ARTICLE

Check for updates

# Self-cyclisation as a general and efficient platform for peptide and protein macrocyclisation

Xinying Jia [1,4✉], Yanni K.-Y. Chin[1,4], Alan H. Zhang[1], Theo Crawford [1], Yifei Zhu[1], Nicholas L. Fletcher [1], Zihan Zhou[2], Brett R. Hamilton[1,3], Martin Stroet [2], Kristofer J. Thurecht[1] & Mehdi Mobli [1✉]

Macrocyclisation of proteins and peptides results in a remarkable increase in structural stability, making cyclic peptides and proteins of great interest in drug discovery—either directly as drug leads or as in the case of cyclised nanodiscs (cNDs), as tools for studies of trans-membrane receptors and membrane-active peptides. Various biological methods have been developed that are capable of yielding head-to-tail macrocyclised products. Recent advances in enzyme-catalysed macrocyclisation include discovery of new enzymes or design of new engineered enzymes. Here, we describe the engineering of a self-cyclising "auto-cyclase" protein, capable of performing a controllable unimolecular reaction for generation of cyclic biomolecules in high yield. We characterise the self-cyclisation reaction mechanism, and demonstrate how the unimolecular reaction path provides alternative avenues for addressing existing challenges in enzymatic cyclisation. We use the method to produce several notable cyclic peptides and proteins, demonstrating how autocyclases offer a simple, alternative way to access a vast diversity of macrocyclic biomolecules.

[1] Centre for Advanced Imaging, Australian Institute for Bioengineering & Nanotechnology, The University of Queensland, St. Lucia, QLD 4072, Australia. [2] School of Chemistry and Molecular Biosciences, The University of Queensland, St. Lucia, QLD 4072, Australia. [3] Centre for Microscopy and Microanalysis, The University of Queensland, St. Lucia, QLD 4072, Australia. [4] These authors contributed equally: Xinying Jia, Yanni K.-Y. Chin. ✉email: jiaxinying@gmail.com; m.mobli@uq.edu.au

Head-to-tail macrocyclisation is a naturally occurring post-translational modification that stabilises the protein fold, leading to enhanced thermal stability and resistance to proteolytic digestion by exoproteases[1–4]. Mimicry of this natural phenomenon has and continues to inspire protein engineering efforts[5,6] including macrocyclic peptide drug leads[7] and highly stable cyclised lipid nanodiscs used in biophysical characterisation of membrane proteins and membrane active peptides[8,9] as well as large protein complexes and biological process, including viral entry, synaptic vesicle fusion, lipid interactions and exocytosis[10–13].

The most common strategy to achieve macrocyclisation is by ligation of the termini of the peptide chain through a peptide bond. Chemically this can be achieved efficiently through native chemical ligation (NCL)[14]. NCL, however, becomes challenging for peptides and proteins longer than 100 amino acids. Consequently, several size-insensitive biological approaches have been developed, such as expressed protein ligation (EPL)[15–18], split-intein mediated protein trans-splicing[19], and genetic-code reprogramming[20]. EPL and split-intein mediated protein cyclisation both require a free cysteine in the sequence to perform an N-S acyl shift and trans-thioesterification, while backbone cyclisation via codon reprogramming requires the introduction of at least one nonproteinogenic amino acid.

The backbones of peptides and proteins are naturally cyclised by a certain group of proteases with unusual enzymatic transpeptidation activity (as a cyclase or ligase). To date, four stand-alone and ATP-independent ligases have been identified: (i) the bacterial transpeptidases, including sortase A[21,22]; (ii) ligases derived from trypsin (trypsiligase)[23]; (iii) ligases derived from subtilisin (subtiligase[24], peptiligase[25], omniligase-1[26]); and (iv) a number of plant-derived ligase-type asparaginyl endopeptidases (AEPs)[27–30] including butelase-1[31] and a number of OaAEPs[32–34]. Of these, the bacterial enzyme, sortase A (SrtA)[21,22] is perhaps the most popular ligase for peptide and protein cyclisation.

Conventional enzymatic protein or peptide cyclisation reactions are achieved through a bimolecular reaction where both components are present at high concentrations. While very powerful, these bimolecular reactions involve competition between the cyclisation reaction and a polymerisation reaction (joining the tail of one protein to the head of another), which can lead to reduced yields, in particular for target proteins prone to self-association[35,36]. Minimising the risk of polymerisation can be achieved by lowering the protein concentration, but doing so also results in a drop in cyclisation rates and makes the overall process particularly challenging to scale[9,35]. Indeed, optimisation of protein and enzyme concentrations remains an important and time-consuming step in enzymatic ligation reactions—leading to various innovations in reaction/reactor design[8,35]. These innovations, however, do not address the root of the polymerisation problem, which is the dual constraint of maximising a desired diffusion-limited (herein always referring to lateral diffusion) step, while minimising an unwanted diffusion-limited step. The desired step being that of enzyme-substrate intermediate formation and the unwanted step being that of the enzyme-substrate intermediate reacting with a second substrate, thereby initiating polymersation.

A particularly attractive method to overcome the challenges of the conventional bimolecular enzymatic cyclisation reaction would be to make the entire process non-diffusion limited. This can in principle be achieved by fusing the ligase to the substrate (target protein), which would fundamentally alter the cyclisation reaction mechanism from a diffusion-limited bimolecular reaction to a non-diffusion-limited unimolecular reaction. The first order kinetics of the unimolecular reaction allows the sample concentration to be lowered without affecting the reaction half-life. Indeed, elements of this design have appeared in the literature[17,37,38] without consideration or characterisation of the fundamental shift in reaction mechanism that can be achieved or the gains in cyclisation efficiency that can be realised.

Here, we describe the design of a modular and general "auto-cyclase" platform for production of head-to-tail macrocyclised proteins and peptides. We show that self-cyclisation (auto-cyclisation) under dilute conditions is scalable by demonstrating that the reaction proceed via a unimolecular mechanism following first order reaction kinetics, while also suppressing unwanted polymeric side products formed via a bimolecular pathway. Conversely, at elevated concentrations we identify the formation of higher order fusion proteins (tandem autocyclases), that follow an alternative reaction path to yield cyclised oligomeric products. The general utility of the autocyclase platform is demonstrated by cyclisation of two popular but very different classes of macro-molecules (1) disulfide-rich cyclic peptides and (2) α-helical membrane scaffold proteins (MSPs) for producing circularized nanodiscs (cNDs). We expect that the versatility of the auto-cyclases will make these of value in the design of novel macro-cyclic peptides and proteins as well as improving access to existing cyclic molecules.

## Results

**The autocyclase design.** An autocyclase comprise six modules (Fig. 1a): (i) a reactive sequence that is liberated by application of a suitable protease (*cleavage site 1*); (ii) a target protein or peptide to be cyclised; (iii) a cyclisation recognition site (*LPGTG*); (iv) a linker of suitable length and flexibility to promote ligation (*linker*); (v) the cyclising enzyme (*sortase*); and finally (vi) a purification tag (*H10*). An additional module (iv'), with a second (orthogonal) protease cleavage site (*cleavage site 2*), allows the ligase to be isolated as a secondary product (lacking a reactive N-terminal nucleophile—only required if recovery of the enzyme is sought). All autocyclase sequences can be found in the supplementary section (see Supplementary Information and Supplementary Data 1). The two key elements that most significantly affect the efficiency of the autocyclase system are the linker and the ligase, these will be discussed in further detail below.

*The linker.* In our design, a linker connects the SrtA recognition site (LPXTG) to the N-terminal end of SrtA. A suitable linker should allow this recognition site to access the SrtA catalytic site[39], across the α1/β2 and β3/β4 loops of SrtA (Fig. 1b). The NMR structure (PDBID: 2KID) of SrtA (hereon referring to the Δ59 truncation variant of SrtA—residue Q60–K206) reveals that the first secondary structure element (β-strand) starts at residue G74 and that the first 5–9 residues are highly disordered in solution (from $^{15}$N spin relaxation experiments)[39]. Based on this we estimate that a ~35–45 Å, or an approximately 10-amino acid long linker is required (Fig. 1b) but also note that some of the disordered N-terminal residues of SrtA may form part of the linker—possibly allowing shorter linkers to be used. A total of five linkers were designed spanning 7–19 residues: 1. [$L_7$ = AAALEGT], 2. [$L_{12}$ = AAALEGTLVPRS], 3. [$L_{7D}$ = GS(GGS) GG], 4. [$L_{14D}$ = GS(GGS)$_4$] and 5. [$L_{19D}$ = GS(GGS)$_4$LVPRS][40], where the subscript number refers to the number of amino acids and a subscript "D" indicates that the amino acid sequence is designed to yield a dynamic linker.

The production of monomeric cNDs with a 9 nm diameter (cNW9—derived from MSP1D1ΔH4H5 or MSP9) has proven to be particularly challenging[9,35], we therefore chose this as a test system to compare the effect of the different linkers on the efficiency of the autocyclase system. Our experiments revealed
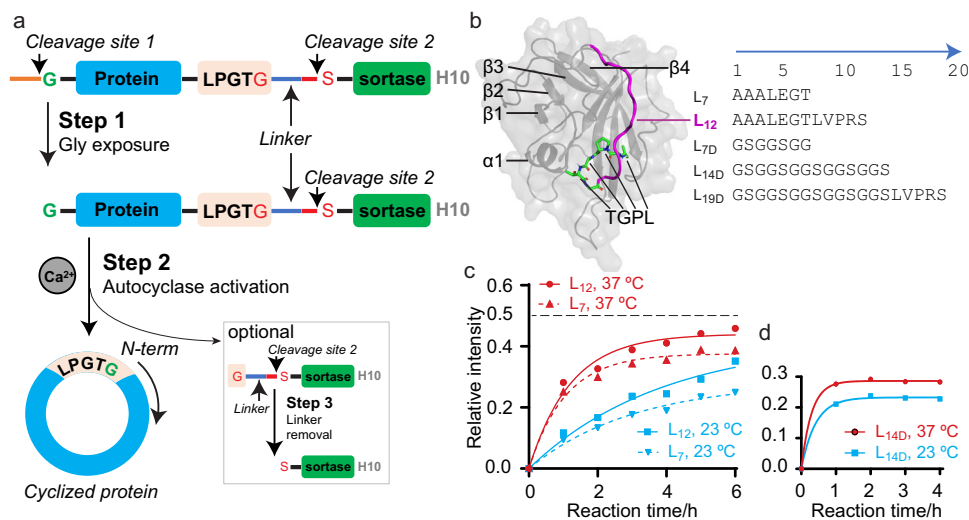
**Fig. 1 Design of autocyclases as a modular, self-cyclising protein family. a** An autocyclase construct features an N-terminal capping sequence (orange line), a protease cleavage site 1 (TEV site—ENLYFQ/G), protein of interest to be cyclised (blue rectangle), SrtA recognition site (pink rectangle), linker (blue line), protease cleavage site 2 (red line, thrombin site—LVPR/S), SrtA (green rectangle) and a C-terminal purification tag (His$_{10}$—H10—in grey). The cyclisation process involves TEV cleavage at cleavage site 1 to expose the N-terminal glycine (step 1), activation of cyclisation reaction using Ca$^{2+}$ ions (step 2) and optionally thrombin cleavage at cleavage site 2 to remove the linker and recover nucleophile free SrtA (step 3). **b** Five linkers (L$_7$, L$_{12}$, L$_{7D}$, L$_{14D}$ and L$_{19D}$) were designed to connect the protein of interest to the SrtA. MD simulations illustrated that the L$_7$ and L$_{12}$ linkers (L$_{12}$ shown in magenta) allows the SrtA recognition site (LPGT residues highlighted in green sticks) to be appropriately positioned at the active site of SrtA. **c, d** Quantitation of SDS-PAGE images (Supplementary Fig. S1) showing the buildup of cMSP9 from autocyclases (aMSP9) over 0–6 h. **c** The cyclisation rate of cMSP9 was faster for the aMSP9-L$_{12}$ than aMSP9-L$_7$ regardless of temperature (23 °C in cyan and 37 °C in red). The theoretical and maximum relative intensity of the cyclic product is ~0.5 (excluding SrtA), indicated by a dash line. The reactions were faster at 37 °C for both constructs. **d** A remarkable increase of reaction rate is observed for the autocyclase containing a dynamic L$_{14D}$ linker.

that autocyclase-MSP9-L$_{12}$ (aMSP9-L$_{12}$) produces an increase in quantity of cMSP9 compared to aMSP9-L$_7$ (Fig. 1c and Supplementary Fig. 1). The higher yield of cMSP9 from aMSP9-L$_{12}$ compared to aMSP9-L$_7$ indicates improved yields as a function of linker length. Next, we investigated if introducing residues that would promote disorder in the linker would affect the reaction rate or yields. We designed a linker of similar length to aMSP9-L$_{12}$, containing a number of GGS repeats (aMSP9-L$_{14D}$) and found a remarkable increase in reaction rate, with the reaction largely complete after ~1 h (cf > 6 h for the L$_7$ linkers, Fig. 1d). We also investigated the effect of temperature on the reaction, and found faster reaction rates of cMSP9 formation at higher temperatures (Fig. 1c, d and Supplementary Fig. 1).

While the longer and more dynamic linkers (L$_{14D}$ and L$_{19D}$) result in a clear increase in reaction rate, this design is also more prone to in vivo hydrolysis leading to a decrease in the overall yields of the aMSP9 protein (Supplementary Fig. 2). We also introduced a shorter dynamic linker to generate aMSP9-L$_{7D}$. Again, we find significant losses due to in vivo hydrolysis, suggesting that linker dynamics alone is sufficient to drive this process (Supplementary Fig. 2c). We then introduced an inhibitory peptide (LPRDA)[41]— that binds to the SrtA active site and inhibits the enzyme—at the N-terminal end of the autocyclase (a-i-MSP9-L$_{14D}$), but found no improvements (Supplementary Fig. 2d), either due to the low affinity of the peptide inhibitor towards SrtA compared to the recognition sequence or due the hydrolysis being driven by other endogenous E. coli proteases (e.g., M23 family of bacterial proteases[42]). Thus, our results indicate that the L$_{12}$ linker provides a good compromise between overall reaction rate and overall yield in recombinant production of cyclic MSPs.

Next, we performed molecular dynamics (MD) simulations on the SrtA recognition sequence (LPGTG) linked to SrtA (Q60-K206) directly (L$_0$, no linker), but found the complex to be unstable with the recognition sequence leaving the active site within 100 ns (Supplementary Figs. 3 and 4). As expected, MD simulations performed with the L$_{12}$ and L$_7$ linkers were stable (Supplementary Fig. 3); however, significantly different conformations of the linker and the N-terminal SrtA region (Q60-G74) were observed (Supplementary Fig. 4). All MD simulations were initiated from the SrtA-substrate structure determined by NMR (PDB ID: 2KID)[39] which was well reproduced in all three cases (Supplementary Fig. 4 - the final configurations after 250 ns of MD are included as Supplementary Data 2). Note that simulations of the SrtA-substrate complex performed using the X-ray structure reported by Zong et al.[43] (PDB ID: 1T2W) were unstable, resulting in the substrate leaving the purported active site in all cases.

*The ligase.* The above experiments were all performed using the wild type SrtA sequence (Δ59). We also generated an aMSP9-L$_7$-eSrtA construct using the evolved sortase pentamutant (eSrtA)[8,44] but found it to produce poor overall yields. The low yields were found to be due to (1) enhanced in vivo hydrolysis by the highly reactive enzyme (Supplementary Fig. 5a–c) and (2) decreased solubility of the protein when expressed at 37 °C (Supplementary Fig. 5d–f).

**The autocyclase reaction mechanism.** The kinetics and mechanism of the SrtA transpeptidation reaction have been characterised in detail previously under suitable steady-state conditions[45], while intramolecular enzymatic cyclisation has not previously been studied. In the traditional (bimolecular) enzymatic cyclisation reaction by SrtA, the initial enzyme concentration is typically similar to that of the substrate[8,9,35], and will be expected to display approximately second order reaction kinetics. In the autocyclase (unimolecular) cyclisation reaction
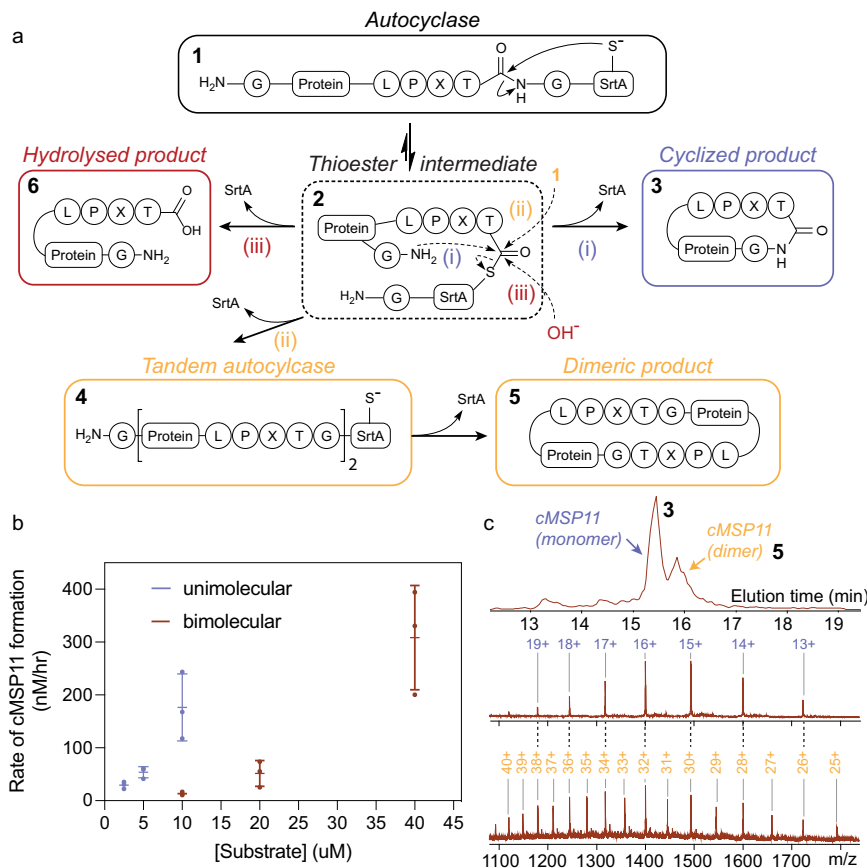
**Fig. 2 The mechanism of autocyclisation. a** A scheme summarising the main reaction pathways of autocyclases. After the generation of a thioester intermediate (2) from a spontaneous intramolecular rearrangement, the reaction can proceed via the following paths: (i) The unimolecular reaction—intramolecular nucleophilic attack by the N-terminal glycine amino group. This pathway results in the cyclised product (3) and free SrtA. (ii) Bimolecular reaction—intermolecular nucleophilic attack. This reaction results in a free SrtA and a protein-protein-SrtA adduct (4—tandem autocyclase). This adduct can self-cyclise to form a cyclic dimer (5). (iii) Hydrolysis—the thioester is resolved by a hydroxide ion and, irreversibly, produces a linear hydrolyzed product (5). **b** The reaction rates of aMSP11-L$_{12}$ cyclisation are compared to the equivalent reaction using the traditional (bimolecular) enzymatic cyclisation method (using MSP11 and WT SrtA in both cases). Rates were determined by monitoring the build-up of cMSP11 using LC/MS. The rates demonstrate that the autocyclase reaction at lower starting concentrations approximates first order reaction kinetics, while the bimolecular reaction approximates second order reaction kinetics. Error bars indicate standard deviations of the mean from three independent replicated experiments (individual data points shown as dots). **c** The MS spectrum (extracted ion chromatogram) of an aMSP11-L$_{12}$ cyclisation reaction at a high starting autocyclase concentration (200 μM) shows the formation of cyclised monomeric and dimeric products. Identity of the two species were confirmed by the observed m/z ion masses.

(Fig. 2a), the initial velocity will be expected to display a mixture of first and second order reaction kinetics (intra- and inter-molecular recognition). Indeed, in the above aMSP9 reactions (using variable linkers) we observe clear evidence of polymerisation at the relatively high concentrations used (~0.1 mM, see also Supplementary Fig. 1), suggesting that the concentration is sufficiently high to allow for intermolecular reactions. However, at lower concentrations, the reaction should favour first order kinetics. It is, therefore, possible to determine a concentration range where the reaction follows a unimolecular mechanism, by following the initial reaction velocity at different (dilute) starting concentration of the autocyclase.

We next set out to characterise the autocyclase reaction mechanism. We used aMSP11-L$_{12}$, which is less prone to self-association, to simplify the analysis. We performed a series of reactions at different concentrations at room temperature, and measured the initial rate of cMSP11 formation, by following the reaction progress using liquid chromatography-mass spectrometry (LC/MS—see also Supplementary Information and Supplementary Data 3). The experiment was also repeated under identical conditions using the same MSP11 and SrtA sequences in

a traditional enzymatic (bimolecular) ligation reaction. The rates of the autocyclase reaction under dilute concentrations were significantly faster than those of the bimolecular reaction, and while the unimolecular autocyclase reaction was feasible in the range of 2–10 μM, the bimolecular rates were only measurable at or above 10 μM—consistent with the predicted change in reaction mechanism. In the range of 2.5 μm–5 μM (Fig. 2b) we observe a linear change in initial reaction velocity, consistent with first order reaction kinetics (a rate constant of ~0.01 h$^{-1}$). Comparing the reaction velocity between 5 and 10 μM, we find a mixture of linear and quadratic behavior suggesting interference from the bimolecular reaction. In the bimolecular reaction we find that increasing the concentration from 10 μM to 20 μM leads to a quadratic increase in reaction velocity (with a rate constant ~130 M$^{-1}$ h$^{-1}$), consistent with second order kinetics[45]. Under dilute reaction conditions (~5 μM), the autocyclase reaction produces the monomeric cyclic product with close to quantitative yields (~90% of theoretical, Supplementary Tables 1, 2).

The MS analysis also revealed the slow build-up of linear MSP over time. This is the product of the third possible reaction pathway in which the reactive SrtA thioester intermediate is being
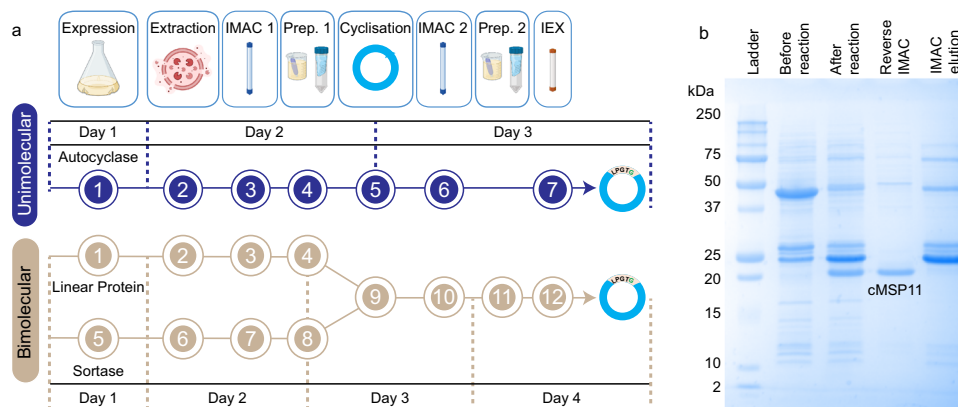
**Fig. 3 Simplification of enzymatic head-to-tail cyclisation using autocyclases. a** A flowchart outlining each distinct step in the process, comparing autocyclisation with the previously described bimolecular reaction[8]. Preparation step 1 (Prep. 1) differs between the two methods. In the unimolecular reaction, this step involves dialysis of the fusion protein into a suitable solution for TEV cleavage and cyclisation reactions, which are performed as a single step (5). In the bimolecular reaction, this step (4) includes the TEV cleavage reaction as well as a number of buffer exchange steps (preceding and following this step). Following reverse IMAC purification (IMAC 2), the sample can be loaded directly onto an ion exchange (IEX) column without the previously described buffer exchange step (Prep. 2). **b** cMSP11 cyclisation using the proposed method, showing SDS-PAGE lanes corresponding to start and end of steps 5 and 6.

hydrolysed irreversibly to produce a linear product (see Supplementary Fig. 5a)—a common feature of both the autocyclase reaction and the traditional bimolecular SrtA reaction. Finally, a small fraction of unreacted fusion protein often remains upon completion of the reaction. We attribute this to fusion proteins that contain misfolded or otherwise inactivated SrtA species.

**Formation of cyclic oligomers**. The formation of polymers in the traditional biomolecular reaction involves the formation of a thioester intermediate, that subsequently reacts with the N-terminal glycine of a neighbouring and unreacted (linear) substrate. In autocyclases, unreacted (linear) substrates are fused to a SrtA molecule, thus the equivalent reaction leads to the formation of products containing multiple, fused, substrate molecules terminated by a single SrtA molecule, which we have termed "tandem" autocyclases.

Indeed, we observe that cyclised SFTI oligomers (i.e., cyclic dimers and trimers of the target protein) at higher starting concentrations, suggesting the formation of tandem autocyclases (containing two or more SFTI and one SrtA molecule). The relative proportion of cyclised polymeric product increases as the starting autocyclase concentration is increased from 50 μM to 200 μM (Supplementary Fig. 6), consistent with the higher order reaction being preferred. As a comparison, no cyclised polymeric SFTI was found in a previous study using the traditional bimolecular reaction when 150 μM linear SFTI was reacted with 50 μM SrtA[46]. Similarly, in the case of MSPs, we find in the MS data that at higher autocyclase starting concentrations (> 50 μM) there is clear evidence of tandem autocyclases (Fig. 2c and Supplementary Fig. 7), which subsequently are cyclised to form oligomeric MSPs. Depending on the intended application of the cyclised protein, this presents an opportunity to adjust the reaction concentrations for generation of higher order cyclic products.

While at high autocyclase reactions, there is a clear bimolecular path to generating polymeric products, there is also great interest in suppressing polymerisation at these elevated concentrations to facilitate scale-up in the production of monomeric products. The concentration at which the bimolecular reaction is preferred to the unimolecular reaction (hence also the formation of polymers) will for autocyclases be chiefly dependent on the $K_m$ of the SrtA

enzyme used and the frequency of intramolecular collisions between the enzyme active site and the recognition sequence (linker sequence). Assuming that the linker is sufficiently long and dynamic to allow for frequent intramolecular collisions, polymerisation in autocyclases may be controlled by increasing the $K_m$ of the SrtA used (reduced affinity for the LPXTG recognition site). To test this hypothesis, we generated an autocyclase mutant containing a A276G mutation (A118G in SrtA sequence), which in SrtA has been shown to increase the $K_m$ 5.3 times[39]. While the autocyclase A276G mutant suffers from low solubility, we find that at concentrations up to 200 μM, no cyclic oligomeric cMSP11 is produced. In contrast, we observe polymeric products at lower concentrations (80 μM) using either the equivalent bimolecular cyclisation reaction or the corresponding autocyclase with the with WT SrtA sequence (Supplementary Fig. 8a–c). This provides a clear strategy for increasing the critical concentration of polymer formation in autocyclases.

**Autocyclases as a general cyclisation strategy**. We next investigated the generality of the method, by generating autocyclase constructs containing different sizes of MSPs and several cyclic peptides of pharmaceutical interest. First, utilising the $L_{12}$ linker, we generated cMSP6, 7 and 15 (in addition to cMSP9 and 11 discussed above). We note that the cMSP15 (cMSP2N2) sequence is roughly twice the length of cMSP11 (1.92x), yet the diameter of MSP2N2 has been reported to be ~15–17 nm in previous reports[47,48]. Thus, while our size measurements, based on low resolution negative stain images of cMSP2N2 (Supplementary Fig. 10), are closer to the theoretically expected diameter of 22 nm, we have followed existing convention, and refer to circular MSP2N2 as cMSP15 and cNW15. Our procedure for cMSP production is very similar to that previously reported[8,35], except that we do not need to separately acquire or produce the SrtA enzyme (Fig. 3 and Supplementary Figs. 9, 10). We also find that the TEV protease cleavage (liberating the N-terminal nucleophile) and cyclisation reactions can be performed concurrently (Supplementary Fig. 11), simplifying the process further. We have quantitated our yields at critical steps of the process (Supplementary Tables 1, 2), and find that our overall yields are generally higher than those previously reported for the bimolecular reaction[8,9,35]. We also find that the solubility of our fusion is higher than previous reports of MSPs expressed in *E. coli*, which
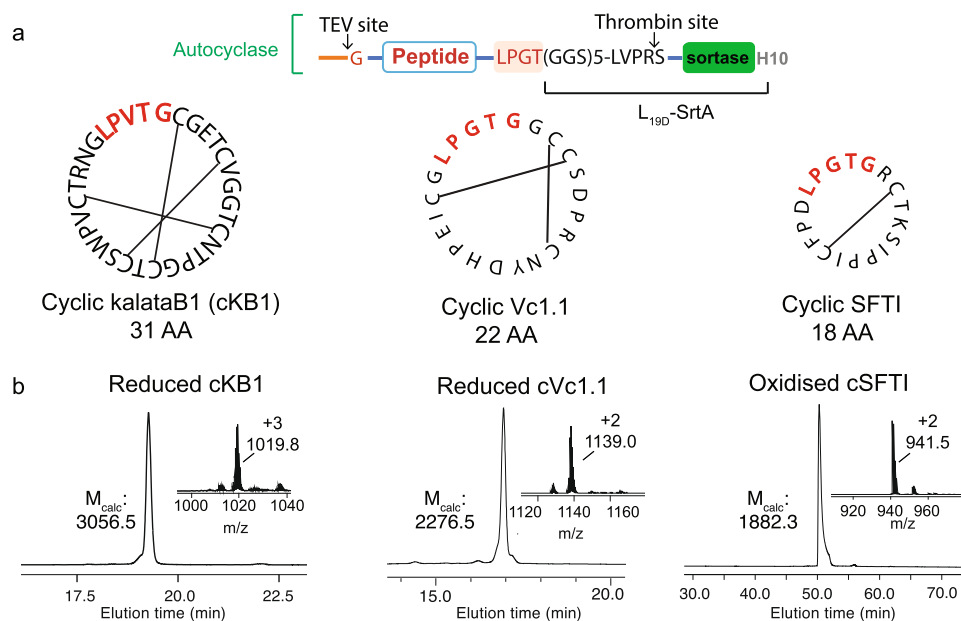
**Fig. 4 The production of three different cyclic peptides from autocyclases. a** The design of the autocyclases used for producing cyclic SFTI, kalataB1 (kB1) and Vc1.1. The native peptide sequences are shown in black, with lines indicating the disulfide bond connectivity. The SrtA recognition sequence (LPXTG) that remains in in the peptide after the cyclisation is highlighted in red. **b** The identity and purity of the cyclised peptides were assessed by rpHPLC and LC/MS. Observed masses were m/z 1019.8 [M + 3H]³⁺ for cKB1, m/z 1139.0 [M + 2H]²⁺ for cVc1.1 and m/z 941.5 [M + 2H]²⁺ for cSFTI. Calculated masses are m/z 1019.8 [M + 3H]³⁺ for cKB1, m/z 1139.3 [M + 2H]²⁺ for cVc1.1 and m/z 942.2 [M + 2H]²⁺ for cSFTI.

may be due to the favourable solution properties of the fused bacterial SrtA protein[8,9,35]. We next used the produced cMSP7, 9, 11, and 15 to assemble cNDs (cNW7, 9, 11, 15) containing 1-Palmitoyl-2-Oleoyl-sn-Glycero-3-Phosphocholine (POPC) lipid bilayers (Supplementary Figs. 9, 10).

We also investigated if the autocyclase method could be employed to produce head-to-tail macrocyclic peptides. We selected three well-characterised disulfide-rich cyclic peptides that vary in their cysteine content, SFTI (a potent protease inhibitor with one disulfide bond)[49], Vc1.1 (a potent inhibitor of a nicotinic acetylcholine receptors with two disulfide bonds)[50] and KalataB1 (kB1; a uterotonic plant peptide with three disulfide bonds)[51]. Initial screening of different linkers in these constructs revealed that these autocyclases were less prone to in vivo hydrolysis compared to the MSPs (Supplementary Fig. 2). The longer dynamic linker (L₁₉D) could therefore be used to facilitate efficient cyclisation of the cyclic peptides (Fig. 4).

**Applications of nanodiscs in structural biology and in vivo imaging.** The applications of cyclic peptides and nanodiscs are diverse[52]. A number of cyclic peptides have pharmaceutical or agrochemical potential[7]; while nanodiscs are commonly employed in studies of membrane proteins in a lipide bilayer[8,53] and used in imaging and drug delivery applications[54]. Here we demonstrate the utility of the autocyclase system to provide further insights in both areas.

First, we used the purified cMSP11 to encapsulate the voltage sensor domain (VSD) of a bacterial voltage gated potassium channel (KvAP)[55]. The ¹⁵N-TROSY spectrum at elevated temperatures (323 K) shows resonances associated with the correct folding of the protein within the nanodisc (Supplementary Fig. 12).

Next, we produced an isotopically-labelled (¹³C/¹⁵N) SFTI peptide, which allowed us to measure, to the best of our knowledge, the first triple resonance 3-dimensional (3D) NMR experiments of a macrocyclised disulfide-rich peptide. This

peptide sequence had previously been produced using a semi-synthetic approach where the chemically synthesized (linear) peptide was cyclised by SrtA[46]. The produced peptide was observed to yield multiple structural isoforms. The structural heterogeneity was also observed here. While the source of the heterogeneity previously remained unknown, we were able to employ 3D NMR experiments at very high resolution using non-uniform sampling and non-Fourier spectral reconstruction methods[56] to perform sequential resonance assignment of all structural isoforms of this peptide in solution (Supplementary Fig. 13a). We can clearly see 3 isoforms, and using the ¹³C chemical shifts of the proline residues[57], we are able to conclusively determine that in the major isoform (65% population) all prolines are in the trans configuration, while the two minor isoforms (20% and 15% populations) are due to a cis-proline configuration at positions P13 and P16 respectively (Supplementary Fig. 13b)—with the relative energy of formation of each cis-isoform from the predominant trans-isoform calculated to be: $\Delta G_{P13} = 12$ and $\Delta G_{P16} = 15$ kcal·mol⁻¹. Furthermore, we are able to unequivocally demonstrate the presence of a covalent bond between T18 and G1 in all three isoforms (Supplementary Fig. 13c).

Another important application of nanodiscs, is their use as carriers of drugs or imaging agents[54]— with implications for peptide drug development[7,51,58,59]. Despite their potential utility in this area, little is known about how the size or cyclisation of nanodiscs affect their in vivo fate (in absence of a cargo)[54,60]. To probe this, we generated both linear and cyclised nanodiscs (NW11 and cNW11) carrying two optical probes that fluoresce at different wavelengths—one inserted into the lipid bilayer, and the other conjugated to the MSP (Fig. 5). We also generated small (cNW7) and large (cNW15) diameter dual-labelled cNDs. The two sets of nanodiscs allow us to monitor the effect of cyclisation and size on biodistribution and metabolism in vivo. Indeed, we find clear differences in both biodistribution and metabolism across the different samples over the 24 h post intravenous injection (Fig. 5, Supplementary Fig. 14, Supplementary Data 4).
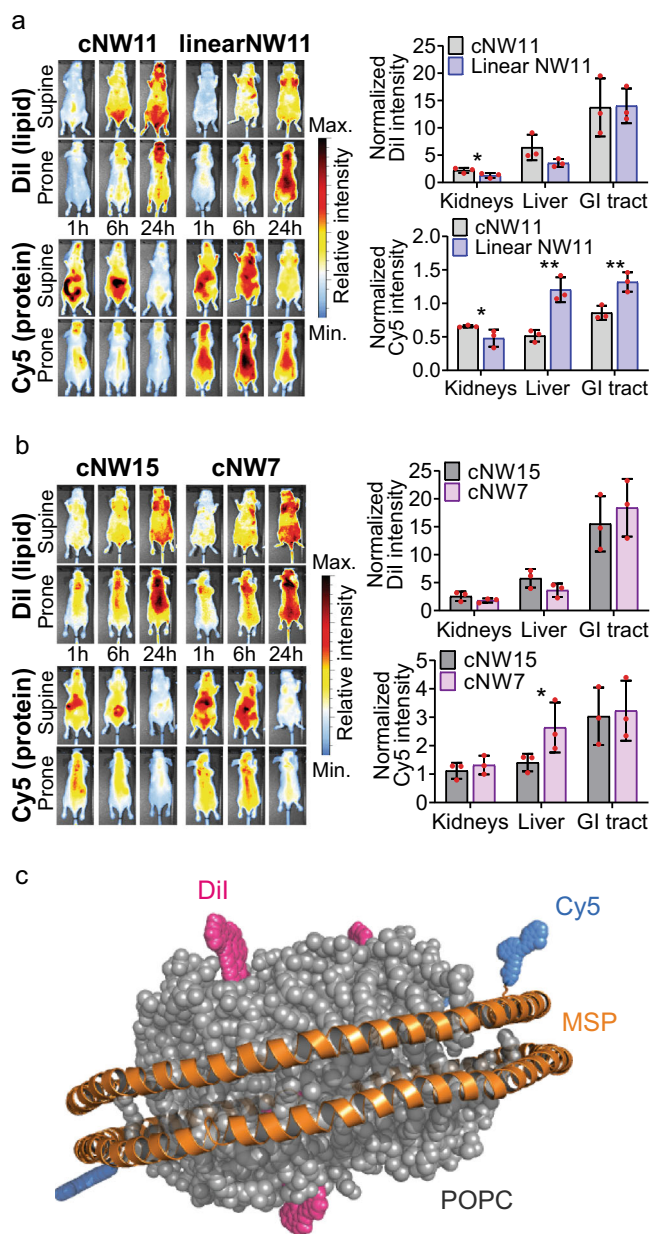
**Fig. 5 In vivo biodistribution of nanodiscs in naïve BALB/c nude mice.** Animals were administered fluorescent nanodiscs via the tail vein and imaged 1, 6, and 24 h post injection. Imaging of the Cy5 conjugated MSP protein was performed using 620 nm excitation and 670 nm emission filters and the lipophilic DiI dye (loaded into the bilayer) using 540 nm excitation and 620 nm emission filters. **a** Comparison of cNW11 and linear NW11 distributions. DiI intensity, localised to regions of fat deposits, increases with time reaching a peak at 24 h post injection for both linear and cyclised nanodiscs. Cy5 distribution demonstrates clearance of both linear and cyclic nanodiscs through the GI tract at early time points. Ex vivo analysis of clearance organs demonstrates higher linear NW11 protein in the liver 24 h post injection compared to cNW11 but lower DiI suggesting increased metabolism. Ex vivo organ fluorescence data is presented as mean normalised radiant efficiency values ($n = 3$) with standard deviation error bars. **b** Comparison of cNW15 and cNW7 distributions. Both cNW15 and cNW7 demonstrate similar distributions to cNW11. Ex vivo analysis demonstrates higher cNW7 protein in the liver compared to cNW15, similar to the trend observed between linear NW11 and cNW11. Statical analysis was performed using unpaired two-tailed t-tests, *p-value $\leq 0.05$, **$p \leq 0.005$. Cy5 linearNW11 vs cNW11 kidneys $p = 0.04$, liver $p = 0.002$, GI $p = 0.005$; DiI linearNW11 vs cNW11 kidneys $= 0.04$; Cy5 cNW7 vs cNW15 liver $p = 0.05$. **c** Graphical representation of a fluorescent nanodisc demonstrating the Cy5 (blue) conjugated to the MSP (orange) lysine sidechain, and the lipophilic DiI dye (magenta) located within the POPC lipid bilayer (grey).

## Discussion

Here, we describe the engineering of a class of proteins, termed autocyclases, which can undergo either intramolecular head-to-tail macrocyclisation to release an enzyme and a cyclic product or intermolecular head-to-tail conjugation, followed by a head-to-tail macrocyclisation to release an enzyme and a cyclic oligomer. We show that conversion of the autocyclase to a monomeric cyclic product, increases as a function of decreasing starting concentration, while approximating first-order reaction kinetics, reaching nearly quantitative yield below concentrations of ~5 μM. We find that under identical conditions the equivalent bimolecular reaction is unfeasible. We also show that the critical concentration at which polymers are formed is related to the $K_m$ of the SrtA enzyme used, providing a path for generation of polymerisation resistant autocyclases.

In recent years, several methods have emerged that enable in vivo production of self-cyclised MSPs[38,61]. This cyclization is achieved by flanking the MSP sequence with two reactive elements that upon expression can spontaneously yield a cyclic product. The first of these, uses the intein-mediated ligation method to achieve head-to-tail cyclisation[61], while the other uses the SpyCatcher/SpyTag technology to produce a head-to-sidechain cyclised product[38]. While the former leaves a modest ligation scar, the latter yields a product carrying a ~120 amino acid insertion. Both methods typically also require the inclusion of an affinity tag within the cyclised product for downstream purification. In vivo cyclisation makes these methods attractive as they dramatically improve throughput. In vivo cyclisation, however, also prohibits detailed characterisation and control of the reaction. The main difference of these methods to autocyclases presented here is that the cyclisation reaction is here triggered in vitro, upon cleavage of the protecting N-terminal capping sequence and addition of calcium. As demonstrated, we can use this property to characterise the reaction in detail, providing opportunities to tune the reaction conditions or explore how small changes to the protein sequence (enzyme or linker) can affect the outcome of the cyclisation reaction—allowing for further optimisation of the self-cyclisation reaction.

The lipid content of the discs appears to preferentially accumulate in brown adipose tissue (see fat pads near the animal head), while the MSPs are generally found in the regions associated with clearance (liver, kidney and GI tract) (Fig. 5). Comparing the linear and cyclic proteins (Fig. 5a), we find higher levels of the linear MSP in the liver (p-value 0.002) and GI tract (p-value 0.005) compared to cyclic MSP (normalised to blood), supporting increased in vivo stability following cyclisation. Similarly, the cND appears to deposit more lipids in the liver compared to the linear nanodiscs, consistent with a longer circulation time. Comparison of the in vivo images of smaller and larger cNDs (cNW7 and cNW15; Fig. 5b), showed the smaller discs to be cleared more rapidly than the larger discs, with similar patterns in clearance organ accumulation at 24 h post administration. The only statically significant difference observed between small and large cNDs was an increase in the smaller cMSP protein in the liver (p-value 0.05), associated with increased clearance. These data provide the first insights into the in vivo properties of cNDs and suggest that larger cNDs are more resistant to clearance than smaller and linear nanodiscs.

In the autocyclase reaction, the key departure from traditional (bimolecular) enzymatic cyclisation is the first step of the reaction (Fig. 2a), i.e., the recognition of the substrate sequence by the enzyme. In the bimolecular SrtA reaction, this is described by the $K_m$ of the enzyme, i.e. directly related to the substrate concentration. In the unimolecular reaction (using the same enzyme and recognition motif), this step is concentration independent and relates instead to the properties of the linker (frequency and orientation of collisions of recognition sequence with the catalytic site). Once the intermediate is formed, all subsequent steps are largely the same—further emphasising the significance of the linker.

Interestingly, several class A and class C sortase enzymes contain a flexible, N-terminal segment that may regulate substrate binding, as their structures reveal that this N-terminal appendage partially shields the active site and wraps around the surface of the catalytic site[62]. It is likely that our design is hijacking this natural autoregulatory function, by replacing the autoinhibitory sequence, that has evolved to access the catalytic site, with a recognition site for autocyclisation. This is supported by our MD simulations, which further show that direct fusion of the target protein to SrtA lacking an N-terminal spacer, leads to an unstable complex (Supplementary Figs. 3, 4). Interestingly, we note that in a reported example of human growth hormone (hGH) cyclisation[37], a direct fusion of hGH to SrtA was created to enable "one-step" purification and cyclisation of hGH. The possibility of intramolecular cyclisation was not considered, and given the results of our MD simulations we can conclude that although this direct fusion reduced the number steps of SrtA mediated cyclisation (see Fig. 3), the cyclisation was achieved instead via the traditional intermolecular reaction.

Another consequence of the departure from traditional (bimolecular) enzymatic cyclisation is the mechanism of polymerisation. Our results show that cyclised oligomeric substrates are formed via a unique reaction path. The first step is the formation of a tandem autocyclase, which contains multiple fused substrates with a single C-terminal SrtA sequence. It is unclear whether this alternative mechanism offers any advantages over oligomer formation in the traditional bimolecular reaction, but comparison of our results with those previously reported for SFTI, suggests that in certain conditions the tandem autocyclase may form oligomeric cyclised products preferentially compared to the traditional bimolecular reaction[46]. Further characterisation of the tandem autocyclases is necessary to fully elucidate the properties of this reaction path.

While the tandem autocyclases provide a stable intermediate from which polymeric materials can be generated, there is also great interest in avoiding the formation of this intermediate. Dilution of the substrate has traditionally been the mechanism by which this has been addressed, however, the shift in reaction mechanism in the autocyclases provides an alternative avenue to tackle this problem. We show that tethering of the enzyme and substrate makes the intramolecular recognition concentration independent (Fig. 2). Tethering is also well known to enhance the local concentration of interacting partners to yield complexes of even weakly binding proteins – used frequently in structural studies[63]. This suggests that increasing the $K_m$ of the SrtA enzyme used in an autocyclase will increase the concentration at which the tandem autocyclase is formed (via an intermolecular reaction), while the tethering will ensure that the intramolecular cyclisation remains feasible. We confirm this by showing that monomeric cMSP proteins can be produced from aMSP11 at elevated concentrations (200 μm) when a mutant of SrtA is used with a $K_m$ that is ~6-fold[39] higher than that of the wild type protein. The mutant generated, unfortunately, suffers from low solubility, possibly due to a loss in stability, and future work in

identifying mutants with high inherent stability and solubility and with a high $K_m$ will likely further improve the system. This of course can also be achieved by altering the recognition sequence. Combining such mutations with an optimised linker would be particularly interesting, as any losses in reaction rate due to the weaker binding of the recognition sequence to the recognition site, would be compensated by the increased effective local concentration at this site due to favourable linker dynamics. Such optimisation can now be pursued following the establishment of the described theoretical framework. Autocyclases, thus, offer a flexible system that through continued development, promises to lead to further advances in enzymatic head-to-tail macrocyclization.

Finally, we demonstrate the utility of macrocyclised proteins produced using the autocyclase approach by (1) uniformly $^{13}C/^{15}N$ isotope-labelling a cyclised disulfide-rich peptide for the first time, (2) generating a range of cNDs of different sizes, including one containing an ion channel voltage sensor domain, and (3) providing the first data on the in vivo biodistribution of cNDs and clearance pathway as a function of cND size and cyclisation. Isotope labelling demonstrates the wealth of structural NMR data that can now be accessed for cyclic peptides produced recombinantly, while our imaging experiments provide evidence that large cNDs may serve as useful carriers of drugs and imaging agents. The detailed characterization of cSFTI also highlights that while the sortase ligation scar (LPXTG) is a minor component of proteins such as MSPs, it may constitute a significant insertion in small cyclic peptides, with possible effects on peptide structure and dynamics (as seen in Supplementary Fig. 13). The position of the insert should therefore be chosen carefully to maintain peptide function. Indeed, many cyclic peptides have been reported that remain active with an LPXTG insertion[64,65].

Clearly the applications of macrocyclised peptides and proteins are both numerous and diverse. It is notable that many of them, including our demonstrated examples in structural biology and in vivo studies, demand large quantities of pure materials. While the cyclic materials can in theory be produced using other cyclisation methods, autocyclases represent a simple, fast and low-cost method of producing these in a scalable manner. We expect our method to provide better access to this exciting class of biomolecules and facilitate future research across fields of structural biology and pharmaceutical sciences.

## Conclusion
Cyclic peptides and proteins have many properties that make them attractive as biochemical tools and for pharmaceutical development. Numerous enzymatic methods have emerged that are capable of cyclising linear substrates. Here we present a method that incorporates both the enzyme and the substrate in the same molecular entity. We show that this leads to a change in reaction mechanism, resulting in a self-cyclisation reaction following first-order reaction kinetics under dilute concentrations. The mechanism of intermolecular oligomerisation at high starting concentrations is also characterised, revealing an alternative path via a tandem autocyclase that can yield cyclised oligomeric products. Finally, we show how the intermolecular reaction leading to polymeric materials can be suppressed by tuning the properties of the enzyme, allowing for monomeric products to be formed even at elevated concentrations. The utility of the method is demonstrated by production and characterisation of a range of cyclic peptides and proteins.

## Methods
**Cloning**. All sequences generated in this study and the primers used in their production are provided in the Supplemental sections (Supplementary Tables 3, 4 and Supplemental Data 1) and available from the authors upon reasonable

request. All autocyclase constructs described here feature an N-terminal capping sequence, a TEV cleavage site, the MSP or peptide of interest, a linker of various lengths and sequences, SrtA (evolved or WT) and a C-terminal His$_{10}$ tag. As a template, a codon-optimized gene encoding an N-terminal His$_6$, TEV site, MSP9, L$_7$ linker, eSrtA and a C-terminal His$_6$ was purchased from IDT. The gene was double digested with *Nde*I and *Xho*I, cleaned up using a macherey-nagel nucleospin gel and PCR clean-up kit and cloned into a pET29a vector. Genetic modifications on this template plasmid were achieved by either gene mutagenesis using NEB Q5 mutagenesis kit or gene replacement using restriction enzyme digestion, to obtain other constructs that encode various MSP, linkers, sortase or histidine tags. The amino acid sequence of the MSPs in these constructs are based on previous reports (including MSP9 and MSP11[8], MSP6 and MSP7[66], and MSP15[67]). The macrocyclic peptide autocyclase constructs were produced by replacing the MSP9 gene in the *a*MSP9 autocyclase template. The peptide sequences were selected based on past literature reports[49–51]. Detailed methods describing specific cloning protocols for each construct are provided in the Supplementary Information.

**Autocyclase expression and purification.** Each autocyclase expression construct (in a pET29a vector) was transformed into *E. coli* BL21(DE3) cells. Freshly transformed colonies or glycerol stock were used to inoculate LB media containing 50 μg/mL of kanamycin. The starter culture was incubated at 30 °C and agitated (shaking at 220 rpm) overnight. 3 mL of the preculture was used to inoculate 300 mL of LB broth containing 50 μg/mL of kanamycin. The culture was then incubated at 37 °C and agitated (shaking at 250 rpm) until the OD$_{600}$ reached ~1.0. Expression was induced by addition of 0.2 or 1 mM IPTG and the culture was left shaking at 250 rpm for 1–6 h at 30 °C. The cells were harvested at 6 h by centrifugation at 6,000 *g* for 10 min at 4 °C.

The cell pellets were resuspended in lysis buffer (25 mM sodium phosphate pH 7.4, 500 mM NaCl, 20 mM imidazole) containing 1 mg/mL lysozyme and stirred at 4 °C for 0.5 h. The resuspended cells were lysed by two 5 min cycles of sonication on ice (digital sonifier 450 Branson; 40% power; repetitions of 3 s on-pulse and 12 s off-pulse) with a 5 min break between the cycles to avoid overheating. The sonicated sample was centrifuged at 30,000 *g* for 30 min at 4 °C to remove the insoluble fractions and the supernatant was loaded onto a gravity column containing Ni-NTA resin (pre-equilibrated with 4 °C lysis buffer). The resin was washed with five column volumes (CVs) of lysis buffer and the autocyclase was eluted with five CVs of elution buffer (25 mM sodium phosphate pH 7.4, 500 mM NaCl, 500 mM imidazole).

**MSP and peptide cyclisation.** The cyclisation process follows the steps described in Fig. 3. For quantitation of the cyclisation reaction, the first preparation step (4) includes an anion exchange chromatography step to remove any free sortase enzyme that is co-purified with the fusion protein. While this is important for quantitation and determination of the reaction mechanism, it can be excluded to make the process more efficient without significantly affecting the final yields. Details of each step in Fig. 3 are provided in the Supplementary Information.

Further details on the reagents, cloning, over-expression, purification, cyclisation, SDS-PAGE analysis, enzyme kinetics, molecular dynamics, empty nanodisc and KvAP-VSD nanodisc assembly, NMR experiments and biodistribution analysis are described in the Supplementary Information.

## Data availability

The data that support the findings of this study are available in the supplementary sections or from the corresponding author upon reasonable request. Plasmids for producing circular membrane scaffold proteins and peptides will be available upon request.

## References

1. Cascales, L. & Craik, D. J. Naturally occurring circular proteins: Distribution, biosynthesis and evolution. *Org. Biomol. Chem.* **8**, 5035–5047 (2010).
2. Clark, R. J., Akcan, M., Kaas, Q., Daly, N. L. & Craik, D. J. Cyclization of conotoxins to improve their biopharmaceutical properties. *Toxicon* **59**, 446–455 (2012).
3. Wong, C. T. et al. Orally active peptidic bradykinin B1 receptor antagonists engineered from a cyclotide scaffold for inflammatory pain treatment. *Angew. Chem. Int Ed. Engl.* **51**, 5620–5624 (2012).
4. Driggers, E. M., Hale, S. P., Lee, J. & Terrett, N. K. The exploration of macrocycles for drug discovery–an underexploited structural class. *Nat. Rev. Drug Disco.* **7**, 608–624 (2008).
5. Hosseinzadeh, P. et al. Anchor extension: A structure-guided approach to design cyclic peptides targeting enzyme active sites. *Nat. Commun.* **12**, 3384 (2021).
6. Camarero, J. A. & Muir, T. W. Biosynthesis of a head-to-tail cyclized protein with improved biological activity. *J. Am. Chem. Soc.* **121**, 5597–5598 (1999).
7. Vinogradov, A. A., Yin, Y. & Suga, H. Macrocyclic peptides as drug candidates: Recent progress and remaining challenges. *J. Am. Chem. Soc.* **141**, 4167–4181 (2019).
8. Nasr, M. L. et al. Covalently circularized nanodiscs for studying membrane proteins and viral entry. *Nat. Methods* **14**, 49–52 (2017).
9. Zhang, A. H. et al. Elucidating the lipid binding properties of membrane-active peptides using cyclised nanodiscs. *Front. Chem.* **7**, 238 (2019).
10. Padmanabha Das, K. M., Shih, W. M., Wagner, G. & Nasr, M. L. Large nanodiscs: A potential game changer in structural biology of membrane protein complexes and viral entry. *Front Bioeng. Biotechnol.* **8**, 539 (2020).
11. Zhao, Z. et al. DNA-Corralled nanodiscs for the structural and functional characterization of membrane proteins and viral entry. *J. Am. Chem. Soc.* **140**, 10639–10643 (2018).
12. Dallo, S., Shin, J., Zhang, S., Ren, Q. & Bao, H. Designer nanodiscs to probe and reprogram membrane biology in synapses. *J. Mol. Biol.* **435**, 167757 (2023).
13. Ren, Q., Zhang, S. & Bao, H. Circularized fluorescent nanodiscs for probing protein-lipid interactions. *Commun. Biol.* **5**, 507 (2022).
14. Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. Synthesis of proteins by native chemical ligation. *Science* **266**, 776–779 (1994).
15. Muir, T. W., Sondhi, D. & Cole, P. A. Expressed protein ligation: A general method for protein engineering. *Proc. Natl. Acad. Sci. USA* **95**, 6705–6710 (1998).
16. Kimura, R. & Camarero, J. A. Expressed protein ligation: A new tool for the biosynthesis of cyclic polypeptides. *Protein Pept. Lett.* **12**, 789–794 (2005).
17. Kimura, R. H., Tran, A. T. & Camarero, J. A. Biosynthesis of the cyclotide Kalata B1 by using protein splicing. *Angew. Chem. Int Ed. Engl.* **45**, 973–976 (2006).
18. Muir, T. W. Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem* **72**, 249–289 (2003).
19. Tavassoli, A. & Benkovic, S. J. Split-intein mediated circular ligation used in the synthesis of cyclic peptide libraries in E. coli. *Nat. Protoc.* **2**, 1126–1133 (2007).
20. Kawakami, T. et al. Diverse backbone-cyclized peptides via codon reprogramming. *Nat. Chem. Biol.* **5**, 888–890 (2009).
21. Mazmanian, S. K., Liu, G., Ton-That, H. & Schneewind, O. Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. *Science* **285**, 760–763 (1999).
22. Mao, H., Hart, S. A., Schink, A. & Pollok, B. A. Sortase-mediated protein ligation: A new method for protein engineering. *J. Am. Chem. Soc.* **126**, 2670–2671 (2004).
23. Liebscher, S. et al. N-terminal protein modification by substrate-activated reverse proteolysis. *Angew. Chem. Int. Ed. Engl.* **53**, 3024–3028 (2014).
24. Chang, T. K., Jackson, D. Y., Burnier, J. P. & Wells, J. A. Subtiligase: a tool for semisynthesis of proteins. *Proc. Natl. Acad. Sci. USA.* **91**, 12544–12548 (1994).
25. Toplak, A., Nuijens, T., Quaedflieg, P. J. L. M., Wu, B. & Janssen, D. B. Peptiligase, an enzyme for efficient chemoenzymatic peptide synthesis and cyclization in water. *Adv. Synth. Catal.* **358**, 2140–2147 (2016).
26. Schmidt, M. et al. Omniligase-1: A powerful tool for peptide head-to-tail cyclization. *Adv. Synth. Catal.* **359**, 2050–2055 (2017).
27. Jackson, M. A. et al. Molecular basis for the production of cyclic peptides by plant asparaginyl endopeptidases. *Nat. Commun.* **9**, 2411 (2018).
28. Hemu, X. et al. Structural determinants for peptide-bond formation by asparaginyl ligases. *Proc. Natl Acad. Sci. USA.* **116**, 11737–11746 (2019).
29. Haywood, J. et al. Structural basis of ribosomal peptide macrocyclization in plants. *Elife* **7**, e32955 (2018).
30. Du, J. Q. et al. A bifunctional asparaginyl endopeptidase efficiently catalyzes both cleavage and cyclization of cyclic trypsin inhibitors. *Nat. Commun.* **11**, 1–11 (2020).
31. Nguyen, G. K. et al. Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat. Chem. Biol.* **10**, 732–738 (2014).
32. Harris, K. S. et al. A suite of kinetically superior AEP ligases can cyclise an intrinsically disordered protein. *Sci. Rep.-Uk* **9**, 1–13 (2019).
33. Harris, K. S. et al. Efficient backbone cyclization of linear peptides by a recombinant asparaginyl endopeptidase. *Nat. Commun.* **6**, 10199 (2015).
34. Yang, R. et al. Engineering a catalytically efficient recombinant protein ligase. *J. Am. Chem. Soc.* **139**, 5351–5358 (2017).
35. Yusuf, Y. et al. Optimization of the production of covalently circularized nanodiscs and their characterization in physiological conditions. *Langmuir* **34**, 3525–3532 (2018).

36. Wu, Z., Guo, X. & Guo, Z. Sortase A-catalyzed peptide cyclization for the synthesis of macrocyclic peptides and glycopeptides. *Chem. Commun. (Camb.)* **47**, 9218–9220 (2011).

37. Rasche, N. et al. PROLink-single step circularization and purification procedure for the generation of an improved variant of human growth hormone. *Bioconjug. Chem.* **27**, 1341–1347 (2016).

38. Zhang, S. et al. One-step construction of circularized nanodiscs using SpyCatcher-SpyTag. *Nat. Commun.* **12**, 5451 (2021).

39. Suree, N. et al. The structure of the Staphylococcus aureus sortase-substrate complex reveals how the universally conserved LPXTG sorting signal is recognized. *J. Biol. Chem.* **284**, 24465–24477 (2009).

40. Warden-Rothman, R., Caturegli, I., Popik, V. & Tsourkas, A. Sortase-tag expressed protein ligation: Combining protein purification and site-specific bioconjugation into a single step. *Anal. Chem.* **85**, 11090–11097 (2013).

41. Wang, J. et al. Oligopeptide targeting sortase a as potential anti-infective therapy for Staphylococcus aureus. *Front Microbiol* **9**, 245 (2018).

42. Becker, S., Frankel, M. B., Schneewind, O. & Missiakas, D. Release of protein A from the cell wall of Staphylococcus aureus. *Proc. Natl Acad. Sci. USA.* **111**, 1574–1579 (2014).

43. Zong, Y., Bice, T. W., Ton-That, H., Schneewind, O. & Narayana, S. V. Crystal structures of Staphylococcus aureus sortase A and its substrate complex. *J. Biol. Chem.* **279**, 31383–31389 (2004).

44. Chen, I., Dorr, B. M. & Liu, D. R. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc. Natl. Acad. Sci. USA* **108**, 11399–11404 (2011).

45. Huang, X. et al. Kinetic mechanism of Staphylococcus aureus sortase SrtA. *Biochemistry* **42**, 11307–11315 (2003).

46. Jia, X. et al. Semienzymatic cyclization of disulfide-rich peptides using Sortase A. *J. Biol. Chem.* **289**, 6627–6638 (2014).

47. Gao, Y., Cao, E., Julius, D. & Cheng, Y. TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature* **534**, 347–351 (2016).

48. Grinkova, Y. V., Denisov, I. G. & Sligar, S. G. Engineering extended membrane scaffold proteins for self-assembly of soluble nanoscale lipid bilayers. *Protein Eng. Des. Sel.* **23**, 843–848 (2010).

49. Luckett, S. et al. High-resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds. *J. Mol. Biol.* **290**, 525–533 (1999).

50. Sandall, D. et al. A novel α-conotoxin identified by gene sequencing is active in suppressing the vascular response to selective stimulation of sensory nerves in vivo. *Biochemistry* **42**, 6904–6911 (2003).

51. Saether, O. et al. Elucidation of the primary and three-dimensional structure of the uterotonic polypeptide kalata B1. *Biochemistry* **34**, 4147–4158 (1995).

52. Sligar, S. G. & Denisov, I. G. Nanodiscs: A toolkit for membrane protein science. *Protein Sci.* **30**, 297–315 (2021).

53. Daniilidis, M., Brandl, M. J. & Hagn, F. The advanced properties of circularized MSP nanodiscs facilitate high-resolution NMR studies of membrane proteins. *J. Mol. Biol.* **434**, 167861 (2022).

54. Wong, P. et al. Antibody targeted PET Imaging of (64)Cu-DOTA-Anti-CEA PEGylated lipid nanodiscs in CEA positive tumors. *Bioconjug. Chem.* **31**, 743–753 (2020).

55. Shenkarev, Z. O. et al. Lipid-protein nanodiscs as reference medium in detergent screening for high-resolution NMR studies of integral membrane proteins. *J. Am. Chem. Soc.* **132**, 5628–5629 (2010).

56. Mobli, M. & Hoch, J. C. Nonuniform sampling and non-Fourier signal processing methods in multidimensional NMR. *Prog. Nucl. Magn. Reson Spectrosc.* **83**, 21–41 (2014).

57. Siemion, I. Z., Wieland, T. & Pook, K. H. Influence of the distance of the proline carbonyl from the beta and gamma carbon on the 13C chemical shifts. *Angew. Chem. Int Ed. Engl.* **14**, 702–703 (1975).

58. Goto, Y. & Suga, H. The RaPID platform for the discovery of pseudo-natural macrocyclic peptides. *Acc. Chem. Res.* **54**, 3604–3617 (2021).

59. Norman, A. et al. Discovery of cyclic peptide ligands to the SARS-CoV-2 spike protein using mRNA display. *ACS Cent. Sci.* **7**, 1001–1008 (2021).

60. Huda, P. et al. PET/CT based in vivo evaluation of 64Cu labelled nanodiscs in tumor bearing mice. *PLoS One* **10**, e0129310 (2015).

61. Miehling, J., Goricanec, D. & Hagn, F. A Split-Intein-based method for the efficient production of circularized nanodiscs for structural studies of membrane proteins. *Chembiochem.* **19**, 1927–1933 (2018).

62. Jacobitz, A. W., Kattke, M. D., Wereszczynski, J. & Clubb, R. T. Sortase transpeptidases: Structural biology and catalytic mechanism. *Adv. Protein Chem. Struct. Biol.* **109**, 223–264 (2017).

63. Kobe, B., Ve, T. & Williams, S. J. Fusion-protein-assisted protein crystallization. *Acta. Crystallogr. F. Struct. Biol. Commun.* **71**, 861–869 (2015).

64. Zhang, J., Yamaguchi, S. & Nagamune, T. Sortase A-mediated synthesis of ligand-grafted cyclized peptides for modulating a model protein-protein interaction. *Biotechnol. J.* **10**, 1499–1505 (2015).

65. Bolscher, J. G. et al. Sortase A as a tool for high-yield histatin cyclization. *FASEB J.* **25**, 2650–2658 (2011).

66. Hagn, F., Etzkorn, M., Raschle, T. & Wagner, G. Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins. *J. Am. Chem. Soc.* **135**, 1919–1925 (2013).

67. Denisov, I. G. & Sligar, S. G. Nanodiscs in membrane biochemistry and biophysics. *Chem. Rev.* **117**, 4669–4713 (2017).

## Acknowledgements

## Author contributions

X.J. and M.M. conceived of the project and supervised all aspects. X.J. and A.Z. designed and performed the molecular biology work. Y.K.C. conducted the reaction kinetic study with advice from B.R.H., X.J., and M.M. on experimental design and data interpretation. X.J., A.Z., Y.K.C., T.C., Y.Z., and M.M. optimised the protocols for autocyclase production and assessed the yields of the systems. X.J., A.Z., Y.K.C., T.C. and Y.Z. optimised the protocols for nanodisc assembly and performed the negative stain imaging. Z.Z. and M.S. performed the molecular dynamic simulations. T.C. and N.L.F. designed and conducted the in vivo nanodiscs study with experimental advice from K.J.T. X.J., Y.Z., T.C. and Y.K.C. and M.M. performed the NMR studies of KvAP and SFTI. X.J. and M.M. wrote the manuscript with contributions from the co-authors. All authors read and approved the final manuscript.

## Competing interests

X.J., Y.K.C., T.C., A.Z, and M.M. have lodged an application for a patent on the described method to produce circular proteins and peptides under Australian Provisional Patent Application Number 2020901892. All other authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s42004-023-00841-5.

**Correspondence** and requests for materials should be addressed to Xinying Jia or Mehdi Mobli.

**Peer review information** *Communications Chemistry* thanks Nicola D'Amelio and the other, anonymous, reviewer for their contribution to the peer review of this work.

**Reprints and permission information** is available at http://www.nature.com/reprints

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.