**Structural mechanism for modulation of functional amyloid and biofilm formation
by Staphylococcal Bap protein switch**

Junfeng Ma^{1,3}, Xiang Cheng^{1,3}, Zhonghe Xu¹, Yikan Zhang¹, Jaione Valle², Iñigo Lasa²,
Xianyang Fang^{1,*}

¹Beijing Advanced Innovation Center for Structural Biology, School of Life Sciences,
Tsinghua University, Beijing 100084, China

²Laboratory of Microbial Pathogenesis, Navarrabiomed-Universidad Pública de Navarra-
Departamento de Salud, IDISNA, Pamplona-31008, Navarra, Spain

³These authors contributed equally

*Correspondence: fangxy@mail.tsinghua.edu.cn

Abstract

The Staphylococcal Bap proteins sense environmental signals (such as pH, $[\text{Ca}^{2+}]$) to build amyloid scaffold biofilm matrices via unknown mechanisms. We here report the crystal structure of the aggregation-prone region of *S. aureus* Bap which adopts a dumbbell-shaped fold. The middle module (MM) connecting the N-terminal and C-terminal lobes consists of a tandem of novel double- Ca^{2+} -binding motifs involved in cooperative interaction networks, which undergoes Ca^{2+} -dependent order-disorder conformational switches. The N-terminal lobe is sufficient to mediate amyloid aggregation through liquid-liquid phase separation and maturation, and subsequent biofilm formation under acidic conditions, such processes are promoted by disordered MM at low $[\text{Ca}^{2+}]$ but inhibited by ordered MM stabilized by Ca^{2+} -binding, with inhibition efficiency depending on structural integrity of the interaction networks. These studies illustrate a novel protein switch in pathogenic bacteria and provide insights into the mechanistic understanding of Bap proteins in modulation of functional amyloid and biofilm formation, and development of anti-biofilm therapeutics.

Introduction

Amyloids are insoluble, highly structured protein aggregates (Eisenberg & Sawaya, 2017) best known for their association with debilitating human diseases such as amyotrophic lateral sclerosis (ALS) or Alzheimer's disease (AD) (Chiti & Dobson, 2017). Recently, the amyloids have also been found to perform physiological functions from prokaryotes to humans (Pham, Kwan et al., 2014). In bacteria, functional amyloids are fulfilling diverse roles relevant for bacterial growth and survival in the environment, including as part of the extracellular biopolymer matrix that ties the bacteria together into multicellular communities called biofilm (Van Gerven, Van der Verren et al., 2018). Biofilm formation provides bacteria with many advantages such as high tolerance to harsh environments and increased recalcitrance to antimicrobial agents (Santos, Galdino et al., 2018), thus extremely difficult to eradicate and poses significant challenges to the health care system today (Bjarnsholt, Ciofu et al., 2013, Santos et al., 2018).

Biofilm-associated proteins (Baps) are a group of bacterial surface proteins mainly involved in intercellular adhesion and mediating biofilm development (Lasa & Penades, 2006, Latasa, Solano et al., 2006). Its first member is identified in bovine *S. aureus* which is a large multidomain protein of 2276 amino acids (aa) (Cucarella, Solano et al., 2001) (**Fig 1A**), of which region B contains three putative EF-hand motifs that may regulate Bap functionality upon Ca^{2+} -binding (Arrizubieta, Toledo-Arana et al., 2004). Although *bap* gene has never been found in *S. aureus* human isolates, *bap* ortholog genes are present in the core genomes of several coagulase-negative staphylococcal species that are frequent colonizers of human skin (Tormo, Knecht et al., 2005). Moreover, functionally related proteins homologous to Bap exist in many phylogenetically unrelated Gram-

positive and Gram-negative bacteria (Lasa & Penades, 2006), suggesting its functional importance in a diverse group of bacteria.

Recent results indicate that the N-terminal of Bap is proteolytically processed, released and form amyloid-like aggregates able to induce bacterial biofilm formation under acidic pHs and low concentrations of Ca^{2+} ($[\text{Ca}^{2+}]$). Restoring pH to neutrality reverses the amyloids formed at acidic buffer; addition of Ca^{2+} in the millimolar range inhibits biofilm formation despite acidification of the medium (Taglialegna, Navarro et al., 2016a). These findings suggest that staphylococcal Baps can sense the bacterial environment through pH and $[\text{Ca}^{2+}]$ changes, and modulate biofilm formation in an amyloid-dependent way. However, how Ca^{2+} and pH signals reconcile to regulate Bap-mediated amyloid and biofilm formation is largely unknown.

Here, we report the crystal structure of the aggregation-prone BSP region of *S. aureus* Bap which adopts a dumbbell-shaped fold and exhibits tandem features. We show that the middle module (MM) connecting the N-terminal and C-terminal lobes consists of a tandem of novel double- Ca^{2+} -binding motifs involved in cooperative interaction networks. While the N-terminal lobe is sufficient to mediate amyloid aggregation through liquid-liquid phase separation (LLPS) and maturation, and subsequent biofilm formation under acidic condition, such processes are regulated by Ca^{2+} -dependent order-disorder conformational switches of MM. The disordered MM at low $[\text{Ca}^{2+}]$ promotes LLPS, functional amyloid and biofilm formation, but the ordered MM stabilized by millimolar $[\text{Ca}^{2+}]$ dictates inhibitory effects which efficiency depends on the structural integrity of the interaction networks. Given the wide distribution of Bap orthologs and high conservation of the MM among pathogenic bacteria, our work lays the framework for

mechanistic understanding of Bap proteins and development of antibacterial and anti-biofilm therapeutics.

Results

Overall architecture of the Ca²⁺-bound BSP

Our bioinformatics analysis predicts that the amyloidogenic hotspots in Bap are mainly covered by the B and SP regions (BSP, aa 361-947) (**Supplementary Extended Section, Supplementary Fig 1A**), we thus target BSP for biochemical analysis and determine its crystal structure at 1.9 Å resolution (**Supplementary Fig 1B and Table 1**).

The BSP monomer has an asymmetric dumbbell-shaped structure that can be further divided into five modules (1 to 5): modules 1 (aa 361-541) and 2 (aa 542-728) constitute the N-terminal lobe, which are held together by a middle module (MM) 3 (aa 729-774) followed by the C-terminal lobe consisting of modules 4 (aa 775-859) and 5 (aa 860-947) (**Fig 1A-C**). Interestingly, BSP exhibits tandem features (**Fig 1B-C**). Modules 1 and 2 share a highly similar β -sandwich core architecture and pack closely against each other with their edge side of β -sandwiches at a near perpendicular angle. The middle module 3 folds into a tandem of novel double-Ca²⁺-binding motifs (aa 729-751 and 752-774) which will be discussed in detail in the next section. Modules 4 and 5 possess the same canonical all-beta, immunoglobulin (Ig)-like fold and the tandem modules 4 and 5 exhibit a V-shaped conformation, in which the Ig-like modules stack on each face of β -sheet.

Middle module 3 folds into a tandem of novel double-Ca²⁺-binding motifs

Previously, three potential Ca²⁺-binding EF-hand motifs (designated as EF1: aa 598-610, EF2: aa 729-741, EF3: aa 752-764, respectively) had been predicted in the BSP region of

Bap(Arrizubieta et al., 2004). The crystal structure of BSP monomer reveals five Ca^{2+} ions bound to BSP in modules 1 and 3, respectively (**Fig 1C**). The putative EF1 located in module 2 does not form at all. Instead, one Ca^{2+} -binding site is observed near the ends of two paired strands from module 1 (**Fig 1C**).

Unexpectedly, the putative EF2 and EF3 located in MM form a tandem of novel Loop-Helix-Loop (LHL) motifs (LHL1: 729-751 and LHL2: 752-774, respectively) that each binds two Ca^{2+} ions (**Fig 1C, 2A-B**) instead of the canonical Helix-Loop-Helix (HLH) EF-hand motif that usually binds with one Ca^{2+} ion (**Fig 2C**)(Gifford, Walsh et al., 2007). While each LHL motif consists of 23 residues, six residues shorter than the canonical 29-residue HLH EF-hand motif, the first 12 residues of LHL1 and LHL2 share 91% and 89% similarity in sequence to the loop of the consensus EF-hand motif, respectively (**Fig 2D**). A comparison of the amino acid composition clearly shows that LHL motif is highly hydrophilic and negatively charged and therefore it lacks the hydrophobic core characteristic of canonical-EF hand motif (**Fig 2E**). In LHL1, the first Ca^{2+} binds to the incoming loop in a similar manner as Ca^{2+} binding to the canonical EF-hand loop(Gifford et al., 2007), acidic residues 1, 3, 5, 12 coordinate the Ca^{2+} through one of the sidechain oxygen atoms in a monodentate mode, residue 7 binds to the Ca^{2+} with backbone oxygen, residue 9 immobilizes Ca^{2+} -bound water through sidechain oxygen (**Fig 2A and 2C**). The difference between LHL motif and canonical EF-hand lies in residue 12. This residue donates only one oxygen atom to the first Ca^{2+} in LHL motif whereas it donates two sidechain oxygen atoms to Ca^{2+} in a bidentate mode in the canonical EF-hand loop (**Fig 2A and 2C**). Furthermore, Ca^{2+} -bound water also participates in hydrogen bond to backbone oxygen of residue 20 from exiting loop, therefore forming a hydrogen bond

network along with residue 9. Given that the $O_{9th}-O_{water}-O_{20th}$ angle (110°) is close to the bond-angle of water, and they are almost coplanar with the Ca^{2+} ions, such hydrogen bonds network is expected to be very stable and thus promotes the coordination of water with Ca^{2+} . Along with the backbone oxygen atoms of residues 19 and 22 from the exiting loop, the other side-chain oxygen atoms of residues 3, 5 and 12 from the incoming loop coordinate with the second Ca^{2+} , thus bridging the two Ca^{2+} ions into a coupled Ca^{2+} -pair with a distance as close as ~ 4 Å. The LHL2 motif shares an almost identical backbone (the RMSD is 0.34 Å) to the LHL1 and harbors a pair of Ca^{2+} as observed in LHL1 (**Fig 2B**). Interestingly, Ca^{2+} - Ca^{2+} coupling exists not only within each motif, but between motifs. As shown in **Fig 2A**, residue 9 in LHL2 (D760) which stabilizes the Ca^{2+} -bound water donates another sidechain oxygen atom to the second Ca^{2+} in LHL1. Such interactions are thought to further strengthen the packing between motifs, therefore promote the folding of module 3. Additionally, E849 from module 4 was found to coordinate the second Ca^{2+} in LHL2.

A search for homologous 3D structures in the Protein Data Bank (PDB) using the Dali Server did not reveal any protein with structures similar to MM, highlighting the novelty of the LHL double- Ca^{2+} binding motifs in tandem. A subsequent search in the database of non-redundant protein sequences (nr) showed that homologues to MM are mainly present in the genus *Staphylococcus*, and a few also exists in *Firmicutes* and *Listeria monocytogenes*. Sequence alignment revealed that residues involved in Ca^{2+} binding in each LHL motif as well as their surrounding residues are highly conserved across species, indicating that the integrity of the novel double- Ca^{2+} -binding motifs in tandem could be crucial to its function (**Fig 2D**).

Cooperative interaction networks dictate high affinity Ca^{2+} binding

Complex protein interaction networks are observed among modules 2, 3 and 4 (**Fig 2F-G, Supplementary Extended Section**), for example, besides coordination with the second Ca^{2+} in the LHL2 motif, E849 in module 4 also forms a salt bridge with K758 in the LHL2 motif of module 3 (**Fig 2G**). As module 3 is highly hydrophilic and negatively charged (**Fig 2D-E**), the interactions of module 3 with its neighboring modules 2 and 4 are predominated by hydrogen bonds or electrostatic interactions (**Fig 2F-G**).

To analyze the importance of the interaction networks, interactions of Ca^{2+} with BSP and its subconstructs such as BSP_{12} , BSP_3 , BSP_{123} , BSP_{45} , BSP_{345} and BSP_{1234} , which consist of the first two, the third, the first three, the last two, the last three, the first four modules of BSP, respectively, were studied using isothermal titration calorimetry (ITC) (**Supplementary Fig 2, Table 2**). Consistent with the crystal structure, the isotherm profile of BSP could be best fitted with a 5-site sequential binding model that gave apparent dissociation constant K_{d} s of 0.12 μM , 3.6 μM , 14 μM , 17 μM and 170 μM for the high to low affinity binding sites, respectively, or fitted with 5 equivalent single-site model that gave K_{d} of 29 μM . The isotherm profiles of both BSP_{12} and BSP_{45} could be best fitted with a one-site binding model and 100-fold decreased binding affinities. Surprisingly, BSP_3 was unable to bind with Ca^{2+} at all. The conjunction of module 3 to BSP_{12} or BSP_{45} result in increased Ca^{2+} -binding affinity to BSP_{123} and BSP_{345} relative to BSP_{12} and BSP_{45} , respectively. The concurrence of both modules 2 and 4 with module 3 in BSP_{1234} leads to similar mode and high affinity of Ca^{2+} binding as in BSP. These data suggest cooperative interaction networks among modules 2, 3 and 4 and the importance of its structural integrity in dictating high affinity Ca^{2+} binding. These findings are further

supported by decreased Ca^{2+} -binding affinities for BSP mutants of LHL1m, LHL2m and LHL12m, which Ca^{2+} -binding sites were disrupted with residues corresponding to positions 1, 3 and 12 of LHL1 and LHL2 alone or in combination replaced by Asn, Ala, and Asn, respectively (**Fig 2H**).

The middle module 3 undergoes Ca^{2+} -dependent conformational switches

We further characterized BSP monomer in solution in the presence of 10 mM Ca^{2+} or 10 mM EDTA at pH 7.5 using SAXS (**Supplementary Fig 3A-C**). The experimental scattering profiles were fitted to the crystal structure of BSP monomer, however, neither fits well ($\chi^2 = 24.63$ and $\chi^2 = 2.66$, respectively) (**Fig 3A**), indicating that BSP monomer in solution adopts conformation significantly different from the crystal structure. To dissect which modules cause the difference, SAXS data were collected for all the aforementioned BSP subconstructs under both conditions (**Fig 3A**). Interestingly, the experimental SAXS profiles of BSP₁₂, BSP₁₂₃ and BSP₁₂₃₄ in 10 mM Ca^{2+} agree well with the crystal structures, but not for BSP₄₅. *Ab initio* 3D shape reconstruction by DAMMIN(Svergun, 1999) as well as rigid-body modeling using Xplor-NIH(Schwieters, Kuszewski et al., 2003) suggest an extended conformation for BSP₄₅ in solution, suggesting that the linker connecting module 4 and 5 is flexible in solution (**Supplementary Fig 3D**). The respective SAXS profiles in 10 mM EDTA were also fitted to the corresponding crystal structures and the rigid-body model of BSP₄₅, respectively. While the SAXS profiles of BSP₁₂ and BSP₄₅ in 10 mM EDTA fit nicely to the crystal structure of BSP₁₂ and the rigid-body model of BSP₄₅, respectively, the SAXS profiles of BSP₁₂₃, BSP₁₂₃₄ and BSP which all contain module 3 don't fit to the respective

Ca²⁺-bound crystal structures (**Fig 3A**). These data indicate that module 3 is responsible for the Ca²⁺-dependent conformational changes in BSP and its subconstructs.

The dimensionless Kratky plots for BSP₃ under both conditions exhibit steady increase at larger q (**Fig 3B**), suggesting that BSP₃ alone is intrinsically disordered protein (IDP)(Kikhney & Svergun, 2015). The IDP feature is further supported with heavily overlapped peaks in 2D ¹H-¹⁵N HSQC NMR spectra of BSP₃ at both conditions (**Fig 3C**; **Supplementary Fig 3G-H**). Given that module 3 is well folded in the Ca²⁺-bound BSP structure (**Fig 1C**), Ca²⁺-depletion (in 10 mM EDTA) has a marginal effect on BSP₁₂ conformation but affects BSP₁₂₃, BSP₃₄₅, BSP₁₂₃₄ and BSP significantly (**Fig 3A**; **Supplementary Table 3**), module 3 in such constructs must undergo Ca²⁺-dependent order-disorder conformational switches and the interaction networks is essential for Ca²⁺-induced folding of module 3. This prediction is further supported by the HSQC spectra of BSP₃₄ at both conditions (**Fig 3D**; **Supplementary 3I-J**), which show strong, well-dispersed peaks (ordered) and weak, overlapped peaks (disordered), respectively.

Atomic models for BSP monomer in 10 mM Ca²⁺ or 10 mM EDTA were then built up to fit the experimental SAXS profiles by rigid-body modeling using Xplor-NIH (see **Methods**). The best fitting models can be nicely fitted into the respective *ab initio* 3D shape envelopes (**Supplementary Fig 3E-F**). Notably, the resulting atomic models for BSP₃₄₅ under both conditions agree well with the respective SAXS profiles (**Fig 3A**), further supporting that module 3 undergoes order-to-disorder conformational switch upon Ca²⁺ depletion.

BSP amyloid aggregation is pH-dependent and Ca²⁺-inhibited

The N-terminal region of Bap was previously reported as aggregation-prone and to self-assemble into amyloid structure under acidic pHs and low $[Ca^{2+}]$ (Taglialegna et al., 2016a). To investigate the effects of pH and $[Ca^{2+}]$ on BSP amyloid formation, the aggregation kinetics was measured for BSP monomer at 5 μ M in the pH range of 4.5-7.5 using the Thioflavin T (ThT) fluorescence assay (Arosio, Knowles et al., 2015). BSP monomer exhibited a prominent pH-dependent on/off aggregation switch in the presence of 10 mM EDTA (**Fig 4A**). BSP monomer did not show any obvious amyloid formation over 24 hours at pH > 4.5, but aggregated rapidly at pH 4.5. As expected, the aggregation rate at pH 4.5 was faster at higher protein concentration (2 - 10 μ M) (**Fig 4B**), however, amyloid formation at each protein concentration was inhibited in the presence of 10 mM Ca^{2+} (**Fig 4C**).

Roles of individual modules in BSP amyloid aggregation *in vitro*

The predicted consensus amyloidogenic hotspots (**Supplementary Fig 1A**) were mapped to BSP monomer structure (**Fig 1D**). Interestingly, the hotspots are distributed across all the modules except for module 3. To probe the roles of the respective modules in BSP amyloid formation, the aggregation kinetics of the aforementioned BSP subconstructs in 10 mM EDTA and at pH 4.5 was measured and compared with BSP. As shown in **Fig 4D**, BSP₁₂ exhibited prominent aggregation at 5 μ M, whereas BSP₃, BSP₄₅ did not show any obvious aggregation at 5 μ M, despite amyloidogenic hotspots were predicted on modules 4 and 5. Therefore, the aggregation propensity of BSP could be mainly attributed to modules 1 and 2. Remarkably, BSP₁₂₃ displayed enhanced aggregation compared to BSP₁₂ and BSP₃ alone (**Fig 4E**), but aggregation was not enhanced when BSP₁₂ was mixed with BSP₃ in 1:1 ratio (**Fig 4D**), suggesting that the cooperativity among modules

1, 2 and 3 promotes BSP₁₂₃ amyloid formation. The effects of protein concentrations on aggregation kinetics for all the subconstructs were analyzed. Overall, increasing protein concentration accelerated aggregation of all the subconstructs (**Supplementary Fig 4A**). Moderate aggregations could be observed for BSP₃ and BSP₃₄₅ at 50 μ M, and not for BSP₄₅ until at a concentration of 200 μ M (**Supplementary Fig 4A**). Thus, the aggregation propensity of the Bap subconstructs at pH 4.5 and low [Ca²⁺] can be ranked as BSP₁₂₃ > BSP₁₂₃₄ > BSP > BSP₁₂ >> BSP₃ > BSP₃₄₅ > BSP₄₅ (**Fig 4D-E; Supplementary Fig 4A**). It's interesting to note that the aggregation of both BSP and BSP₃₄₅ was retarded compared to BSP₁₂₃ and BSP₃, respectively, suggesting that modules 4 and 5 play a role in retarding BSP amyloid formation at pH 4.5.

Aggregation kinetics of all the subconstructs were also investigated at elevated pH range of 5.5-7.5 in the presence of 10 mM EDTA. Striking pH-dependent on/off aggregation switches were observed for all the subconstructs containing modules 1 and 2, such as BSP₁₂, BSP₁₂₃, BSP₁₂₃₄, BSP and BSP₃, whereas amyloid fibrils formation by BSP₄₅ and BSP₃₄₅ was not affected by the pH (**Supplementary Fig 4B**), indicating that module 1, 2 and 3 play an essential role in acidic pH-induced BSP amyloid formation.

Next, we investigated the effects of Ca²⁺ on the aggregation of the respective subconstructs at pH 4.5 by measuring the ratio of achieved maximum fluorescence intensities in 10 mM EDTA to 10 mM Ca²⁺ (**Fig 4F**). As shown in **Fig 4F**, the presence of Ca²⁺ exhibited prevailing inhibitory effects on all the subconstructs, interestingly, the strongest inhibitory effects were observed for BSP₁₂₃₄ and BSP, which very likely the cooperative interaction networks among modules 2, 3 and 4 were preserved for the inhibition. Accordingly, the inhibitory effects of Ca²⁺ on aggregation of BSP mutants

(LHL1m, LHL2m and LHL12m) were subtle (**Fig 4F**), in which the cooperative interaction networks among modules 2, 3 and 4 were absent for the inhibition, due to deficient Ca^{2+} binding to MM.

Roles of individual modules in Bap-mediated biofilm formation *in vivo*

To further dissect the roles of the individual modules in Bap-mediated biofilm formation *in vivo*, variants of BSP including BSP₁₂, BSP₁₂₃, BSP, LHL1m, LHL2m and LHL12m were cloned and introduced into *bap*-deficient *S. aureus* (denoted as *S. aureus* Δ bap) (see Methods, **Fig 4G**). The bacteria were grown in media containing glucose, because of the accumulation of acidic by-products from glucose fermentation, the pH drops below 5 when the bacteria reaches stationary. After an overnight incubation V329 wild type strain clearly showed a biofilm adhered to the microtiter plate and bacterial clumps at the bottom and wall of the tube, while Δ bap strain did not (**Fig 4H**). While the empty shutter vector (EV) showed no effects on mediating biofilm formation to *S. aureus* Δ bap, all the BSP variants conferred similar capacity to induce biofilm formation by *S. aureus* Δ bap to that of the wild type *S. aureus* V329 (**Fig 4I**). As all the BSP variants contain modules 1 and 2, these results indicated that the N-terminal lobe of BSP were sufficient to mediate biofilm development *in vivo* under acidic culture conditions. Notably, the presence of 10 mM Ca^{2+} inhibited the biofilm formation mediated by BSP, where the interaction networks among modules 2, 3 and 4 are preserved. In contrast, the inhibitory effects of Ca^{2+} were subtle on biofilm formation capacity of other BSP variants. These data were consistent with the inhibitory effect of Ca^{2+} to amyloid aggregation by BSP variants *in vitro*. Together, our data suggest that Bap amyloid aggregation *in vitro* and Bap-mediated biofilm development *in vivo* under acidic conditions are mainly contributed by its N-

terminal lobe, and the inhibitory effects of Ca^{2+} highly depends on the structural integrity of the interaction networks among modules 2, 3 and 4.

BSP undergoes acidic pH-induced but Ca^{2+} -inhibited LLPS and maturation

To understand the mechanism of acidic pH-induced BSP amyloid aggregation, we studied the LLPS behavior of BSP. The tandem feature of modules 1 and 2 and the IDP property of module 3 is reminiscent of a variety of proteins that undergo LLPS and underlie the pathological amyloid assembly in a variety of neurodegenerative diseases (Aguzzi & Altmeyer, 2016, Babinchak & Surewicz, 2020), but the roles of LLPS in bacteria functional amyloid formation remain largely unexplored.

To examine the phase separation behavior of BSP, BSP was expressed and purified fused to the enhanced green fluorescent protein (eGFP) (BSP_eGFP) through a C-terminal flexible linker (**Fig 5A**). To mimic the crowded environment of extracellular matrix, polyethylene glycol 8000 (PEG 8K) was included into the solutions. Different concentrations of BSP_eGFP protein (2 - 30 μM) were subjected to a wide range of pH values (7.5 - 4.5) and PEG 8K concentrations (2% - 8%), and analyzed by confocal microscopy (**Fig 5B-C; Supplementary 5A**). BSP_eGFP exhibited pH-dependent LLPS behavior and large droplets with fused structures can be distinctly observed at lower pH 4.5 (**Fig 5B**). We next assessed the mobility of BSP_eGFP between the droplet and the bulk phase by fluorescence recovery after photobleaching (FRAP) measurement (**Fig 5D**). FRAP experiment performed 15 min after mixing showed that the majority of the fluorescence signal (~80%) recovered efficiently, suggesting that BSP_eGFP was highly dynamic, with rapid exchange of molecules between the droplets and the surrounding solution. However, the droplets undergo maturation, evidenced by slower rates and

smaller extents of recovery when photobleaching was initiated 2 and 6 hours after mixing, respectively (**Fig 5D**). After 83 days, the material of BSP_eGFP droplets appeared with clustered amyloid-like fibers, consistent with a liquid-to-solid transitions (**Supplementary Fig 5B**).

To gain insights into the roles of individual modules, the phase separation behaviors of BSP subconstructs (BSP₁₂, BSP₃ and BSP₁₂₃) (**Fig 5A**) were characterized (**Fig 5B-C**; **Supplementary Fig 5C**). Similar pH-dependent liquid droplets formation was observed for both BSP₁₂ and BSP₁₂₃. Unexpectedly, no obvious phase-separated liquid droplets was observed for BSP₃ under all the conditions. At pH 5.0 and the same protein concentration, BSP₁₂ formed larger droplets than BSP, but BSP₁₂₃ formed even larger ones (**Fig 5B**). At pH 4.5 and the same protein concentrations, while all of BSP, BSP₁₂ and BSP₁₂₃ formed some fused droplets, the droplet density for BSP₁₂ and BSP₁₂₃ dropped markedly as compared to BSP (**Fig 5B**). These data suggest that LLPS of BSP is mainly related to modules 1 and 2, which can be promoted by the IDP module 3 in the context of BSP₁₂₃, although module 3 alone doesn't undergo phase separation. FRAP experiments for both BSP₁₂ and BSP₁₂₃ also support time-dependent droplets maturation (**Fig 5D**), the recovery efficiencies decrease against the time delays to initiate photobleach after mixing, but the transitions for BSP and BSP₁₂₃ are much faster than BSP₁₂, consistent with the aggregation propensities of the respective constructs.

The effects of Ca²⁺ on LLPS of BSP₁₂, BSP₁₂₃ and BSP were also assayed. The presence of Ca²⁺ strongly inhibited the droplets size and area (**Supplementary Fig 5D**). While the presence of Ca²⁺ generally cause smaller droplet size for BSP₁₂₃ than BSP₁₂ at [Ca²⁺] ranging from 2.5 to 30 mM, no obvious droplets formation can be observed for BSP at

[Ca²⁺] larger than 10 mM (**Fig 5E**). Overall, Ca²⁺-binding causes similar inhibitory effects to the phase-separated liquid droplets formation as that to amyloid aggregation, and the inhibition efficiency also depends on the structural integrity of the interaction networks among modules 2, 3 and 4.

Ca²⁺-binding counteracts pH-induced conformational changes

To provide structural insights into the counteracting effects of Ca²⁺-binding on acidic pH-induced LLPS and amyloid aggregation of BSP, conformational changes of BSP and its subconstructs at various pHs were probed by SAXS. The scattering profiles, the pair distance distribution function PDDF transformed from the scattering profiles, and the structural parameters of R_g and D_{max} derived from PDDFs for BSP and its subconstructs at various pHs are plotted in **Supplementary Fig 6**. In 10 mM EDTA, that the scattering profiles, the PDDFs, the R_g and D_{max} of BSP₁₂ and BSP₄₅ were similar across the varying pHs, by contrast, the scattering profiles, the PDDFs, the R_g and D_{max} of all the constructs containing module 3 such as BSP₃, BSP₁₂₃, BSP₁₂₃₄, BSP₃₄₅ and BSP were prominently affected by varying pH (**Supplementary Fig 6A, 6C**). These results suggest that module 3 plays an essential role in dictating the pH-dependent conformational changes at low [Ca²⁺]. In 10 mM Ca²⁺, the conformations of BSP and all its subconstructs except for BSP₃ were very similar across the varying pHs (**Supplementary Fig 6B, 6C**), suggesting that millimolar [Ca²⁺] were able to counteract the effects of varying pH to BSP and its subconstructs. Notably, the presence of 10 mM Ca²⁺ cause further conformational compaction to all the constructs except for BSP₁₂, BSP₃ and BSP₄₅ under acidic conditions, consistent with disorder-to-order transitions of module 3 upon Ca²⁺ binding in the context of BSP₁₂₃, BSP₃₄₅, BSP₁₂₃₄, and BSP. These data further highlight the

significance of cooperative interaction networks in mediating the counteracting effects of Ca^{2+} -binding on pH-induced conformational changes.

Discussion

A detailed mechanistic understanding of bacteria amyloid formation will be of great value in development of new anti-biofilm therapeutics. In this work, we determined the crystal structure of the aggregation-prone region of *S. aureus* Bap protein and studied its conformational dynamics in response to $[\text{Ca}^{2+}]$ and pH changes, and dissected the roles of the individual modules and the cooperative interaction networks in conferring and regulating BSP amyloid aggregation and biofilm formation. Our results identify modules 1 and 2 of BSP as the scaffold of LLPS and amyloid aggregation, and module 3 as the sensor of pH and $[\text{Ca}^{2+}]$ changes, consistent with the essential role of the N-terminal in *S. aureus* biofilm formation (Taglialegna, Navarro et al., 2016b), allowing us to propose a mechanistic model for staphylococcal Bap in modulation of amyloid aggregation and biofilm formation in response to environmental signals (**Fig 6**).

LLPS has been implicated in a variety of biological processes and pathological protein amyloid aggregation in neurodegenerative diseases (Aguzzi & Altmeyer, 2016, Babinchak & Surewicz, 2020). However, the role of LLPS in the formation of bacteria functional amyloid remains the least explored. So far, only a few examples of phase separation in bacteria have been reported (Al-Husini, Tomares et al., 2018, Monterroso, Zorrilla et al., 2019, Wang, Yan et al., 2019, Yoo, Triandafillou et al., 2019). To our knowledge, *S. aureus* Bap is one of the first bacterial proteins secreted into the extracellular matrices that undergoes LLPS and progressively aggregates into amyloid fibrils by liquid-to-solid maturation (Yoo et al., 2019). Our results thus support an

important role of LLPS and maturation in bacteria functional amyloid formation. Recently, a variety of bacteria proteins including the FapC, TasA and Sbp have been identified to form functional amyloid (Van Gerven et al., 2018). It would be important to decipher whether LLPS plays similar roles in the aggregation process of such proteins.

Our results provide mechanistic insights into amyloidogenesis of Bap proteins. Our structural data reveals that modules 1 and 2 are duplicated folded domains and module 3 is intrinsically disordered region at low $[Ca^{2+}]$, reminiscent of those multivalent proteins producing phase-separated liquid droplets (Banani, Lee et al., 2017). Thus, LLPS of BSP is likely to be driven by multivalent weak interactions between modules 1 and 2. Our data reveal coherent pH-dependence of BSP-related conformational changes, LLPS, amyloid aggregation and biofilm formation at low $[Ca^{2+}]$, establishing a general paradigm on the conversion between soluble protein and its high-order amyloid as well as functional relevance (Wu & Fuxreiter, 2016). The pH-dependent behaviors of BSP can be explained by the changes in surface charges upon varying pHs. BSP contains a total of 89 negatively charged residues (31 glutamic acids and 58 aspartic acids) against 48 positively charged residues (10 arginines and 38 lysines) and 4 histidines, most of these charged residues are exposed, so the surface of BSP at neutral pH is predicted to be negatively charged, likely resulting in Coulombic repulsion. As pH decreases from neutral to acidic, some aspartic acids and glutamic acids are expected to change from negatively charged to neutral, and some histidines are expected to change from neutral to positively charged. This would result in a reduction of surface electrostatic repulsion that would favor conformational compaction, concomitantly, evolving into liquid droplets. However, acidic pH-induced phase-separated liquid droplets are only metastable and

prone to conversion into more stable, solid-like fibrous states over time. Therefore, amyloid aggregation of BSP may proceed from acidic pH-induced conformational compaction to phase-separated liquid droplets, and to fibrous aggregates (Wu & Fuxreiter, 2016). Increased amyloid fibrils formed by Bap at low pHs could help biofilm formation under harsh conditions, which is of central importance to bacteria pathogenesis.

Another important finding in this study is the discovery of a tandem of novel double- Ca^{2+} -binding motifs involved in cooperative interaction networks. Our crystal structure confirms a novel LHL motif in tandem structurally distinct from a variety of Ca^{2+} -binding motifs found in both the eukaryotic and prokaryotic Ca^{2+} binding proteins in recent years, including the canonical EF-hand motif, EF-hand like motifs, Ca^{2+} -binding β -roll motif, Ca^{2+} -binding Greek key motif and the bacterial immunoglobulin-like domains (Dominguez, Guragain et al., 2015). Module 3 acts as a Ca^{2+} and pH sensor through conformational switches. However, the Ca^{2+} -binding affinity and its inhibitory efficiency counteracting pH effects highly depends on the structural integrity of the cooperative interaction networks, suggesting an allosteric pathway to propagate environmental signals (Xiao, Shaw et al., 2017).

Recently, the N-terminal of Esp, a Bap-orthologous protein produced by *Enterococcus faecalis*, displays a similar amyloidogenic behavior as Bap under acidic pH (Taglialegna, Matilla-Cuenca et al., 2020). As Bap is present in coagulase-negative staphylococci, and many other bacteria use Bap-like proteins to build the biofilm matrix, the structural mechanism described here for *S. aureus* Bap may be widespread among pathogenic bacteria and of great value in search for antibacterial and anti-biofilms compounds.

Materials and Methods

Amyloidogenic hotspots prediction

The propensity of the peptide segments across full-length Bap to form amyloid fibrils is analyzed using AMYLPRED2 (<http://aias.biol.uoa.gr/AMYLPRED2/>) server, which employs a consensus of different methods that have been specifically developed to predict features related to the formation of amyloid fibrils. The consensus of short peptide segments in Bap are predicted as in **Supplementary Fig 1A**.

Construct cloning

The DNA fragments encoding Bap_B (amino acids 361-784) and Bap_BSP (amino acids 361-947) of Bap were PCR amplified from the pBT2 plasmid encoding full length Bap of *Staphylococcus aureus* (GenBank: AAK38834.2), then subcloned into pET-28a plasmid with an N-terminal hexa-histidine tag and verified by DNA sequencing. All the other mutants were generated using the TIANGEN's Fast Mutagenesis System.

Protein expression and purification

Plasmids containing the respective constructs were transformed into *E. coli* BL21(DE3) competent cells and cultured in Luria-Bertani media. Protein expression was induced with 0.5 mM IPTG (isopropyl β -D-1-thiogalactopyranoside) when OD600 reached 0.6. Bacteria was further cultured at 37 °C for 4h and then centrifuged at 4,000rpm, 4 °C for 30min. Cell pellet were resuspended with lysis buffer (20 mM Tris-HCl pH 7.4, 500 mM NaCl) and then homogenized by a high-pressure cell disruptor system. 1 mM PMSF (phenylmethyl sulfonylfluoride) was added to lysate during this process. After centrifugation at 18,500 $\times$ g for 60 min, the supernatant containing the His-tagged protein was collected and loaded onto a Ni-NTA affinity resin which was pre-equilibrated with

the binding buffer (20 mM Tris-HCl pH 7.4, 500 mM NaCl). The resin was extensively washed using wash buffer1 (20 mM Tris-HCl pH 7.4, 1 M NaCl, 20 mM imidazole) and wash buffer2 (20 mM Tris-HCl pH 7.4, 200 mM NaCl, 20 mM imidazole) successively. Target protein was eluted using buffer containing 20 mM Tris-HCl pH 7.4, 100 mM NaCl, 250 mM imidazole. His-tag was cleaved by TEV protease which then was removed by anion-exchange chromatography by using HiTrap Q column (GE Healthcare). Preparative size-exclusion chromatography (SEC) was performed for further purification using HiLoad 16/600 Superdex 75 or 200 column (GE Healthcare) depending on the molecular weights of desired proteins. Fractions containing target protein were visualized with SDS-PAGE gel and pooled together. Selenomethionine-labeled Bap_B was overexpressed in *E. coli* BL21(DE3) B834 strain cultured with selenomethionine-supplemented M9 medium as described by the manufacturer (Molecular Dimensions), and purified using the same protocol as above.

Crystallization, data collection and structure determination

Proteins prior to crystallization trial were concentrated to about 20 mg/ml in buffer consisting of 20 mM Tris-HCl pH 7.4, 100 mM NaCl, 10 mM CaCl₂. Both initial screening and optimization were performed using the sitting-drop or hanging-drop vapor diffusion method. The crystallization optimization of Bap_BSP protein was performed by mixing 1.3 µl protein with 1 µl crystallization buffer containing 0.1 M Sodium Citrate pH 5.0, 10% 2-propanol, 26% PEG 400 at 16 °C. The crystallization optimization of Bap_B protein was performed by mixing 3 µl protein with 1 µl crystallization buffer consisting of 0.1 M Sodium Cacodylate pH 6.5, 40% MPD, 5% PEG 8000 at 16 °C. The crystallization optimization of Selenomethionine-labeled Bap_B protein were performed

by mixing 2.5 ul protein with 1 ul crystallization buffer containing 0.1 M Sodium Cacodylate pH 6.5, 38% MPD, 5% PEG 8000. For cryoprotection, the single crystals were fished with nylon loops gently and dipped into crystallization mother liquor for several seconds before transferred into liquid nitrogen.

X-ray diffraction data from all the crystals were collected at the wavelength of 0.9792 Å and temperature of -170 °C at the BL17U and BL19U beamlines of Shanghai Synchrotron Radiation Facility (SSRF). The data were indexed, integrated, merged and scaled with HKL2000 (<http://www.hkl-xray.com>) or X-ray Detector Software (XDS). The phasing of the Bap_B crystal structure was performed by the selenium single-wavelength anomalous dispersion method using Phaser from the program PHENIX and CCP4. For the Bap_BSP crystal structure, the Bap_B structure was used as a search model to determine the phase. Initial models were built using the AutoBuild wizard in PHENIX, these models were further refined by iterative cycles of manual model building using COOT and refinement against initial X-ray diffraction data with PHENIX. Validation of the refined models was carried out using MolProbity on the PHENIX platform. Ramachandran analysis of the backbone dihedral angles indicated that all residues are located in the most favorable and generally allowed regions. The data collection and structural refinement statistics are summarized in **Supplementary Table 1**.

Isothermal titration calorimetry

All the ITC measurements were performed with a MicroCal ITC 200 calorimeter (GE Healthcare) at 25 °C and the data were processed with the Origin7 software package from MicroCal. Purified Bap_BSP and its sub-constructs were extensively exchanged into buffer of 50 mM Tris-HCl pH 7.5, 100 mM NaCl using a Superdex 75 or 200 10/300 GL

column (GE Healthcare), then concentrated to 20-800 μM and placed in the cell. The syringe was filled with 2-10 mM CaCl_2 solution. Measurements were repeated twice, and similar results were obtained. The background data obtained from the buffer blank titration were subtracted before the data analysis.

Small Angle X-ray Scattering

The data collection and processing procedures are similar as described before (Zhang, Zhang et al., 2019). Briefly, SAXS measurements were carried out at room temperature at the beamline 12 ID-B of the Advanced Photon Source, Argonne National Laboratory or the beamline BL19U2 of the Shanghai Synchrotron Radiation Facility (SSRF). The scattered X-ray photons were recorded with a PILATUS 2M detector (Dectris) at 12 ID-B and a PILATUS 100k detector (Dectris) at BL19U2. The setups were adjusted to achieve scattering q values of $0.005 < q < 0.89 \text{ \AA}^{-1}$ (12ID-B) or $0.009 < q < 0.415 \text{ \AA}^{-1}$ (BL19U2), where $q = (4\pi/\lambda)\sin\theta$, and 2θ is the scattering angle. Thirty two-dimensional images were recorded for each buffer or sample solution using a flow cell. The 2D images were reduced to one-dimensional scattering profiles using Matlab (12ID-B) or BioXTAS Raw (BL19U2). Scattering profiles of the proteins were calculated by subtracting the background buffer contribution from the sample-buffer profile using the program PRIMUS (Konarev, Volkov et al., 2003) following standard procedures. Concentration series measurements (4- and 2- fold dilution and stock solution) for the same sample were carried out to remove the scattering contribution due to inter-particle interactions and to extrapolate the data to infinite dilution. The forward scattering intensity $I(0)$ and the radius of gyration (R_g) were calculated from the data of infinite dilution at low q values in the range of $qR_g < 1.3$, using the Guinier approximation:

$\ln I(q) \approx \ln(I(0)) - R_g^2 q^2 / 3$. These parameters were also estimated from the scattering profile with a broader q range of 0.006-0.30 $\text{\AA}^{-1}$ using the indirect Fourier transform method implemented in the program GNOM (Svergun, 1992), along with the pair distance distribution function (PDDF), $p(r)$, and the maximum dimension of the protein, $D_{\max}$. The Volume-of-correlation (V_c) were calculated using the program Scatter, and the molecular weights of solutes were calculated on a relative scale using the R_g/V_c power law developed by Rambo et al (Rambo & Tainer, 2013), independently of protein concentration and with minimal user bias. The theoretical scattering intensity of the atomic structure model was calculated and fitted to the experimental scattering intensity using CRY SOL (Svergun, Barberato et al., 1995).

Low resolution bead models were built up with the program DAMMIN, which generate models represented by an ensemble of densely packed beads (Svergun, 1999), using scattering data within the q range of 0.006 -0.30 $\text{\AA}^{-1}$. 32 independent runs were performed, and the resulting models were subjected to averaging by DAMAVER (Volkov & Svergun, 2003) and were superimposed by SUPCOMB (Kozin & Svergun, 2001) based on the normalized spatial discrepancy (NSD) criteria and were filtered using DAMFILT to generate the final model. NSD is a measure of quantitative similarity between sets of three dimensional models, if two models systematically differ from each other, their NSD exceeds 1, for identical objects, it is 0.

With the available coordinates for the individual modules, rigid-body modeling was carried out to construct atomic models for Bap_BSP in 10 mM Ca^{2+} or 10 mM EDTA using Xplor-NIH package (Schwieters et al., 2003), during which the respective individual modules were kept as rigid bodies (module 1-4 and module 5 as two rigid

bodies for BSP in 10 mM Ca^{2+} , module 1-2, module 4 and module 5 as three rigid bodies for BSP in 10 mM EDTA), the linkers between were allowed to translate or rotate freely, and a simulated annealing algorithm were performed to optimize the best position and orientation of the individual modules against SAXS data.

NMR Spectroscopy

To make uniformly ^{15}N -labeled BSP₃ or BSP₃₄ samples for NMR analysis, cells were grown in 2L LB medium at 37 °C until A_{600} was ~ 0.6. Cells were then spun down, washed with 1×M9 salt, and transferred into 1 liter of M9 minimal medium containing $^{15}\text{NH}_4\text{Cl}$ as the sole nitrogen sources to produce ^{15}N -labeled proteins. The samples used for NMR experiment are as follows: 0.5 mM ^{15}N -labeled BSP₃ or BSP₃₄ in buffers containing 10% D_2O (v/v), 20 mM $\text{K}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$, 100 mM NaCl, 50 μM NaN₃, pH 7.20, with or without additional 10 mM Ca^{2+} . 2D ^1H - ^{15}N HSQC spectra were acquired at 25 °C on a Bruker 600 MHz spectrometer equipped with a cryogenic probe (Bruker BioSpin GmbH, Germany). All data were processed using the software package NMRPipe (Delaglio, Grzesiek et al., 1995) and analyzed with NMRView (Johnson, 2004).

Structural illustration

All of the illustrations of atomic models were generated using the PyMOL Molecular Graphics System. The bead models generated by DAMMIN were transformed into volumetric maps using Situs (Wriggers, 2012) and displayed using Chimera (Pettersen, Goddard et al., 2004).

Aggregation kinetics

Protein samples were diluted by different pH buffer with 10 mM EDTA or 10 mM Ca^{2+} to the final concentration. 100 μM thioflavin T (ThT) was added into samples before using PerkinElmer Enspire plate reader monitoring fluorescence during amyloidogenesis. The protein samples were shaken at 800 rpm under the temperature of 25 °C. Every 15 minutes, the fluorescence intensity was measured through the bottom of the plate with excitation filter of 440 nm and emission filter of 480 nm, to record the fluorescence intensity results at this time. At the same time the buffer without protein was measured as the background ThT fluorescence, so the fluorescence intensity of the protein sample was divided by the fluorescence intensity of the ThT-only sample after subtracting initial fluorescence intensity value, to calculate fluorescence fold change.

Construction of shuttle vectors

The signal peptide (SP), the R region of clumping factor A gene containing an LPXTG motif were cloned from the pCN51 plasmid, the different domains of the Bap protein were cloned from the BSP-pET-28a plasmid. Plasmid DNA was isolated from *E. coli* or *S. aureus* RN4220 strains whose cells were lysed by lysostaphin (12.5 $\mu\text{g}/\text{ml}$, Sigma-Aldrich) at 37 °C for 2 h before plasmid purification, then transformed into *Staphylococcus aureus* by electroporation, using a previously described procedure (Cucarella et al., 2001).

Biofilm formation assay *in vivo*

Briefly, *S. aureus* strains were grown overnight in TSB at 37 °C. The culture was diluted 1:40 in TSB containing 0.15% glucose, 10 $\mu\text{g}/\text{ml}$ erythromycin, 1 μM CdCl_2 , 10 mM or 20 mM CaCl_2 . 200 μl of this diluted culture was used to inoculate sterile 96-well polystyrene microtiter TC-plates (Corning). After 24 h, the wells were gently washed twice by submerging in sterile water, dried overnight, and stained with 0.1% crystal

violet for 10 min. Wells were gently rinsed twice with sterile water, dried and eluted with 33% acetic acid, then the absorbance was determined at 595 nm (Perkin EnVision instruments). Each assay was repeated three times.

***In vitro* phase separation assay**

In vitro phase separation assay was performed in different buffer mixed with PEG 8000 (Sigma). Droplets were assembled in 384 low-binding multi-well microscopy glass plates (384-well microscopy plates), sealed with optically clear adhesive film to prevent evaporation and observed under a Nikon A1 microscope equipped with 100× oil immersion objectives.

Fluorescence recovery after photobleaching (FRAP)

For the *in vitro* experiments, FRAP was carried out with samples in 384-well microscopy plates using a Nikon A1 microscope equipped with 100× oil immersion objectives. Droplets were bleached with a 488-nm laser pulse (3 repeats, 100% intensity, dwell time 0.5 s). FRAP analysis involved assessment of mean fluorescence intensity from each bleached region of interest as well as non-bleached regions and background to correct for microscope drift. The recovery from photobleaching was recorded for the indicated time.

Additional information

Additional Information including **Supplementary Figures 1-6** and **Supplementary Tables 1-3** can be found online.

Acknowledgements

We thank the staffs at beamlines BL17U1, BL19U1 and BL19U2, Shanghai Synchrotron Radiation Facility, China and Dr. Xiaobing Zuo at the beamline 12-ID-B, Advanced

Photon Source, USA for assistances during data collection. We thank Prof. Jiawei Wang at Tsinghua University in X-ray diffraction data analysis, Prof. Jing-Ren Zhang at Tsinghua University for providing *S. aureus* RN4220. This work was supported by grants from the National Natural Science Foundation of China (No. 31872712), the National Key Research and Development Project of China (2016YFA0500700), the Beijing Advanced Innovation Center for Structural Biology, the Tsinghua-Peking Joint Center for Life Sciences to X.F..

Author contributions

X. F. conceived and supervised the project. J. M., X. C. performed the majority of the experiments and analyzed the data. J. M., Z. X. contributed to the crystal structure analysis, Y. Z. contributed to SAXS and NMR data analysis, J. V., I. L. provided Bap plasmid, *S. aureus* strains and edited the manuscript. X. F. wrote the paper with inputs from all authors.

Competing interests

The authors declare no competing interests.

REFERENCES

- Aguzzi A, Altmeyer M (2016) Phase Separation: Linking Cellular Compartmentalization to Disease. *Trends Cell Biol* 26: 547-558
- Al-Husini N, Tomares DT, Bitar O, Childers WS, Schrader JM (2018) alpha-Proteobacterial RNA Degradosomes Assemble Liquid-Liquid Phase-Separated RNP Bodies. *Mol Cell* 71: 1027-1039 e14
- Arosio P, Knowles TP, Linse S (2015) On the lag phase in amyloid fibril formation. *Phys Chem Chem Phys* 17: 7606-18
- Arrizubieta MJ, Toledo-Arana A, Amorena B, Penades JR, Lasa I (2004) Calcium inhibits bap-dependent multicellular behavior in *Staphylococcus aureus*. *J Bacteriol* 186: 7490-7498
- Babinchak WM, Surewicz WK (2020) Liquid-Liquid Phase Separation and Its Mechanistic Role in Pathological Protein Aggregation. *J Mol Biol* 432: 1910-1925
- Banani SF, Lee HO, Hyman AA, Rosen MK (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* 18: 285-298

Bjarnsholt T, Ciofu O, Molin S, Givskov M, Hoiby N (2013) Applying insights from biofilm biology to drug development - can a new approach be developed? *Nat Rev Drug Discov* 12: 791-808

Chiti F, Dobson CM (2017) Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu Rev Biochem* 86: 27-68

Cucarella C, Solano C, Valle J, Amorena B, Lasa I, Penades JR (2001) Bap, a *Staphylococcus aureus* surface protein involved in biofilm formation. *J Bacteriol* 183: 2888-96

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6: 277-93

Dominguez DC, Guragain M, Patrauchan M (2015) Calcium binding proteins and calcium signaling in prokaryotes. *Cell Calcium* 57: 151-65

Eisenberg DS, Sawaya MR (2017) Structural Studies of Amyloid Proteins at the Molecular Level. *Annu Rev Biochem* 86: 69-95

Gifford JL, Walsh MP, Vogel HJ (2007) Structures and metal-ion-binding properties of the Ca²⁺-binding helix-loop-helix EF-hand motifs. *Biochem J* 405: 199-221

Johnson BA (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol* 278: 313-52

Kikhney AG, Svergun DI (2015) A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett* 589: 2570-7

Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI (2003) PRIMUS - a Windows-PC based system for small-angle scattering data analysis. *J Appl Cryst* 36: 1277-1282

Kozin MB, Svergun DI (2001) Automated matching of high- and low-resolution structural models. *Journal of Applied Crystallography* 34: 33-41

Lasa I, Penades JR (2006) Bap: a family of surface proteins involved in biofilm formation. *Res Microbiol* 157: 99-107

Latasa C, Solano C, Penades JR, Lasa I (2006) Biofilm-associated proteins. *C R Biol* 329: 849-57

Monterroso B, Zorrilla S, Sobrinos-Sanguino M, Robles-Ramos MA, Lopez-Alvarez M, Margolin W, Keating CD, Rivas G (2019) Bacterial FtsZ protein forms phase-separated condensates with its nucleoid-associated inhibitor SlmA. *EMBO Rep* 20

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25: 1605-12

Pham CL, Kwan AH, Sunde M (2014) Functional amyloid: widespread in Nature, diverse in purpose. *Essays Biochem* 56: 207-19

Rambo RP, Tainer JA (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* 496: 477-+

Santos A, Galdino ACM, Mello TP, Ramos LS, Branquinho MH, Bolognese AM, Columbano Neto J, Roudbary M (2018) What are the advantages of living in a community? A microbial biofilm perspective! *Mem Inst Oswaldo Cruz* 113: e180212

Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* 160: 65-73

Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Cryst* 25: 495-503

Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* 76: 2879-2886

Svergun DI, Barberato C, Koch MHJ (1995) CRY SOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Cryst* 28: 768-773

Taglialegna A, Matilla-Cuenca L, Dorado-Morales P, Navarro S, Ventura S, Garnett JA, Lasa I, Valle J (2020) The biofilm-associated surface protein Esp of *Enterococcus faecalis* forms amyloid-like fibers. *NPJ Biofilms Microbiomes* 6: 15

Taglialegna A, Navarro S, Ventura S, Garnett JA, Matthews S, Penades JR, Lasa I, Valle J (2016a) Staphylococcal Bap Proteins Build Amyloid Scaffold Biofilm Matrices in Response to Environmental Signals. *Plos Pathog* 12

Taglialegna A, Navarro S, Ventura S, Garnett JA, Matthews S, Penades JR, Lasa I, Valle J (2016b) Staphylococcal Bap Proteins Build Amyloid Scaffold Biofilm Matrices in Response to Environmental Signals. *PLoS Pathog* 12: e1005711

Tormo MA, Knecht E, Gotz F, Lasa I, Penades JR (2005) Bap-dependent biofilm formation by pathogenic species of *Staphylococcus*: evidence of horizontal gene transfer? *Microbiology* 151: 2465-75

Van Gerven N, Van der Verren SE, Reiter DM, Remaut H (2018) The Role of Functional Amyloids in Bacterial Virulence. *J Mol Biol* 430: 3657-3684

Volkov VV, Svergun DI (2003) Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography* 36: 860-864

Wang H, Yan X, Aigner H, Bracher A, Nguyen ND, Hee WY, Long BM, Price GD, Hartl FU, Hayer-Hartl M (2019) Rubisco condensate formation by CcmM in beta-carboxysome biogenesis. *Nature* 566: 131-+

Wriggers W (2012) Conventions and workflows for using Situs. *Acta Crystallogr D Biol Crystallogr* 68: 344-51

Wu H, Fuxreiter M (2016) The Structure and Dynamics of Higher-Order Assemblies: Amyloids, Signalosomes, and Granules. *Cell* 165: 1055-1066

Xiao Y, Shaw GS, Konermann L (2017) Calcium-Mediated Control of S100 Proteins: Allosteric Communication via an Agitator/Signal Blocking Mechanism. *J Am Chem Soc* 139: 11460-11470

Yoo H, Triandafillou C, Drummond DA (2019) Cellular sensing by phase separation: Using the process, not just the products. *J Biol Chem* 294: 7151-7159

Zhang Y, Zhang Y, Liu ZY, Cheng ML, Ma J, Wang Y, Qin CF, Fang X (2019) Long non-coding subgenomic flavivirus RNAs have extended 3D structures and are flexible in solution. *EMBO Rep*: e47016

Figure and legend

Figure 1. Overall architecture of aggregation-prone region of *S. aureus* Bap.

- A** Schematic domain organization of *S. aureus* Bap. S, signal peptide (aa 1-44); A, region A (aa 45-360); B, region B (aa 361-818); SP, spacer region (aa 819-947); C, repeat region (aa 948-2139); D, serine-aspartate repeats region (aa 2148-2208); W, cell wall anchor region. The aggregation-prone regions of BSP consists of 5 modules: 1 (magenta), 2 (orange), 3 (green), 4 (blue) and 5 (cyan) based on its crystal structure.
- B** A topology diagram of BSP.
- C** Overall structure of Ca^{2+} -bound BSP monomer. Ca^{2+} ions are shown as gray spheres. Color code is the same in (a-c).
- D** The consensus amyloidogenic hotspots (red) predicted with AMYLPRED2 are mapped on BSP structure (light blue).

Figure 2. Structural details of the novel double- Ca^{2+} -binding motifs in tandem.

- A** Ribbon diagram of module 3 showing a tandem of double- Ca^{2+} -binding motifs with loop-helix-loop (LHL) topology (LHL1 (limegreen) and LHL2 (limon)). Residues involved in Ca^{2+} coordination are highlighted in stick representation. The Ca^{2+} ions and Ca^{2+} -bound water are shown in magenta and red spheres, respectively.
- B** Superimposition of the LHL1 and LHL2 motifs together with the Ca^{2+} ions, with a backbone RMSD of $\sim 0.3 \text{ \AA}$.
- C** Ribbon diagram of a canonical EF-hand (PDB: 2NXQ) and details of its Ca^{2+} -binding sites.
- D** Sequence alignment of module 3 orthologs using ESPript 3.0. NCBI accession

numbers for the respective bacteria are given in parentheses. The positions of secondary structure elements are indicated above the sequences (α represents α -helix).

- E** Amino acid composition comparison of the LHL motifs and a canonical EF-hand (PDB: 1J7O).
- F-G** The interaction networks between modules 3 and 2 (**F**) as well as between modules 3 and 4 (**G**).
- H** Schematic diagram of LHL1m, LHL2m and LHL12m mutants, which have D1N, D3A, E12N mutations in LHL1 and/or LHL2.

Figure 3. The middle module 3 undergoes Ca^{2+} -dependent order-disorder conformational switches.

- A** Experimental SAXS profiles of monomeric BSP and its subconstructs such as BSP_{12} , BSP_{123} , BSP_{1234} , BSP_{45} , BSP_{345} in 10 mM Ca^{2+} (green) and 10 mM EDTA (red) are fitted to the crystal structures directly (black, 10 mM Ca^{2+} ; blue, 10 mM EDTA) or to the rigid-body models refined using Xplor-NIH against SAXS profiles (magenta, 10 mM Ca^{2+} ; cyan, 10 mM EDTA). The residual is defined as $I_{\text{calc}}/I_{\text{exp}}$.
- B** The dimensionless Kratky plots of SAXS profiles of BSP_3 in 10 mM Ca^{2+} (red) and 10 mM EDTA (black).

C-D Overlays of 2D ^1H - ^{15}N heteronuclear single quantum coherence (HSQC) NMR spectra of BSP₃ (**C**) and BSP₃₄ (**D**) in 10 mM Ca^{2+} (red) and 10 mM EDTA (black).

Figure 4. Amyloid aggregation *in vitro* and biofilm formation *in vivo* by BSP.

A-C BSP amyloid aggregation *in vitro* is pH-dependent in 10 mM EDTA (**A**), protein concentration-dependent in 10 mM EDTA and at pH 4.5 (**B**) and strongly inhibited by millimolar Ca^{2+} at various protein concentrations (**C**).

D-E Amyloid aggregation by BSP subconstructs. BSP₁₂ exhibits more prominent aggregation propensity than BSP₃ and BSP₄₅ in 10 mM EDTA and at pH 4.5 (**D**). The conjunction of module 3 with modules 1 and 2 promotes aggregation (**E**).

- F** Millimolar Ca^{2+} inhibits aggregation of BSP and its subconstructs. The inhibition efficiency is defined as the ratio of maximum fluorescence intensity achieved in 10 mM EDTA to in 10 mM Ca^{2+} after 24 h incubation.
- G** Structural organization of BSP chimeric subconstructs. S (red), signal peptide; module 1 (magenta); module 2 (orange); module 3 (green); module 4 (blue); module 5 (cyan); R-ClfA (yellow), clumping factor R domain; W (purple), cell wall anchor.
- H** Bacterial clumping and biofilm formation of *S. aureus* V329 and the corresponding *S. aureus* Δbap strain. For bacterial clumping (bottom), bacteria were cultured overnight with agitation (220 rpm) in LB-glu (pH<5). For biofilm formation (top), bacteria were cultured overnight at 37°C in TSB-glu (pH<5) using sterile 96-well polystyrene microtiter TC-plates under static conditions.
- I** Biofilm formation of *S. aureus* Δbap mutants expressing BSP chimeric subconstructs in the absence (cyan) or presence (red) of Ca^{2+} assayed with crystal violet staining.

Figure 5. Liquid-liquid phase separation and maturation of BSP is regulated by pH and Ca^{2+} .

A Schematic of BSP constructs with eGFP fusions.

- B** Representative fluorescence microscopy images of BSP and its subconstructs (20 μ M) at varying pHs. PEG 8K concentration is 2 % for BSP, BSP₁₂, BSP₁₂₃ and 8 % for BSP₃. Scale bar, 20 μ m.
- C** Phase diagram of 20 μ M proteins as a function of pH and PEG 8K concentration, which was scored by the presence or absence of droplets in the samples.
- D** Time lapse FRAP of liquid droplets formed by fluorescently labeled BSP, BSP₁₂ and BSP₁₂₃. Bleaching was performed at the indicated time points after mixing, and the recovery occurring was recorded. Data are reported as mean $\pm$ SD. SD is depicted as shadows in a light color. $n = 3$ liquid droplets. n.s., not significant ($P > 0.05$); *** $P < 0.001$; [one-way analysis of variance (ANOVA)].
- E** Quantification of droplet area is shown for each protein under various Ca²⁺ concentration conditions. Each sample is quantified by total area (left) and mean area (right).

Figure 6. Mechanistic model for staphylococcal Bap proteins to build amyloid scaffold biofilm matrices in response to environmental changes.

After secretion, Bap is proteolytically processed, and the N-terminal containing regions A and B is released. Module 3 acts as a pH and Ca^{2+} sensor by significant conformational switches. When the pHs become acidic and $[\text{Ca}^{2+}]$ is low, modules 1 and 2 of BSP act as scaffold and aggregate into amyloid fibrils over time through LLPS and maturation, contributing to subsequent staphylococcal biofilm formation, which process can be promoted by intrinsically disordered module 3. When $[\text{Ca}^{2+}]$ reaches millimolar in the ECM, Ca^{2+} -binding to module 3 stabilizes its ordered folding and the interaction networks among modules 2, 3 and 4, resulting in reduced conformational dynamics across BSP, hindering its self-assembly and aggregation.