT

Self-binding peptides containing zipper-like sequences, such as the Leu/Ile zipper sequence within the coiled coil regions of proteins and the cross-β spine steric zippers within the amyloid-like fibrils, could bind to the protein-of-origin through homophilic sequence-specific zipper motifs. These self-binding sequences represent opportunities for the development of biochemical tools and/or therapeutics. Here, we report on the identification of a putative self-binding β-zipper-forming peptide within the severe acute respiratory syndrome-associated coronavirus spike (S) protein and its application in viral detection. Peptide array scanning of overlapping peptides covering the entire length of S protein identified 34 putative self-binding peptides of six clusters, five of which contained octapeptide core consensus sequences. The Cluster I consensus octapeptide sequence GINITNFR was predicted by the Eisenberg’s 3D profile method to have high amyloid-like fibrillation potential through steric β-zipper formation. Peptide C6 containing the Cluster I consensus sequence was shown to oligomerize and form amyloid-like fibrils. Taking advantage of this, C6 was further applied to detect the S protein expression in vitro by fluorescence staining. Meanwhile, the coiled-coil-forming Leu/Ile heptad repeat sequences within the S protein were under-represented during peptide array scanning, in agreement with that long peptide lengths were required to attain high helix-mediated interaction avidity. The data suggest that short β-zipper-like self-binding peptides within the S protein could be identified through combining the peptide scanning and predictive methods, and could be exploited as biochemical detection reagents for viral infection.

1. Introduction

Severe acute respiratory syndrome (SARS) is a contagious atypical pneumonia that caused an epidemic between November 2002 and July 2003 with a 9.6% mortality rate (8096 known cases and 774 deaths) (http://www.who.int/csr/sars/country/table2004_04_21/en/). While the SARS-CoV epidemic has subsided, the closely related Middle East respiratory syndrome coronavirus (MERS-CoV) infection is still presenting severe public health threats locally and globally. Generation of effective anti-viral therapeutic and detection agents is warranted, and the SARS-CoV spike (S) protein, which mediates viral entry and initiates viral infection, represents attractive drug development target. The SARS-CoV surface spike (S) glycoprotein is composed of the surface-exposed S1 subunit and the transmembrane S2 subunit which further hetero-trimerize on the viral envelope, a characteristic structural prerequisite for Class I fusion glycoproteins to facilitate viral entry (Fig. 1) (Bosch et al., 2003; Xiao and Dimitrov, 2004). Specifically, following the recognition of the cellular receptor angiotensin-converting enzyme 2 (ACE2) by the receptor binding domain (RBD) within the surface S1 subunit, structural changes are induced within the S2 subunit, exposing the fusion peptide (FP) within S2 and enabling it to be inserted into the target cell membranes. This leads to the heptad repeat 1 (HR1) regions C-terminal to the FP to oligomerize in an antiparallel fashion with the more C-terminal HR2 regions of the S protein trimers, resulting in a coiled-coil structure, i.e., the six-helix bundle (6-HB) fusion cores (Bosch et al., 2003). The 6-HB structure is further strengthened by hetero-dimerization of the hydrophobic regions flanking the two HRs, specifically the membrane proximal external region (MPER) and the S2

---

Abbreviations: S, Spike protein; HR, Heptad repeat; HIV-1, Human immunodeficiency virus type I; SARS-CoV, Severe acute respiratory syndrome associated coronavirus; RBD, Receptor binding domain; FP, Fusion peptide; 6-HB, Six-helix bundle; MPER, Membrane proximal external region; IFP, Internal fusion peptide; HA, Hemagglutinin; PAGE, Polyacrylamide gel electrophoresis; CD, Circular dichroism; ThT, Thioflavin T

⁎ Corresponding author.
E-mail address: jptam@ntu.edu.sg (J.P. Tam).

https://doi.org/10.1016/j.biocel.2018.05.012
Received 7 December 2017; Received in revised form 3 May 2018; Accepted 21 May 2018
Available online 22 May 2018
1357-2725/ © 2018 Elsevier Ltd. All rights reserved.
internal fusion peptide (IFP) (Guillen et al., 2008; Liao et al., 2015). Collectively, the opposing viral and cellular membranes are juxtaposed into proximity for the subsequent fusion and viral entry (Bosch et al., 2003).

One of the approaches to generate anti-viral therapeutics has been focused on self-binding peptides. Self-binding peptides bind to their cognate proteins through sequence-specific interactions (Beisswenger and Cabrele, 2014; Hard and Lendel, 2012), with examples including Leu/Ile zipper sequences in coiled-coil motifs (Hakoshima, 2001; Perutz, 1994), cross-β spine steric zipper sequences (abbreviated as steric β-zipper) in amyloid and amyloid-like fibrils (Nelson et al., 2005; Parry, 2005; Sawaya et al., 2007), and aromatic-residue-rich sequences in aromatic zippers (Cochran et al., 2001), among others. Their abilities to conduct sequence-specific interactions have been exploited and been further developed into peptidyl therapeutics and biochemical reagents (Azzarito et al., 2013; Hakoshima, 2001). For anti-viral applications, Leu/Ile zipper-like sequences within HRs have successfully been developed as entry inhibitors against infections such as the human immunodeficiency virus type I and SARS-CoV (Borrego et al., 2013; Bosch et al., 2004; Liu et al., 2005, 2004; Medinas et al., 2002; Wang et al., 2011; Wild et al., 1994; Yao and Compans, 1996; Young et al., 1999).

Specifically, the HR-derived Leu/Ile zipper-like peptides can bind to the fusion glycoprotein in a dominant-negative fashion, and thereby halt the viral entry and infection.

According to our understanding of the literature, aside from the Leu/Ile zipper-like sequences within S2 subunit, other potentially self-interacting zipper-like sequences in other S protein domains that mediate the multi-step viral entry process of SARS-CoV are less known.

---

**Fig. 1.** Screening of the 20-residue peptide array library with S1188HA expressed in Sf9 cells. a. Screening result of the 20-residue peptide array library. Retained S1188HA was identified as dark spots on the peptide-spotted membrane. b. Densitometry analysis of the positive spots in A revealed the S1188HA-binding pattern. Five high-affinity-binding clusters (I, II, III, IV, and V), which contained three out of six consecutive peptides exhibiting optical density values at least 2.5 times of the background level (Rigter et al., 2007), were shown with the representative peptides (*). RBD-derived peptides bound to S1188HA with high affinity were clustered as Cluster R. False positive peptides that bound to the detection antibodies were indicated by the arrow heads and excluded in data analysis.
and studied. The globular S1 subunit is structurally different from S2 subunit, and is largely composed of β-sheets, such as within the RBD (Gianneccini et al., 2005). Logically, identification of zipper-like sequences within the S1 subunit should focus on β-zipper-like motifs, which include steric β-zippers. These complementary, dry, puzzle-like interfaces, form between as few as four and as many as seven tightly interdigitating side chains of two facing peptide β-sheets (Azriel and Gazit, 2001; Balbach et al., 2000; Balbirnie et al., 2001; Nelson et al., 2005; Reches et al., 2002; Tenidis et al., 2000; von Bergen et al., 2000). Self-binding peptides containing this motif have been shown to recognize and homo-dimerize with the same sequence within amyloid-like fibrils and then inhibit the fibrillation process (Hard and Lendel, 2012). This suggests that the steric beta-zipper sequences within a viral fusion glycoprotein, such as SARS-CoV S protein, could also bind to the cognate viral fusion protein through homophilic interaction, and could further serve as a biochemical reagent for viral detection and/or as anti-viral therapeutics.

In the current study, we identified the self-binding β-zipper-like peptides within the S protein of SARS-CoV. Firstly, putative self-binding peptides were identified by screening a peptide array library that was composed of overlapping peptides covering the entire length of S protein (Kaukinen et al., 2003; Rigter et al., 2007; Uttamchandani and Yao, 2008). The identified peptides were further analyzed by the 3D profile method to predict the steric β-zipper segments. Subsequently, a peptide that potentially oligomerized through forming steric β-zippers was identified experimentally and further used to detect S protein expression in vitro. Finally, we could demonstrate the differential influences of peptide lengths on β- or helix-mediated interactions. Altogether, these data suggest that endogenous steric β-zipper peptides within viral fusion proteins, such as SARS-CoV S protein, could be a practical source for developing detection reagents against viral infections.

2. Materials and methods

2.1. Cell lines and plasmids

Plasmid pXJ40-S encoding the full-length S protein of SARS-CoV, S1255, (Sin 2774, accession number: P50594) was kindly provided by the Institute of Molecular Cell Biology, Singapore. Plasmid pCDNA3.1-OPTS-HA encoding a HA-tagged codon-optimized full-length S protein under a 5’ T7 promoter was kindly provided by Dr. Zhang Linqi from the Aaron Diamond AIDS Research Center (Chen et al., 2005). Plasmid pVL1392(S1188HA) encoding the ectodomain (aa 1-1188) of the S protein (Sin 2774) with a C-terminal human influenza hemagglutinin (HA) tag was constructed from pXJ40-S (Lu et al., 2008a).

2.2. Peptide array library construction and Fmoc solid phase peptide synthesis

Peptide array libraries containing SARS-CoV (Sin 2774) S protein-derivatized overlapping peptides were synthesized and simultaneously anchored through the C-termini on PEG-derivatized membranes (Amino-PEG500 –UC540 Sheet, Intavis AG, Germany), using Fmoc chemistry and MultiPep (SPOT) synthesizer (Intavis AG, MultiPep, Germany) (Frank, 2002; Frank and Overwin, 1996). S protein-derived peptides were chemically synthesized with or without C-terminal biotin tags using 0.1 mM biotinylated Rink and Wang resins, respectively, using the Liberty™ automated microwave peptide synthesizer (Applied Photo-physics) between 180 nm and 260 nm, with a 0.5 nm step resolution. The secondary structure of the peptide was estimated as previously described (Zhang et al., 2015). Brieﬂy, peptides dissolved in ddH2O were measured by a Chirascan circular dichroism spectrometer (Applied Photo-physics) between 180 nm and 260 nm, with a 0.5 nm step resolution. The secondary structure of the peptide was estimated as previously described (Zhang et al., 2015).

2.3. Peptide array screening and pull down experiments

Peptide array library on PEG-derivatized membranes or biotinylated peptides mixed with UltraLink™ Immobilized NeutrAvidin™ resins (Pierce) were blocked with Superblock™ Blocking Buffer (Pierce) and probed with S1188HA-containing S9 cell lysates (Lu et al., 2008a), at RT for 1 h in 200 µl binding buffer (PBS, 1% NP40, 0.1% Sodium decyl sulfate (SDS), 2 M urea). S1188HA retained on the peptide array membrane or resin-bound peptides was detected via chemiluminescence immunoassay or Western blot, respectively, using rabbit anti-HA (Y-11) (Santa Cruz Biotechnology, USA) followed by HRP-conjugated swine anti-rabbit secondary antibody (Dako, Denmark).

2.4. Chemical crosslinking

Biotinylated peptides (1 mM) were incubated with 0, 1, 5, or 10 mM glutaraldehyde at RT for 1 h, followed by quenching with 100 mM glycine. Peptide oligomers were then separated on 16% Tricine SDS-PAGE and analyzed by Western blot using avidin-HRP (Santa Cruz Biotechnology, USA).

2.5. Prediction of the amyloid-like fibril-forming segments by the 3D profile method

Sequence of the SARS-CoV S protein (Sin 2774) was analyzed with the 3D profile method (available online at http://services.mbi.ucla.edu/zipperdb/) to predict amyloid-like sequences, with the cut-off ROSETTA energetic scores set at −3.7 kcal/mol/residue (Thompson et al., 2006).

2.6. Circular dichroism (CD) spectroscopy

CD spectroscopy analysis was performed as previously described (Zhang et al., 2015). Briefly, peptides dissolved in dH2O were measured by a Chirascan circular dichroism spectrometer (Applied Photo-physics) between 180 nm and 260 nm, with a 0.5 nm step resolution. The secondary structure of the peptide was estimated as previously described (Zhang et al., 2015).

2.7. Congo Red and Thioflavin T (ThT) assays

Amyloid-like fibrils were identified through the Congo Red and ThT assays, as previously described (Nilsson, 2004). Briefly, following a three-day incubation in DMSO at 10 mg/ml, peptides were incubated with 1 ml Congo Red solution (0.035 g/L Congo Red, 5 mM potassium phosphate, 150 mM NaCl, pH 7.4) for 30 min, or with 1 ml ThT solution (0.016 g/L ThT, 10 mM phosphate, 150 mM NaCl, pH 7.4). The absorbance spectrum of the peptide/Congo Red mixture was recorded between 400 and 700 nm, and the spectrum of the Congo Red solution mixed with 10 µL DMSO subtracted. The fluorescence emission spectrum of the peptide/ThT mixture was recorded between 460 nm and 540 nm with a bandwidth of 10 nm and an excitation wavelength at 440 nm.

2.8. Immunofluorescence staining

293T cells were transiently transfected with pcDNA3.1-OPTS-HA or pXJ40-S, with or without pre-infection with recombinant vaccinia virus-T7 (VTF7.3) (MOI = 1), respectively, using Lipofectamine 2000 (Life Technologies, Singapore). Twenty-four h post-transfection, cells were fixed with 4% paraformaldehyde and permeabilized with 0.5% Triton X-100. Cells were co-stained with 2 µM C6 and anti-HA (Y-11) or anti-S antibody (Lu et al., 2008b), followed by Dylight 488-conjugated NeutrAvidin™ (Thermo Fisher Scientific) and Alexa 594-conjugated anti-rabbit antibody (Molecular Probes) for anti-HA primary staining or Alexa 488-conjugated anti-mouse antibody for anti-S primary staining.

S.M. Zhang et al.  International Journal of Biochemistry and Cell Biology 101 (2018) 103–112
Slides were mounted in Prolong Gold Anti-Fade medium with DAPI (Invitrogen), and analyzed with an Axiovert 200M inverted fluorescence compound microscope (Carl Zeiss, Germany).

3. Results & discussion

3.1. Peptide array mapping of S1188HA-binding sequences

To map putative self-binding sequences within the S protein, peptide array libraries of overlapping 15- to 20-residue peptides with offsets of two to four residues were derived from SARS-CoV S protein (Frank, 2002; Frank and Overwin, 1996), synthesized on PEG-derivatized cellulose membranes, and then probed by S1188HA, a recombinant HA-tagged S protein ectodomain (aa 1-1188) that is structurally more stable than the full length S protein (Lu et al., 2008a).

S1188HA in this study was expressed in the Sf9 insect cell system, based on the observations that the Sf9 insect cell offers higher expression level of S1188HA without affecting the protein immunogenicity, receptor-binding properties or its ability to construct virus-like particles, as compared to the mammalian system (Jefferis et al., 2004; Mortola and Roy, 2004). Peptide-retained S1188HA was then detected as dark spots on the membranes by an anti-HA chemiluminescence imaging system (Supplementary S1 Fig.). False-positive sequences were identified by immunoblotting the peptide-anchored membranes with an HRP-conjugated secondary antibody alone or in combination with a primary anti-HA antibody, without probing the membranes with S1188HA (S2 Fig.). Neighboring overlapping peptides that did not bind to S1188HA served as negative controls, further ensuring the specificity of the peptide-S1188HA interactions.

Based on primary screening data using libraries of different peptide lengths, the 20-residue peptide library with an offset of 4 aa provided the most comprehensive hit peptide profile, and was chosen for further study. In brief, probing 310 peptides in the 20-residue library with S1188HA identified thirty-four S1188HA-binding peptides, sequences of which were located to the N-terminal region of S1, RBD, the region between S1 and S2, IFP, HR1, and the loop region connecting HRs (Fig. 1a and Table 1). Additionally, these identified peptides appeared in clusters. Five clusters shared a common feature of containing three to four overlapping peptides, each yielding a consensus putative self-binding octapeptide motif (Fig. 1b and Table 1). These were the motif GINITNFR (aa 225-232) in Cluster I peptides C6–C9 derived from the region upstream to the RBD; the motif VLTPSKR (aa 537-544) in Cluster II peptides F12–F15 derived from the region immediately following RBD; the motif LNRALSGI (aa 745-752) in Cluster III peptides C6–C9 derived from the region immediately following RBD; the motif AMQMAVRF (aa 881-888) in Cluster IV peptides J2, 3 and 5 derived from IFP; and finally, the motif RITGQR (aa 977-984) in Cluster V peptides K2–K5 derived from the HR1 (Table 1). Besides the five continuous clusters of self-binding peptides, the sixth cluster R contained a cluster of eight discontinuous RBD-derived sequences that bound to S1188HA with relatively high affinity.

It was also noted that no HR2- and only four HR1-derived sequences, K2, K3, K4 and K5, were identified as putative self-binding sequences from the 20-residue peptide array library. The HR regions of Class 1 viral fusion glycoproteins, including SARS-CoV S protein, are known to oligomerize into helix trimers through Leu/Ile zippers (Hakansson-McReynolds et al., 2006; Xu et al., 2004). Accordingly, HR2-derived peptides with a high tendency to form homophilic interactions were synthesized and applied in pull down experiments. These eight peptides were chosen based on their relatively high affinity to S1188HA in peptide array screening and their inclusion of the consensus self-binding sequences (Table 2). Peptide D4A (302 to 321 aa) contained the sequence shifted downstream from the scanning-identified D4 sequence, to facilitate peptide synthesis and purification. The scanning-identified H18 sequence contained an N-terminal cysteine residue. To investigate if the H18-S protein interaction was mediated through homo-oligomerization.

To confirm and investigate the peptide-S protein interaction, 8 representative peptides, C6, D4A, D24, E11, H18A, J1A, K2A and K4A, were synthesized and applied in pull down experiments. These eight peptides were chosen based on their relatively high affinity to S1188HA in peptide array screening and their inclusion of the consensus self-binding sequences (Table 2). Peptide D4A (302 to 321 aa) contained the sequence shifted downstream from the scanning-identified D4 sequence, to facilitate peptide synthesis and purification. The scanning-identified H18 sequence contained an N-terminal cysteine residue. To investigate if the H18-S protein interaction was mediated through

---

Table 1

| Cluster | Peptide | Sequence | Position (Domain) |
|---------|---------|----------|------------------|
| A       | A6      | TDVDQAPNYNTGHSSRKGV | 21-40 |
|         | A5      | NYQTSSGMSYVPPYDPRFR | 29-48 |
|         | C6      | FNTKLFKPKLGINIFNFRI | 213-232 |
|         | C7      | KFLPKPFLGLGINIFNRAIT | 217-240 |
|         | C8      | KLPGINCNFNRAITPASFQDI | 221-244 |
|         | C9      | GINIFNRAITPASFQDI | 225-254 |
|         | D1      | SVKRFKFLGKFQIQNHK | 289-308 |
|         | D2      | DTQSNFVRVCGDVVRPFINIT | 310-320 (RBD) |
|         | D3      | EERKICNSVADCEVSLYNSTF | 341-360 (RBD) |
|         | D23     | LCVSNVADCEVSLYNSTF | 377-396 (RBD) |
|         | D24     | NVYTSFPVKKVGDVRQIAPE | 381-400 (RBD) |
|         | E7      | NYKLPDFDCMCCWALNRTNFIF | 409-428 (RBD) |
|         | E11     | TRNATDSTGGNYKRYIFIR | 415-444 (RBD) |
|         | E13     | TDQNYNRYKLYRKGKIFPE | 433-452 (RBD) |
|         | E20     | PNENGKFTYTTGQGQHK | 477-496 (RBD) |
|         | H1      | MHKSYQGCSTQCNIYK | 721-740 |
|         | H11     | XLLQGVGQFCTNLARSGI | 733-752 |
|         | H19     | PCTQLRNLGSAEQAARPQDRTVF | 741-760 |
|         | I2      | IFVGLGCGTTLPSKRQFP | 525-544 |
|         | I8      | FNGLGTGCTVTLSKRQFP | 529-548 |
|         | I14     | GTTOGKTSLPSEQFPFOQGQ | 533-552 |
|         | I15     | VLTPSKRQFPQFOQFQRV | 537-556 |
|         | V5      | RDKWKEEAQVPIQGLRDLQR | 965-984 (HR1) |
|         | V3      | VEADVQDILGRKGLQQL | 969-988 (HR1) |
|         | K4      | VOVRKIDSSLQGKLQTVQ | 973-992 (HR1) |
|         | K5      | RITGQR | 977-996 (HR1) |
|         | L5      | MRSQFEKYSQTTKQIYK | 1097-1114 (Loop) |
|         | L6      | HNFRTSWSQITGNPLFSPQDI | 1077-1096 (Loop) |
|         | L7      | FITQRWFSQPOQDITTDFTV | 1085-1104 (Loop) |
|         | L12     | NTSTQGCNVDWYQGINSTTV | 1101-1120 (Loop) |

*Consensus putative self-binding sequences were underlined.

---

3.2. S protein-derived peptides pulled down S1188HA putatively through homo-oligomerization

To confirm and investigate the peptide-S protein interaction, 8 representative peptides, C6, D4A, D24, E11, H18A, J1A, K2A and K4A, were synthesized and applied in pull down experiments. These eight peptides were chosen based on their relatively high affinity to S1188HA in peptide array screening and their inclusion of the consensus self-binding sequences (Table 2). Peptide D4A (302 to 321 aa) contained the sequence shifted downstream from the scanning-identified D4 sequence, to facilitate peptide synthesis and purification. The scanning-identified H18 sequence contained an N-terminal cysteine residue. To investigate if the H18-S protein interaction was mediated through

---

Table 2

| Peptide | Sequence | Position (Domain) |
|---------|----------|------------------|
| C6      | FNTKLFKPKLGINIFNFRI | 213-232 |
| D4A     | TSNKRFVPSGDVVRPFINIT | 302-321 (RBD) |
| D24     | NVYTSFPVKKVGDVRQIAPE | 381-400 (RBD) |
| E11     | TRNATDSTGGNYKRYIFIR | 415-444 (RBD) |
| H18A    | TQLRNLGSAEQAARPQDRTVF | 743-762 |
| J1A     | AGWNTGCAAGQCLPSKRQFP | 866-888 (RBD) |
| K2A     | RDKWKEEAQVPIQGLRDLQR | 965-988 (HR1) |
| K4A     | VOVRKIDSSLQGKLQTVQ | 973-996 (HR1) |

---

GB38 CDGISGINSVQNLQKIRKRESEVNAKLLNFLNSSL1513ELG 1149-1186 (HR2)
disulfide bond formation, peptide H18A (743 to 762 aa) with sequence shifted downstream from the H18 by two residues was generated. Peptide J1A (866 to 888 aa) contained sequences of the overlapping J1 and J2 peptides that were identified from the 15-residue and 20-residue peptide array library, respectively. Finally, peptide K2A contains sequences of the overlapping K2 and K3 sequences, while K4A contains sequences of the overlapping K4 and K5 sequences. Each of these eight peptides was entirely or largely composed of sequences identified from peptide array scanning, where the specificities of their interactions with S1188HA were ensured by including continuous overlapping peptides. Therefore, no scrambled peptide was included in the subsequent pull down experiments. Finally, peptide GG38 (aa 1149-1186), which was
the cell lysates were pre-treated with their respective non-biotinylated analogues (Fig. 2b). A possible explanation for this phenomenon is that the peptides C6, E11, K2A and K4A contained oligomerization sequences, as does the positive control HR2 peptide GG38, which contains a trimmerizing sequence (Celigny et al., 2011; Hakansson-McReynolds et al., 2006). The potential oligomerization sequences within C6, E11, K2A and K4A could enable the non-biotinylated peptides to bind to S1188HA during pre-treatment, and further serve as oligomerization sites for their biotinylated analogues during pull down experiments. These oligomerization sequences within the non-biotinylated peptides would be more exposed and accessible than the corresponding region within S1188HA, which could serve as a ‘bridge’ between the biotinylated peptides and S1188HA and thereby lead to an enhanced retention of S1188HA.

To strengthen our evidence for the oligomerization potential of the self-binding peptides C6, E11, K2A and K4A, chemical crosslinking experiments were performed. With increasing concentrations of the chemical crosslinker glutaraldehyde, peptides C6, D24 and E11 subsequently formed dimer, trimer, and higher-order oligomers, in the same way as the HR2 control peptide GG38 (Fig. 2c). We also observed that the HR1-derived peptides K2A and K4A formed dimers under chemical crosslinking experiments, consistent with observation in previous studies by others [46] and by our group (S5 Fig.) that HR1 sequences can oligomerize (Supek et al., 2004). However, we did not observe any higher-order oligomers even at relatively high concentrations of the chemical crosslinker (Fig. 2c), possibly owing to their incomplete inclusion of the HR1 sequence.

Our data suggest that the peptides C6, D24, E11, K2A and K4A self-adhered to S1188HA through a sequence-based, not glycan-dependent, and likely homophilic-interaction-mediated mechanism. All of these oligomerizing peptides, C6, D24, E11, K2A and K4A, were mapped onto the surface exposed regions within the S protein model structure (PDB: 1T7G) (Fig. 2d) (Spiga et al., 2003). Furthermore, most of them have been shown to contain potential antigenic epitopes, which further confirms that these sequences are located in the accessible regions on the S protein (S1 Table) (Duan et al., 2005; Greenough et al., 2005; He et al., 2004). Their homophilic-interaction potential and the accessibility of their corresponding regions on the S protein indicated that these peptides could bind to S protein during viral infection and potentially be used as viral detection reagents, particularly if the peptide-S protein interactions are maintained through sequence-specific interactions such as the β-zipper motifs within amyloid-like fibrils. Indeed, amyloid-derived steric β-zipper peptides have been shown to specifically bind to the cognate amyloid protein through oligomerization with the corresponding sequence within the protein (Hard and Lendel, 2012).

### 3.3. Peptide C6 formed amyloid-like fibrils potentially through a steric β-zipper motif

We next identified the β-zipper-like self-binding peptides within the S protein by sequentially predicting β-zipper segments and then comparing the prediction results with the peptide array scanning results. The S protein-derived peptides with high potential to form steric β-zippers, or the self-complementary motif between facing β-strands, were predicted by the 3D profile method developed by the Eisenberg group (Bowie et al., 1991; Goldschmidt et al., 2010; Thompson et al., 2006); and followed by assessing β-sheet-favorable residue pairings within the peptide self-dimers using pair correlation data from the work of Wouters and Curmi (Wouters and Curmi, 1995). Among the peptides that oligomerized under chemical crosslinking, both C6 and E11 could form multiple residue pairings that were positively correlated with antiparallel β-sheet formation (S7 Fig.), and were further predicted to contain steric β-zipper-like fibril-forming segments.

Particularly, the Cluster I consensus sequence, GINITNFR within C6

![Fig. 3. Peptide C6 exhibited characteristics of amyloid-like fibrils. a. Peptide C6 contained a predicted amyloid-like fibril-forming segment and exhibited β-strand content in an aqueous environment. The Cluster I consensus sequence is in bold and the predicted amyloid-like β-strand sequence is underlined. The CD spectrum of C6 (1 mM) in H2O was shown with its estimated secondary structure contents. NRMSD, normalized root mean square deviation. b. The aged C6 exhibited maximal absorbance at 540 nm when mixed with Congo Red. c. The aged C6 had an increased fluorescence emission at 482 nm, when mixed with ThT.](image-url)
Fig. 4. C6 detected the expression of full-length S protein in 293T cells and the effect of peptide homodimerization on C6-S1188HA interaction. a. Peptide C6 and anti-HA antibodies co-stained 293T cells that were transiently expressing HA-tagged full-length S protein. b. Peptide C6 had the same staining pattern of 293T cells that were transiently expressing full-length S protein, as an anti-S monoclonal antibody did. Control cells were stained with the fluorescein-conjugated secondary antibodies only.
showed similar charge distribution and side chain sizes as the steric β-zipper-like amyloid-forming sequence GNNQQNY (Nelson et al., 2005), and its partial sequence GINITNF was predicted to be fibril-forming (Fig. 3a). Subsequently, the structural profile and fibrillation behavior of C6 was investigated, using CD spectroscopy and Congo red and thioflavin T (ThT) spectroscopic assays, respectively. Peptide C6 was dissolved at 10 mg/ml, a relatively high concentration that could accelerate nucleation stage and allow the observation of fibrillation within the experiment time-span of three days (Teng and Eisenberg, 2009). It has been previously shown that amyloid-like fibrils could bind to Congo Red, resulting in characteristic apple-green birefringence under polarized light and enhanced absorbance at 540 nm (Nilsson, 2004). It is shown here that peptide C6, upon 3-day incubation, exhibited an increase of absorbance at 540 nm when mixed with Congo Red, suggesting its Congo red-binding and fibrillation formation behavior (Fig. 3b). The incubated peptide C6 also resulted in an approximately 2.5 times increase of the fluorescence emission at 482 nm when mixed with ThT, another characteristic behavior of amyloid fibrils (Biancalana and Koide, 2010; Nilsson, 2004), further confirming its ability to form amyloid-like fibrils (Fig. 3c). Furthermore, CD spectroscopy study revealed that C6 adopts a β-strand conformation in aqueous environment, supporting its ability to form β-strand-containing structures such as the β-zipper (Fig. 3a). Meanwhile, the other potentially amyloidogenic peptide E11 did not demonstrate these amyloid-like characteristics (S6 Fig.).

Fig. 5. Peptide C6-S1188HA interaction avidity could be enhanced through homodimerization of peptide C6. a. Lys-based linker molecule for the synthesis of C6-D. b. C6-D aggregated and formed fibril-like network, as observed under light microscopy at 40× (left), and 100× (right) magnifications. c. C6-D had enhanced pull-down efficiency of S1188HA than C6.
Taking advantage of its self-interacting amyloidogenic behavior, C6 was further applied to detect S protein expression by fluorescence staining in 293 T cells, a human embryonic kidney cell line permissive to SARS-CoV infection (Gillim-Ross et al., 2004). The 293 T cells transiently expressing an HA-tagged full-length S protein, were stained with either anti-HA (Y-11) primary antibody followed by Alexa Fluor 594-conjugated secondary antibody (Fig. 4a, panel ‘anti-HA’), or with 2 μM of the biotinylated peptide C6 followed by Dylight 488-conjugated Neutravidin™ (Fig. 4a, panel ‘2μM C6’). The staining pattern of anti-HA antibody targeting the S protein, and that of fluorescence-conjugated Neutravidin™ targeting the C6 peptide, were found to co-localize (Fig. 4a). The cells expressing high levels of the S protein were also bound extensively by peptide C6 (Fig. 4a). We further applied C6 and anti-S antibody separately to detect the expression of an un-tagged full-length S protein extensively by peptide C6 (Fig. 4a). We further applied C6 and anti-S antibody targeting the S protein, and that of Dimeric C6-D had a higher propensity for oligomerization and a stronger affinity towards the S protein, as compared to the monomeric C6, which warrants further investigation of C6-D as a detection reagent for S protein expression.

4. Conclusions

The data presented in this paper demonstrate that β-zipper-like self-binding peptides, which bind to the protein-of-origin by forming sequence-specific steric β-zippers, could be identified by employing peptide scanning and zipper sequence prediction using programs such as the 3D profile method. Self-binding peptides, such as C6 from SARS-CoV S protein, might mimic certain functions of an antibody for protein detection. With an enhanced peptide-protein interaction avidity and specificity, this approach may be further developed as an animal-free way to rapidly generate and produce peptide-based viral diagnostic and research tools.

Acknowledgements

This research was in part supported by the Singapore Ministry of Education AcRF Tier 3 funding (MOE2016-T3-1-003) and the Singapore Ministry of Education AcRF Tier 2 funding (ACR47/14).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi: https://doi.org/10.1016/j.biocel.2018.05.012.

References

Azriel, R., Gazit, E., 2001. Analysis of the M minimal amyloid-form ing fragment of the letid amyloid polypeptide: an experimental support for the key role of the phenyl an ile residue in amyloid formation. J. Biol. Chem. 276 (36), 34156–34161.
Azzarito, V., Long, K., Murphy, N.S., Wilson, A.J., 2013. Inhibition of alpha-helix-mediated protein-protein interactions using designed molecules. Nat. Chem. 5 (3), 161–173.
Balbach, J.J., Ishti, Y., Anztrutzkin, O.N., Leapman, R.D., Rizzo, N.W., Dyda, F., Reed, J., Tycko, R., 2000. Amyloid fibril formation by AP16-22, a seven-residue fragment of the Alzheimer’s β-amyloid peptide, and structural characterization by solid-state NMR. Biochemistry 39 (45), 13748–13759.
Balbien, M., Grothe, R., Eisenberg, D.S., 2001. An amyloid-forming peptide from the protein Sup35 reveals a dehydrated beta-sheet structure for amyloid. Proc. Natl. Acad. Sci. U.S.A. 98 (55), 2375–2380.
Beisswenger, M., Cabrele, C., 2014. Self-recognition behavior of a helix-loop-helix domain by a fragment scan. Biochim. Biophys. Acta 1844 (9), 1673–1683.
Biancalana, M., Kioe, S., 2010. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. Biochim. Biophys. Acta 1804 (7), 1405–1412.
Borrego, P., Calado, R., Marcelino, J.M., Pereira, P., Quintas, A., Barroso, H., Taveira, N., 2013. An ancestral HIV-2/minam immunodeficiency virus peptide with potent HIV-1 and HIV-2 fusion inhibitor activity. AIDS 27 (7), 1081–1090.
Bosch, B.J., Martina, B.E., Van Der Zee, R., Osterhaus, A.D., Rottier, P.J., 2004. Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides. Proc. Natl. Acad. Sci. U.S.A. 101 (22), 8455–8460.
Bosch, B.J., Van der Zee, R., de Haan, C.A.M., Rottier, P.J.M., 2003. The coronavirus spike protein is a Class I virus fusion protein: structural and functional characterization of the fusion core complex. J. Virol. 77 (16), 8801–8811.
Bowie, J.U., Lutyh, R., Eisenberg, D., 1991. A method to identify protein sequences that display the β/α-parallel helices. Protein Sci. : Publ. Protein Soc. 20 (12), 2125–2129.
Chen, Z., Zhang, L., Qin, C., Bai, L., Cei, Z., Zhang, F., Wei, Q., He, T., Yu, W., Yu, J., Hu, X., Tu, X., Gettie, A., Farzan, M., Yuan, K.-Y., Ho, D.D., 2005. Recombinant modified vaccinia virus Ankara expressing the spike glycoprotein of severe acute respiratory syndrome coronavirus induces protective neutralizing antibodies primarily targeting the receptor binding region. J. Virol. 79 (5), 2679–2688.
Cochran, A.G., Skelton, N.J., Starovasnik, M.A., 2001. Tryptophan zippers: stable, monomeric C6 followed by Dylight 488-conjugated Neutravidin™ alone. C6 detected S1255 expression in 293 T cells, with the staining pattern and intensity similar to that obtained with the anti-S monoclonal antibody (Fig. 4b). As our first step to enhance the binding avidity of peptide C6 to S1188HA, C6 was homodimerized on a biotinylated Lys-based linker molecule, a strategy that has been applied successfully in enhancing the peptide immunogenicity and anti-viral potency (Fig. 5a) (Tam, 1988; Tam et al., 2002). The dimeric C6 peptide, C6-D, formed visible fibril-like aggregates upon four to five days of incubation at 4 °C (Fig. 5b). However, such aggregates were not observed in C6 monomer when incubated for the same duration at 4 °C. Furthermore, in the pull down experiment, peptide C6-D retained approximately 2.5 times higher amount of S1188HA than C6 (Fig. 5c). Collectively, the data indicate that the dimeric C6-D had a higher propensity for oligomerization and a stronger affinity towards the S protein, as compared to the monomeric C6, which warrants further investigation of C6-D as a detection reagent for S protein expression.

4. Conclusions

The data presented in this paper demonstrate that β-zipper-like self-binding peptides, which bind to the protein-of-origin by forming sequence-specific steric β-zippers, could be identified by employing peptide scanning and zipper sequence prediction using programs such as the 3D profile method. Self-binding peptides, such as C6 from SARS-CoV S protein, might mimic certain functions of an antibody for protein detection. With an enhanced peptide-protein interaction avidity and specificity, this approach may be further developed as an animal-free way to rapidly generate and produce peptide-based viral diagnostic and research tools.

Acknowledgements

This research was in part supported by the Singapore Ministry of Education AcRF Tier 3 funding (MOE2016-T3-1-003) and the Singapore Ministry of Education AcRF Tier 2 funding (ACR47/14).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi: https://doi.org/10.1016/j.biocel.2018.05.012.
Jeffers, S.A., Tusell, S.M., Gillim-Ross, L., Hemmila, E.M., Achenbach, J.E., Rebcock, G.J., Thomas Jr., W.D., Thackray, L.B., Young, M.D., Mason, R.J., Ambrosino, D.M., Wentworth, D.E., Demartini, J.C., Holmes, K.V., 2004. CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus. Proc. Natl. Acad. Sci. U. S. A. 101 (44), 15748–15753.

Jiang, S., Lin, K., 1995. Effect of amino acid replacements, additions and deletions on the antiviral activity of a peptide derived from the HIV-1 gp41 sequence. Pept. Res. 8 (6), 345–348.

Kaukinen, P., Valeri, A., Pysunin, A., 2003. Mapping of the regions involved in homotypic interactions of Tula hantavirus N protein. J. Virol. 77 (20), 10910–10916.

Liao, Y., Zhang, S.M., Neo, T.L., Tam, J.P., 2015. Tryptophan-dependent membrane insertion of SARS-CoV spike protein. Biochemistry 54 (9), 1819–1830.

Liu, S., Lu, H., Niu, J., Xu, Y., Wu, S., Jiang, S., 2005. Different from the HIV fusion inhibitor C54, the anti-HIV drug Fuzeon (T-20) inhibits HIV-1 entry by targeting multiple sites in gp41 and gp20. J. Biol. Chem. 280 (12), 11259–11273.

Liu, S., Xiao, G., Chen, Y., He, Y., Niu, J., Escalante, C.R., Xiong, H., Farmar, J., Debnath, A.K., Tien, P., Jiang, S., 2004. Interaction between heptad repeat 1 and 2 regions in spike protein of SARS-associated coronavirus: implications for virus fusogenic mechanisms and identification of fusion inhibitors. Lancet 363 (9413), 938–947.

Lu, Y., Liu, D.X., Tam, J.P., 2008a. Lipid rafts are involved in SARS-CoV entry into Vero E6 cells. Biochem. Biophys. Res. Commun. 369 (2), 344–349.

Lu, Y., Neo, T.L., Liu, D.X., Tam, J.P., 2008b. Importance of SARS-CoV spike protein Trp-rich region in viral infectivity. Biochem. Biophys. Res. Commun. 371 (3), 356–360.

Medinas, R.J., Lambert, D.M., Tompkins, W.A., 2002. C-Terminal gp40 peptide analogs inhibit feline immunodeficiency virus: cell fusion and virus spread. J. Virol. 76 (18), 10979–10986.

Mortola, E., Roy, P., 2004. Efficient assembly and release of SARS-coronavirus-like particles by a heterologous expression system. FEBS Lett. 576 (1–2), 174–178.

Nelson, R., Sawaya, M.R., Balbi-Mirne, M., Madsen, A.O., Riekel, C., Grothe, R., Eisenberg, D., 2005. Structure of the cross-beta spine of amyloid-like fibrils. Nature 435 (7043), 773–778.

Nilsson, M.R., 2004. Functional and structural implications of sequence repeats in fibrous proteins. Adv. Protein Chem. 70, 1–36.

Perutz, M., 1994. Polar zippers: their role in human disease. Protein Sci. : Publ. Protein Soc. 3 (10), 1629–1637.

Reches, M., Porat, Y., Gazit, E., 2002. Amyloid fibril formation by pentapeptide and tetrapeptide fragments of human calcitonin. J. Biol. Chem. 277 (38), 35475–35480.

Rigter, A., Langeveld, J.P., Timmers-Parodi, D., Jacobs, J.G., Moonen, P.L., Bosser, A., 2007. Mapping of possible prion protein self-interaction domains using peptide arrays. BMC Biochem. 8, 6.

Sawaya, M.R., Sambashivan, S., Nelson, R., Ivanova, M.I., Sievers, S.A., Apostol, M.I., Thompson, M.J., Balbi-Mirne, M., Wiltzius, J.J., McFarlane, H.T., Madsen, A.O., Riekel, C., Eisenberg, D., 2007. Atomic structures of amyloid cross-beta spines reveal varied steric zippers. Nature 447 (7143), 453–457.

Spiga, O., Bernini, A., Ciotti, A., Chielini, S., Mencicelli, N., Finetti, F., Causarono, V., Anselmi, F., Priftchi, F., Nicolai, N., 2003. Molecular modelling of S1 and S2 subunits of SARS coronavirus spike glycoprotein. Biochem. Biophys. Res. Commun. 310 (1), 78–83.

Supekar, V.M., Bruckmann, C., Ingallinella, P., Bianchi, E., Pesti, A., Carfi, A., 2004. Structure of a proteolytically resistant core from the severe acute respiratory syndrome coronavirus S2 fusion protein. Proc. Natl. Acad. Sci. 101 (52), 17958–17963.

Tam, J.P., 1988. Synthetic peptide vaccine design: synthesis and properties of a high-density multiple antigenic peptide system. Proc. Natl. Acad. Sci. U. S. A. 85 (15), 5409–5413.

Tam, J.P., Li, Y.A., Yang, J.L., 2002. Antimicrobial dendrimeric peptides. Eur. J. Biochem. / FEBS 269 (3), 923–932.

Teng, P.K., Eisenberg, D., 2009. Short protein segments can drive a non-fibrillising protein into the amyloid state. Protein Eng. Des. Sel. : PEDS 22 (8), 531–536.

Tenidis, K., Waldner, M., Bernhagen, J., Fischle, W., Bergmann, M., Weber, M., Merkle, M.L., Voelter, W., Brunner, H., Kapurniotu, A., 2000. Identification of a pentad- and hexapeptide of islet amyloid polypeptide (IAPP) with amyloidogenic and cytotoxic properties. J. Mol. Biol. 295 (4), 1055–1071.

Thompson, M.J., Sievers, S.A., Karanicolas, J., Ivanova, M.I., Baker, D., Eisenberg, D., 2006. The 3D profile method for identifying fibril-forming segments of proteins. Proc. Natl. Acad. Sci. U. S. A. 103 (11), 4074–4078.

Uttamchandani, M., Yao, S.Q., 2008. Peptide microarrays: next generation bioscips for detection, diagnostics and high-throughput screening. Curr. Pharm. Des. 14 (24), 2428–2438.

Wang, X.J., Li, C.G., Chi, X.J., Wang, M., 2011. Characterisation and evaluation of anti-viral recombinant peptides based on the heptad repeat regions of NDV and IBV fusion glycoproteins. Virology 416 (1–2), 65–74.

Wild, C.T., Shugars, D.C., Greenwell, T.K., McDanal, C.B., Matthews, T.J., 1994. Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. Proc. Natl. Acad. Sci. U. S. A. 91 (21), 9770–9774.

von Bergen, M., Friedhoff, P., Biernat, J., Heberle, J., Mandelkow, E.-M., Mandelkow, E., 2000. Assembly of 14 α protein into Alzheimer paired helical filaments depends on a local sequence motif (306LVQVK311) forming β structure. Proc. Natl. Acad. Sci. 97 (10), 5129–5134.

Wouters, M.A., Curnu, P.M., 1995. An analysis of side chain interactions and pair correlations within antiparallel beta-sheets: the differences between backbone hydrogen-bonded and non-hydrogen-bonded residue pairs. Proteins 22 (2), 119–131.

Xiao, X., Dimitrov, D.S., 2004. The SARS-CoV S glycoprotein. Cell. Mol. Life Sci. : CMLS 61 (19–20), 2428–2430.

Xu, Y., Zhu, J., Liu, Y., Lou, Z., Yuan, F., Liu, Y., Cole, D.K., Ni, L., Su, N., Qin, L., Li, X., Bai, Z., Bell, J.I., Pang, H., Tien, P., Gao, G.F., Rao, Z., 2004. Characterization of the heptad repeat regions, HR1 and HR2, and design of a fusion core structure model of the spike protein from severe acute respiratory syndrome (SARS) coronavirus. Biochemistry 43 (44), 14064–14071.

Yao, Q., Compans, R.W., 1996. Peptides corresponding to the heptad repeat sequence of human parainfluenza virus fusion protein are potent inhibitors of virus infection. Virology 223 (1), 103–112.

Ying, W., Hao, Y., Zhang, Y., Peng, W., Qin, E., Cai, Y., Wei, K., Wang, J., Chang, G., Sun, Dai, S., Li, X., Zhu, Y., Li, J., Wu, S., Guo, L., Dai, J., Wang, J., Wan, P., Chen, Y., Du, C., Li, D., Wan, J., Kuzi, X., Li, W., Shi, R., Wei, H., Cao, C., Yu, M., Liu, H., Dong, F., Wang, D., Zhang, X., Qian, X., Zhu, Q., He, F., 2004. Proteomic analysis on structural proteins of severe acute respiratory syndrome coronavirus. Proteomics 4 (2), 492–504.

Young, J.K., Li, D., Abramowicz, M.C., Morrison, T.G., 1999. Interaction of peptides with sequences from the Newcastle disease virus fusion protein heptad repeat regions. J. Virol. 73 (7), 5945–5956.

Zhang, S.M., Jejicic, A., Tam, J.P., Vahine, A., 2015. Membrane-active sequences within gp41 Membrane Proximal External Region (MPER) modulate MPER-containing peptide fusinhibitor activity and the biosynthesis of HIV-1 structural proteins. PloS One 10 (7), e0134851.