A conserved β-bulge glycine residue facilitates folding and increases stability of the mouse α-defensin cryptdin-4

Richard J. Clark[1], Thanh Huyen Phan[1], Angela Song[1], André J. Ouellette[2], Anne C. Conibear[1], and K. Johan Rosengren[1]

[1]The University of Queensland School of Biomedical Sciences
[2]University of Southern California Keck School of Medicine

September 25, 2021

Abstract

Defensins are key components of both innate and adaptive immune responses to pathogens. Cryptdins are mouse alpha-defensins that are secreted from Paneth cells in the small intestine and have disulfide-stabilised structures and antibacterial activities against both Gram-positive and Gram-negative bacteria. The folding and three-dimensional structures of alpha-defensins are thought to depend on a conserved glycine residue that forms a β-bulge. Here we investigated the role of this conserved glycine at position 19 of cryptdin-4 (Crp4) in terms of the folding, structure and stability. A Crp4 variant with D-Ala at position 19 folded efficiently, was stabilised by a large number of hydrogen bonds, and resisted proteolysis in simulated intestinal fluid. Although a variant with L-Ala at position 19 was able to adopt the correct fold, it showed less efficient folding and was degraded more rapidly than the D-Ala variant. These results demonstrate the key role that glycine residues can have in folding of bioactive peptides and can provide insights to guide design of stable antimicrobial peptides that fold efficiently.

A conserved β-bulge glycine residue facilitates folding and increases stability of the mouse α-defensin cryptdin-4

Richard J. Clark,[1] Thanh Huyen Phan,[1] Angela Song,[1] André J. Ouellette,[2] Anne C. Conibear,[1*] and K. Johan Rosengren[1*]

[1]The University of Queensland, School of Biomedical Sciences, Brisbane, QLD 4072, Australia.

[2]Department of Pathology and Laboratory Medicine and USC Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA.

[*]Address Correspondence to:

Anne C. Conibear

The University of Queensland, School of Biomedical Sciences, St Lucia 4072, Brisbane, Australia.

E-mail: a.conibear@uq.edu.au

Phone: + 61-7-3365-1738

K. Johan Rosengren

The University of Queensland, School of Biomedical Sciences, St Lucia 4072, Brisbane, Australia.

E-mail: j.rosengren@uq.edu.au

Phone: + 61-7-3365-1403

Abstract

Defensins are key components of both innate and adaptive immune responses to pathogens. Cryptdins are mouse alpha-defensins that are secreted from Paneth cells in the small intestine and have disulfide-stabilised structures and antibacterial activities against both Gram-positive and Gram-negative bacteria. The folding and three-dimensional structures of alpha-defensins are thought to depend on a conserved glycine residue that forms a β-bulge. Here we investigated the role of this conserved glycine at position 19 of cryptdin-4 (Crp4) in terms of the folding, structure and stability. A Crp4 variant with D-Ala at position 19 folded efficiently, was stabilised by a large number of hydrogen bonds, and resisted proteolysis in simulated intestinal fluid. Although a variant with L-Ala at position 19 was able to adopt the correct fold, it showed less efficient folding and was degraded more rapidly than the D-Ala variant. These results demonstrate the key role that glycine residues can have in folding of bioactive peptides and can provide insights to guide design of stable antimicrobial peptides that fold efficiently.

Keywords

antimicrobial peptides, defensins, Nuclear Magnetic Resonance spectroscopy, peptide folding, peptide stability, peptide synthesis

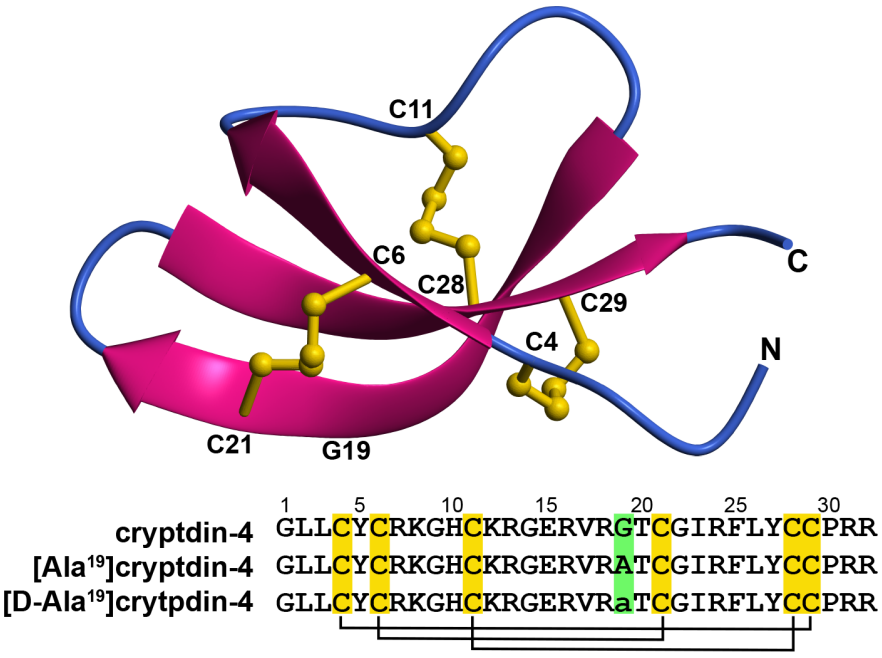
Introduction

Defensin peptides are key mediators of the mammalian innate immune response and have garnered scientific and commercial interest as antimicrobial agents.[1-4] Mammalian defensins are cysteine-rich, cationic, predominantly β-sheet-rich, and have broad-spectrum activities against Gram-positive and -negative bacteria. They are divided into three classes: α-defensins have 32-40 residues and Cys^{I}-Cys^{VI}, Cys^{II}-Cys^{IV}, Cys^{III}-Cys^{V} disulfide bonds; β-defensins have approximately 45 residues and Cys^{I}-Cys^{V}, Cys^{II}-Cys^{IV}, Cys^{III}-Cys^{VI} disulfide bonds; and ϑ-defensins have 18 residues, Cys^{I}-Cys^{VI}, Cys^{II}-Cys^{V}, Cys^{III}-Cys^{IV} disulfide bonds and a head-to-tail cyclic backbone.[5, 6] The three-dimensional structures of defensins are stabilised by their disulfide bonds, enabling them to withstand harsh environments and to maintain their amphipathic nature. Mammalian defensins are typically amphipathic and are thought to act by disrupting bacterial membranes,[7, 8] although alternative mechanisms are likely involved in other activities such as immunomodulatory, antiviral and chemotactic activities.[3, 4, 9, 10]

Mammalian α-defensins are found in neutrophils and in the small intestine, where they are secreted into crypts in the lumen by Paneth cells, sometimes at concentrations of 25-100 mg/mL – far higher than their minimal inhibitory concentrations.[11]Paneth cell α-defensins from mice are termed cryptdins (Crps). Of these cryptdins, Crp4 was reported to have the highest antibacterial activity and has served as a model for understanding the structure-activity relationships of mammalian α-defensins.[11-13]Crp4 is derived from a 73-residue pro-peptide (proCrp4) that is cleaved by matrix metalloproteinase-7 (MMP7, matrilysin) to the 32-residue mature peptide before secretion.[14, 15] Crp4 is rich in Arg residues, which are important for membrane interactions and bioactivity.[16] The unstructured and negatively-charged pro-segment is thought to shield the positively-charged regions of the mature folded peptide, keeping it in an inactive state prior to secretion.[17, 18]

The three-dimensional structure of mature Crp4 contains several structural features that are thought to underpin its antibacterial activity and stability. Three disulfide bonds cross-brace the β-sheet structure (Figure 1) and have been shown to confer protection to the defensin against cleavage by the activating proteinase MMP7.[7] Replacing pairs of Cys residues with Ala did not abrogate *in vitro* antibacterial activity but resulted in cleavage of the mature domain by MMP7. Following initial structural assessment,[19] a subsequent study of the three-dimensional structure of mature Crp4 by NMR spectroscopy identified a salt bridge between Arg7 and Glu15 that is conserved throughout the α-defensin family.[13] When mutated to Asp15 with a shorter side-chain, the overall native fold was retained, with rearrangements of nearby side-chains compensating for disruption of the salt bridge. Neither *in vitro* antibacterial activity nor proteolytic

stability were affected by this change, leading to the hypothesis that the salt bridge might facilitate *in vivo*folding or trafficking.[13] Further mutations of the residues in the salt bridge to Gly were structurally tolerated but mutation of Glu15 to the larger hydrophobic residue Leu15 was not, although, once again, *in vitro* antibacterial activity was maintained.[20] In the latter study, the salt bridge was found to facilitate *in vitro* folding of the α-defensin and resistance to digestion by trypsin.[20] A third structural feature of Crp4 is the β-bulge formed by the conserved residue Gly19 in the second β-strand (Figure 1). β-sheets are common structural motifs in proteins and are stabilised by mainchain hydrogen bonds. β-bulges are irregularities; they are distortions to the classic hydrogen bond pattern as a result of insertion of one or more residues into one of the strands.[21, 22] β-bulges have been classified into five different types based on the hydrogen bond pattern formed and backbone conformation of the residues involved. The hydrogen bond pattern described for the bulge in α-defensins conforms to the "classic" type, with residues Arg18 and Gly19 both donating hydrogen bonds to Tyr27, which in turn donates a hydrogen bond to Gly19.[13] As a Gly residue does not have a sidechain, it can adopt /ψ angle combinations inaccessible to L-amino acids. Glycine is therefore believed to reduce structural strain in the bulge in α-defensins, although it is not a prerequisite for this type of bulge in other proteins.[21, 22] In human neutrophil α-defensin 2 (HNP2), replacement of the corresponding conserved residue (Gly17) with L-Ala17 resulted in a failure to fold, however, replacement with D-Ala or several other D-residues resulted in efficient folding.[23] These results raised the question of the role of the β-bulge in Crp4, and more generally, the structural and functional role of β-bulges, which can alter the direction of β-strands and side chain projections, and prevent aggregation.[21]



Φιγυρε 1. *Στρυστυρε ανδ σεχυενσε οφ μουσε α-δεφενσιν ῾ρπ4 σηοωινγ τηε β-σηεετ στρυστυρε σταβιλισεδ βψ σροσσ-στρανδ δισυλφιδε βονδς (PDB: 2GW9). Gly19 (highlighted in green) was replaced by either L-Ala or D-Ala to study the role of this conserved residue in structure and stability. The structure is shown in cartoon representation with disulfide bonds in yellow. Amino acids are represented by their one-letter codes with disulfide bonds shown as black lines.*

Here we studied the role of the conserved Gly19 residue in the structure, folding and stability of mouse α-defensin Crp4. By replacing the unconstrained Gly residue with either L-Ala or D-Ala, we show that whereas the D-Ala is readily accommodated, L-Ala at this position results in less efficient folding, fewer hydrogen bonds and decreased stability in simulated intestinal fluid. This study helps to complete our understanding

of how the structural features of α-defensins work together to stabilise the overall structure and support antibacterial activity in the challenging environment of the small intestine. The results also illustrate that structure-activity relationships of bioactive peptides might not be as straightforward as correlating overall structure to biological activity; consideration might need to be given to how subtle structural features such as backbone torsion angles, salt bridges and hydrogen bonding influence folding, stability and dynamics and thereby facilitate biological activity.

Materials and methods

Solid phase peptide synthesis

All peptides were synthesized on a CSBio CS336X synthesizer on 2-chlorotrityl chloride (2-CtC) resin (1.0 mmol/g) using 9-fluorenylmethoxycarbonyl (Fmoc)-based solid-phase peptide synthesis. Couplings were performed in N,N'-dimethylformamide (DMF) using 2-(1*H* -benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU, 0.5 M in DMF) with diisopropylethylamine. N-terminal deprotection was achieved with piperidine (20% in DMF). Peptides were cleaved from the resin by treatment with trifluoroacetic acid/triisopropylsilane/water (95:2.5:2.5) for 1.5 hours. The cleavage mixture was then concentrated under vacuum, the peptide precipitated by addition of cold diethyl ether and then collected by filtration. The crude peptide was redissolved in 1:1 solvent A/B (A: H_2O with 0.05% trifluoroacetic acid; B: 90% acetonitrile, 10% H_2O, 0.045% trifluoroacetic acid) and lyophilised. Peptides were purified by reverse-phase HPLC (RP-HPLC) on a C_{18} column using a gradient of 0-80% B in 80 min with UV monitoring at 214 and 280 nm. Electrospray-mass spectroscopy (ES-MS) confirmed the molecular mass of the synthesised peptide and analytical RP-HPLC was used to analyse purity. To form the disulfide bonds, pure reduced peptides were dissolved in 1% sodium acetate (pH 8.0) containing glutathione (2 mol of oxidized glutathione and 1 mol of reduced glutathione per 20 cysteine-SH molar equivalents) at a concentration of 0.3 mg/mL and stirred at room temperature for 24 hours. The oxidised peptides were purified and analysed by RP-HPLC as described for the reduced peptides.

NMR data collection, processing, and structure calculations

[D-Ala^{19}]Crp4 and [Ala^{19}]Crp4 (1 mg) were dissolved in either 550 µL of 90% H_2O/10% deuterium (D_2O) or 100% D_2O at 298 K at pH 4. Sodium 2-2-dimethyl-2-silapentane-5-sulfonate (DSS) was added for internal referencing of chemical shifts.[24] NMR spectra were acquired on an Avance 700 MHz NMR spectrometer equipped with a cryoprobe at 298 K at the University of Queensland, Centre for Advanced Imaging. Standard Bruker pulse programs were used for all two-dimensional spectra. Excitation sculpting with gradients was used to achieve water suppression for TOCSY and NOESY experiments. NMR experiments included TOCSY using a MLEV-17 spin lock sequence with a 80 ms mixing time, NOESY with a 200 ms mixing time, DQF-COSY, E.COSY,^{1}H- ^{13}C HSQC and^{1}H-^{15}N HSQC. Spectra were recorded with 4096 data points in the *F* 2 dimension and 512 increments in the *F* 1 dimension for TOCSY, NOESY, DQF-COSY and E.COSY experiments and 2048 x 240 for^{1}H-^{13}C HSQC and 2048 x 128 for^{1}H-^{15}N HSQC data points in the*F2* dimension and increments in the *F1* dimension, respectively. The *F* 1 dimension was zero-filled to 1024 real data points before Fourier transformation. Spectra were processed in TOPSPIN (Bruker) and assigned with CARA[25] using the sequential assignment method. Secondary shifts were calculated using the random coil values reported by Wishart *et al* .[26] ^{1}H-^{2}H exchange experiments were carried out by adding 99.96% D_2O to lyophilized peptide and immediately acquiring a series of^{1}H and TOCSY experiments at 298 K. Amide protons giving rise to cross-peaks detected in TOCSY spectra after >1 h were considered to be slow exchanging and potentially involved in hydrogen bonds. For all these hydrogen bond donors the corresponding acceptors were identified in early structural models and were used as additional structural restraints during refinement.

Distance restraints for [D-Ala^{19}]Crp4 and [Ala^{19}]Crp4 were derived from the intensity of cross-peaks in NOESY spectra recorded at 298 K with a mixing time of 200 ms. The NOEs were assigned, and their intensities translated automatically into interproton distances using CYANA 3.97.[27] Prediction of the φ and ψ dihedral angles constraints was performed in TALOS-N.[28]The program CYANA was used for automated

NOE assignment and torsion angle dynamics to generate initial structures. Several rounds of structure calculations were performed to resolve distance and angle constraint violations. Using protocols from the RECOORD database,[29] ensembles of 50 structures were subsequently calculated with CNS[30] using the force field distributed with Haddock 2.0[31] and further refined in a water shell. For the structure calculations involving D-Ala, a CYANA library file was generated from [Ala^{19}]Crp4. For the CNS calculations, the stereochemistry was switched using a command that changes the order of the atoms in the 'improper' parameters. Final sets of structures with the lowest energy and no NOE violations greater than 0.2 Å or dihedral- angle violations greater than 2° were selected for MolProbity analyses[32] . The best 20 structures were visualized using MOLMOL[33] and figures generated using MOLMOL and PyMOL.[34] The structures of [D-Ala^{19}]Crp4 and [Ala^{19}]Crp4 were deposited in the Protein Data Bank under accession numbers 7RC8 and 7RC7, respectively.

Stability assays

Simulated intestinal fluid (SIF) was prepared by combining 68 mg of KH_2PO_4 in 250 uL of MilliQ water, 770 µL of 0.2 M NaOH in 5 mL of water, and 100 µg of pancreatin. The pH was then adjusted to pH 6.8 with 0.2 M NaOH and the solution diluted to 10 mL with MilliQ water. Peptide was then dissolved in 500 µL of SIF at a concentration of 0.1 mg/mL. The solution was incubated at 37°C and 50 µL aliquots taken at 0 min, 5 min, 20 min, 1 h, 2 h, 4 h, 6 h, 8 h and 24 h. Aliquots were quenched by addition of 50 µL of 4% aqueous trifluoroacetic acid and stored at -20°C until analysis by LC-MS/MS on an API2000 (AB Sciex).

Results and Discussion

Synthesis, purification and folding of Crp4 analogues

The peptide chains of both [Ala^{19}]Crp4 and [D-Ala^{19}]Crp4 were successfully assembled by Fmoc-based solid phase peptide synthesis and purified using RP-HPLC followed by ES-MS and analytical HPLC to confirm their identity and purity (Figure 2a). For oxidative folding, peptides were incubated under basic conditions in 1% sodium acetate in the presence of reduced and oxidised glutathione. To monitor the progress and efficiency of folding, aliquots of the reaction mixture were taken at 2, 4 and 22 h timepoints and analysed by analytical RP-HPLC as the peptide peak shifts to a shorter retention time upon oxidation. As shown in Figure 2b-d, both [Ala^{19}]Crp4 and [D-Ala^{19}]Crp4 folded correctly, indicated by the appearance of a sharp peak eluting at a retention time of approximately 26 min. The RP-HPLC peaks were integrated to compare efficiency of folding of the two Crp4 analogues: [D-Ala^{19}]Crp4 folded more efficiently [34.7% (2 h); 44.5% (4 h); 47.7% (22 h)] than [Ala^{19}]Crp4 [5.5% (2 h); 12.4% (4 h); 26.6% (22 h)]. Nevertheless, both folded analogues were subsequently isolated at high purity (> 97%) and mass spectra were consistent with the formation of three disulfide bonds in each peptide (Figure 2e).

The ability of [Ala^{19}]Crp4 to fold correctly, albeit at reduced efficiency compared to [D-Ala^{19}]Crp4, demonstrates that an L-amino acid can be tolerated at this position. Although the same final fold is attained, the [Ala^{19}]Crp4 variant folds more slowly, with lower yield and is less stable than [D-Ala^{19}]Crp4. This result is in contrast to the finding in a similar study on HNP2 that replacement of the corresponding conserved residue (Gly17) with L-Ala17 resulted in a complete inability to fold, and led to aggregation.[23] Although different folding conditions were used for HNP2, this result might indicate that Crp4 is more tolerant of mutations in this position than HNP2. As confirmed by NMR spectroscopy, the native Cys^{I}-Cys^{VI}, Cys^{II}-Cys^{IV}, Cys^{III}-Cys^{V} disulfide connectivity and α-defensin fold are favoured during undirected folding conditions and no mis-folded disulfide isomers were identified. However, as shown in Figure 2d, neither [Ala^{19}]Crp4 nor [D-Ala^{19}]Crp4 reached a 100% folding yield, suggesting that folding conditions could be further optimised. Furthermore, it is not known whether *in vivo* folding is assisted by other proteins, or whether folding proceeds via shuffling of disulfide bonds to reach the thermodynamically stable native conformation.

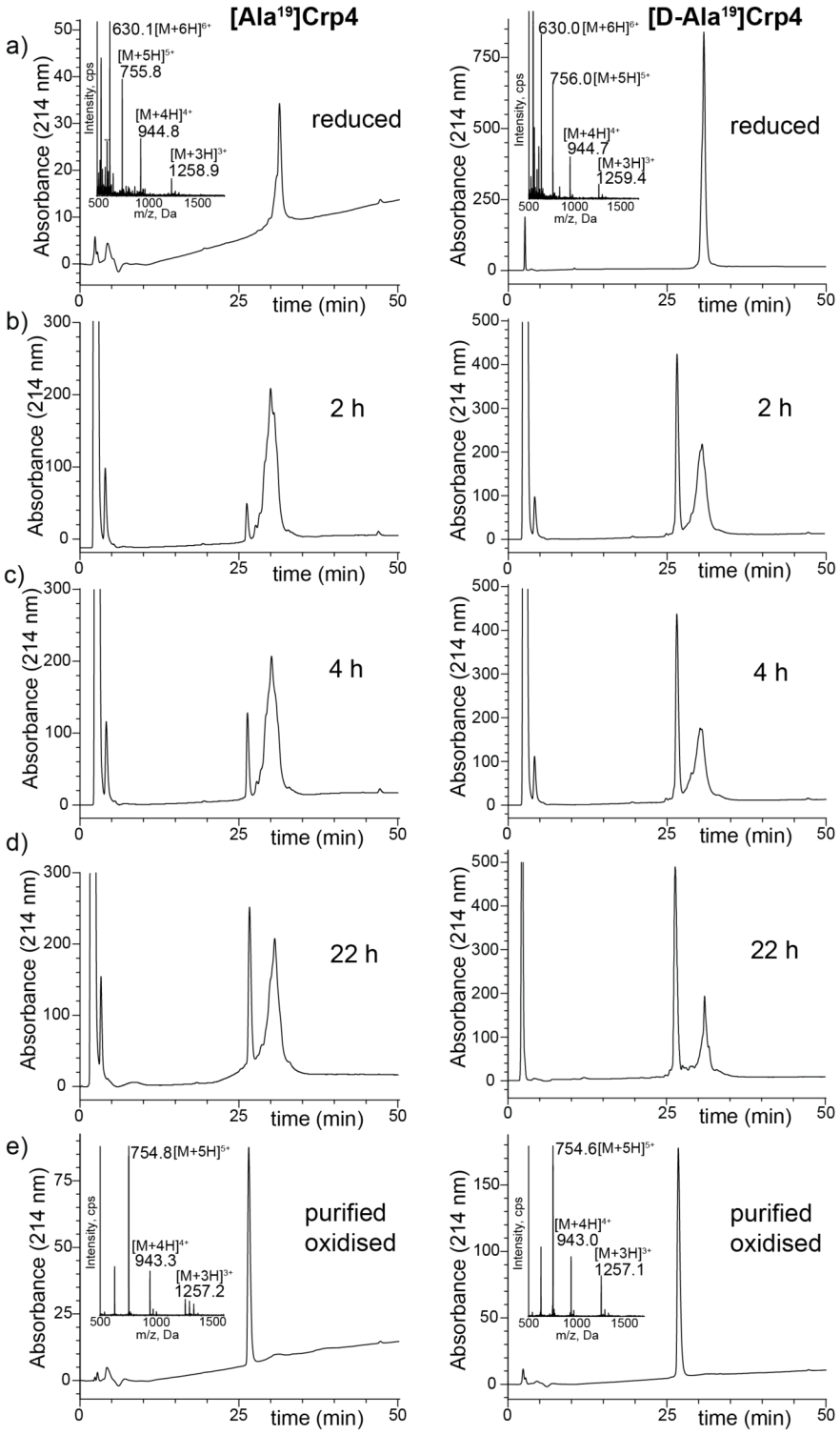


Figure 2. *Purification and folding of [Ala^{19}]Crp4 and [D-Ala^{19}]Crp4. a) RP-HPLC traces and mass spectra (inset) of purified reduced peptides [Ala^{19}]Crp4 ([M+3H]$^{3+}{}_{calc}$ = 1259.3; [M+3H]$^{3+}{}_{obs}$ = 1258.9) and [D-Ala^{19}]Crp4 ([M+3H]$^{3+}{}_{calc}$ = 1259.3; [M+3H]$^{3+}{}_{obs}$ = 1259.4). b)-d) RP-HPLC traces following the*

timecourse of folding (oxidation). d) RP-HPLC traces and mass spectra (inset) of purified oxidised peptides [Ala[19]]Crp4 ([M+3H]$^{3+}{}_{calc}$ = 1257.5; [M+3H]$^{3+}{}_{obs}$ = 1257.2) and [D-Ala[19]]Crp4 ([M+3H]$^{3+}{}_{calc}$ = 1257.5; [M+3H]$^{3+}{}_{obs}$ = 1257.1).

Structural analysis by NMR spectroscopy

The purified Crp4 analogues gave dispersed ^{1}H NMR spectra, implying that they adopt ordered structures in solution. Both analogues were further analysed by two-dimensional NMR to confirm that they had correctly formed the characteristic α-defensin fold and disulfide connectivity.[13, 19] To assign the spectra, Hα-NH$_{i+1}$ connectivities obtained from the NOESY spectra were used in the sequential assignment of individual spin systems determined from the TOCSY spectra.[35]Annotated TOCSY and NOESY data are provided in the supplementary data for [Ala[19]]Crp4 and [D-Ala[19]]Crp4, and in Jing *et al.* (2004) for native Crp4.[19] For both analogues, sequential Hα-NH$_{i+1}$ connectivities were observed for the entire peptide chain, with the exception of Pro[30], which lacks an amide proton. A strong NOE correlation between the Cys[29] Hα proton and the Pro[30] Hδ protons indicated that the Cys[29]-Pro[30] amide bond was in the*trans* configuration. Significantly, the chemical shift of the Hε side chain proton of Arg[7] is at 9.6 ppm in both [Ala[19]]Crp4 and [D-Ala[19]]Crp4, matching that of native Crp4 Arg[7] Hε (9.61 ppm)[13] and indicative of a key salt-bridge between Arg[7] and Glu[15]. Secondary Hα chemical shifts were determined by comparison with the respective random coil shifts to indicate secondary structural elements.[26] Figure 3 shows the comparison of the Hα secondary shifts for [Ala[19]]Crp4 and [D-Ala[19]]Crp4 in comparison to published secondary shifts of native Crp4 (PDB: 2GW9).[13] The Hα secondary shifts for all three peptides show the same trends and the positive values for residues 4 – 7, 15 – 21, and 25 – 30 indicate that the triple-stranded β-sheet is conserved in the L-Ala[19] and D-Ala[19] analogues.

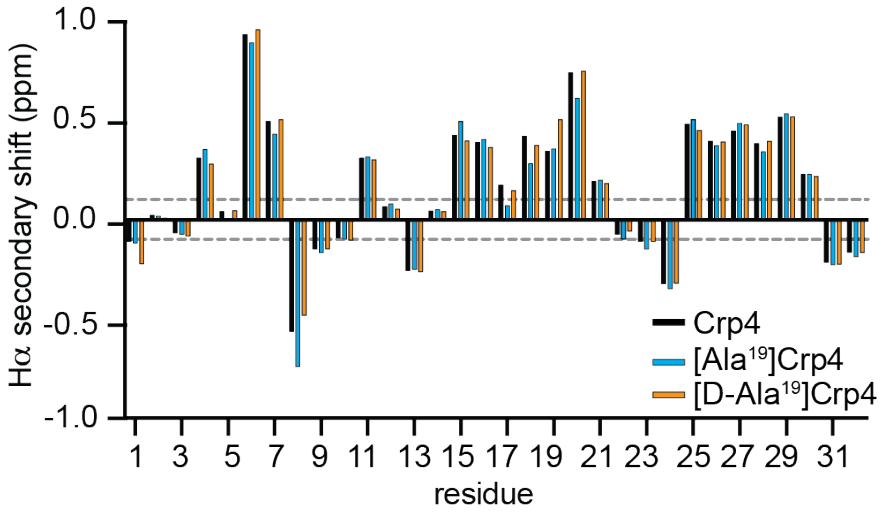
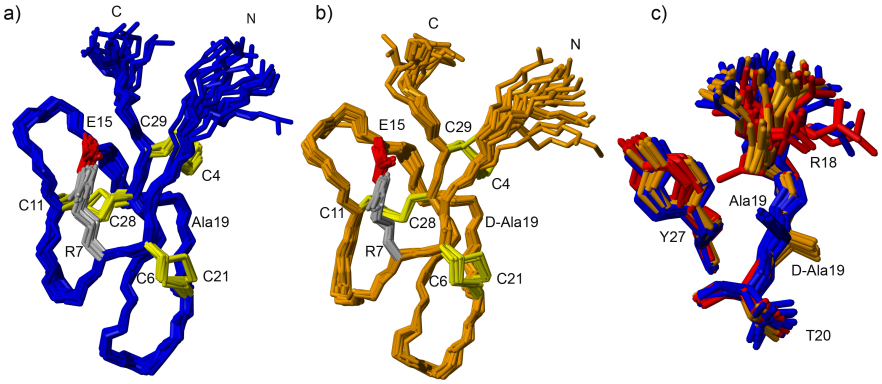


Figure 3. *Secondary Hα chemical shifts of native Crp4 (black, PDB: 2GW9),[13][Ala[19]]Crp4 (blue) and [D-Ala[19]]Crp4 (orange) showing conserved β-sheet regions (sequences of positive Hα secondary shifts).*

Three-dimensional structures of [Ala[19]]Crp4 and [D-Ala[19]]Crp4 were calculated in CYANA[27] and refined in CNS[30] using TALOS-N-derived dihedral angle restraints,[36] and hydrogen bond restraints derived from D_2O exchange experiments. The structural ensembles of the 20 representative structures for each analogue are shown in Figure 4 and statistics for the structures are provided in Table 1.



Φιγυρε 4. *Στρυστυρες οφ (α) [Αλα^{19}]ᶜrπ4 (βλυε, ΠΔΒ:7Pᶜ7) ανδ (β) [Δ-Αλα^{19}]ᶜrπ4 (ορανγε, ΠΔΒ: 7Pᶜ8). Συπερποσιτιον οφ τηε στρυστυραλ ενσεμβλες αρε σηοων ωιτη δισυλφιδες ιν ψελλοω, Αργ^{7} ιν γρεψ ανδ Γλυ^{15} ιν ρεδ φορμινγ τηε σαλτ βριδγε. Κεψ ρεσιδυες αρε λαβελλεδ ωιτη ρεσιδυε νυμβερς. (ς) Συπερποσιτιον οφ τηε β-βυλγε ρεγιον οφ [Αλα^{19}]ᶜrπ4 (βλυε), [Δ-Αλα^{19}]ᶜrπ4 (ορανγε) ανδ ᶜrπ4 (ρεδ, ΠΔΒ: 2ΓΩ9), ηιγηλιγητινγ τηε μινιμαλ διφφερενσες ιν βασκβονε σονφορματιον ανδ προθεστιονς οφ σιδεσηαιν, εξσεπτ τηε ρεσιδυε 19 μετηψλ γρουπς ρεσυλτινγ φρομ τηε ινερτεδ στερεοσηεμιστρψ.*

Table 1. NMR distance and dihedral statistics for [Ala^{19}]Crp4 and [D-Ala^{19}]Crp4

	[Ala^{19}]Crp4	**[D-Ala^{19}]Crp4**
Distance restraints		
Sequential (|i-j| [?] 1)	267	311
Medium range (1 < |i-j| < 5)	29	56
Long range (|i-j| [?] 5)	93	112
Hydrogen bonds	20	24
Total	409	503
Dihedral angle restraints		
Φ	21	24
Ψ	19	22
Total	40	46
Energies (kcal/mol)[a)]		
Overall	-908.93 ± 37.64	-897.71 ± 26.78
Bonds	14.61 ± 1.44	20.24 ± 1.23
Angles	41.91 ± 4.08	50.48 ± 3.96
Improper	15.49 ± 2.10	15.63 ± 1.39
Dihedral	143.04 ± 2.08	146.28 ± 3.11
Van der Waals	-153.34 ± 4.65	-136.13 ± 4.82
Electrostatic	-971.35 ± 43.20	-995.07 ± 26.22
NOE (experimental)	0.02 ±0.01	0.17 ± 0.03
Constrained dihedrals (experimental)	0.68 ± 0.37	0.69 ± 0.46
Atomic RMSD (Å)[a)]		
Mean global backbone (3-30)	0.52 ± 0.09	0.42 ± 0.08
Mean global heavy atoms (3-30)	1.72 ± 0.23	1.52 ± 0.26
Molprobity[a)]		
Clash Score, all atoms	6.74 ± 2.85	9.77 ± 3.91
Poor rotamers	0 ± 0	0.58 ± 0.51
Favored rotamers (%)	97.04 ± 3.09	83.15 ± 4.25
Ramachandran Outliers (%)	0 ± 0	3.50 ± 0.75[b)]

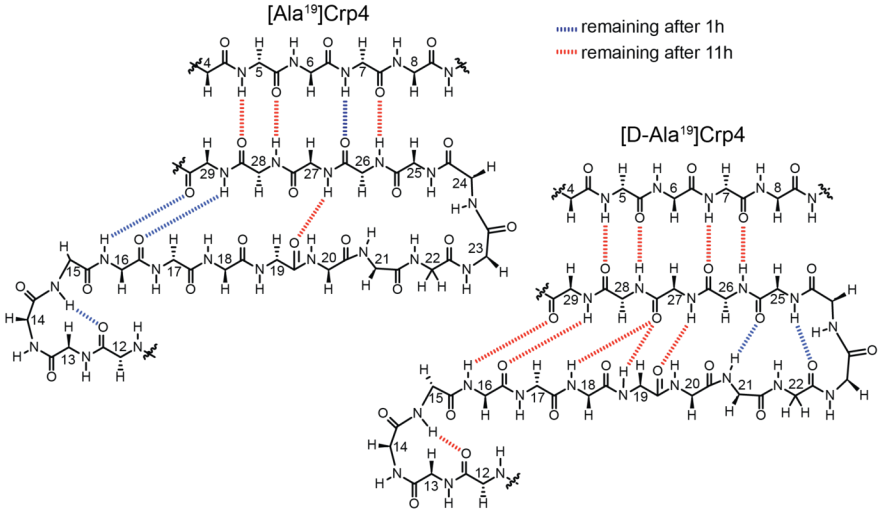
	[Ala^{19}]Crp4	**[D-Ala^{19}]Crp4**
Ramachandran Favored (%)	98.67 ± 2.51	96.67 ± 1.84
MolProbity score	1.41 ± 0.16	2.01 ± 0.15
MolProbity score percentile	96.21 ± 2.35	74.00 ± 6.51
Violations from experimental restraints		
NOE violations exceeding 0.2 Å	0	0
Dihedral violations exceeding 2.0°	0	0

[a] Statistics are given as mean ± standard deviation.

[b] D-Ala is considered an outlier by Molprobity but does in fact adopt a 'Favored' conformation for a D-amino acid. This also affects the overall Molprobity score.

As predicted by the secondary Hα shifts, both [Ala^{19}]Crp4 and [D-Ala^{19}]Crp4 structures show the characteristic three-stranded β-sheet structure cross-braced by disulfide bonds. For the [Ala^{19}]Crp4 structure, however, significantly fewer distance and dihedral angle restraints were obtained than for [D-Ala^{19}]Crp4 and higher RMSD values (Table 1), indicating that the [Ala^{19}]Crp4 structure is less stable and well-defined than the [D-Ala^{19}]Crp4 structure. These results are consistent with the respective folding efficiencies and show that, while L-Ala is tolerated at position 19, it is unfavourable for folding and structure stabilisation. The overall difference in the backbone conformation between the two variants is however relatively limited, with the D-Ala^{19} adopting a φ/ψ combination of +160/-150 (a region typically only favoured by Gly) and the [Ala^{19}] -160/180 (a region characteristic of residues in a β-sheet).

Exchange of the H_N protons on dissolution of the lyophilized peptides in D_2O was monitored by acquiring ^1H and TOCSY spectra at 1 h and 11 h after dissolution to examine further the differences in the structures of [Ala^{19}]Crp4 and [D-Ala^{19}]Crp4. Comparison of the observable H_N resonances at each timepoint provides an indication of the comparative stability of the hydrogen bond network within the triple-stranded β-sheet of each peptide. As shown in Figure 5, there were substantially more observable H_N protons in the TOCSY spectrum of [D-Ala^{19}]Crp4 than [Ala^{19}]Crp4 after 11 h, adding further evidence that the β-sheet structure and the overall fold of [D-Ala^{19}]Crp4 is more stable than that of [Ala^{19}]Crp4.



Φιγυρε 5. *Βασκβονε στρυστυρε οφ β-σηεετ ρεγιονς οφ [Αλα^{19}]ῖρπ4 ανδ [Δ-Αλα^{19}]ῖρπ4 σηοωινγ ηψδρογεν βονδς τηατ περσιστ αφτερ 1 η (βλυε) ανδ 11 η (ρεδ) αφτερ δισσολυτιον οφ τηε λψοπηλισεδ πεπτιδε ιν $Δ_2O$.*

Stability to proteolytic cleavage and relevance for antimicrobial function

Mutational studies have shown that for all classes of mammalian defensins, despite their highly optimised and disulfide-stabilised fold, the 3D structure is largely decoupled from direct antimicrobial activity; removal of disulfide bonds or other structural features often has minimal or unpredictable effects on activity *in vitro*.[7, 8, 10, 20] In fact, for human β-defensin-1 the disulfide array even serves to inhibit activity, and it is only under specific reducing conditions that antimicrobial activity is revealed.[37] While some defensins have additional specific functions such as chemotaxis, which do rely on correct structure for receptor interactions,[10]it is likely that in many cases the optimised structure only serves to protect the peptide from degradation. The harsh environment experienced by Paneth cell α-defensins upon secretion into the small intestine means that they need to be able to withstand basic pH and proteolytic digestion to carry out their antibacterial functions. We compared the stabilities of [Ala^{19}]Crp4 and [D-Ala^{19}]Crp4 to proteolytic degradation in simulated intestinal fluid to determine whether the structural stability differences observed correspond to proteolytic stability.[38] The peptides were incubated in simulated intestinal fluid and aliquots were analysed by LS-MS/MS over a 24 h period to determine the % peptide remaining. As shown in Figure 6, [D-Ala^{19}]Crp4 was more stable than [Ala^{19}]Crp4; after 24 h, 42% of [D-Ala^{19}]Crp4 remained, whereas [Ala^{19}]Crp4 had been completely degraded. These results indicate that structural stability correlates with proteolytic stability and that the β-bulge enables Crp4 to carry out its antibacterial roles in the natural environment of the intestine. This is analogous to the data from studies on salt-bridge deficient variants, which, while able to retain a native overall fold and resistance to the protease MMP-7 during processing,[13] still show impaired folding and, because of increased flexibility, are susceptible to degradation by trypsin once secreted into the small intestine.[20] Thus, we conclude that all key conserved structural features of α-defensins – the disulfide-array, the salt-bridge and the glycine-mediated β-bulge – serve to ensure efficient folding, structural stability and protease resistance *in vivo* of Crp4.

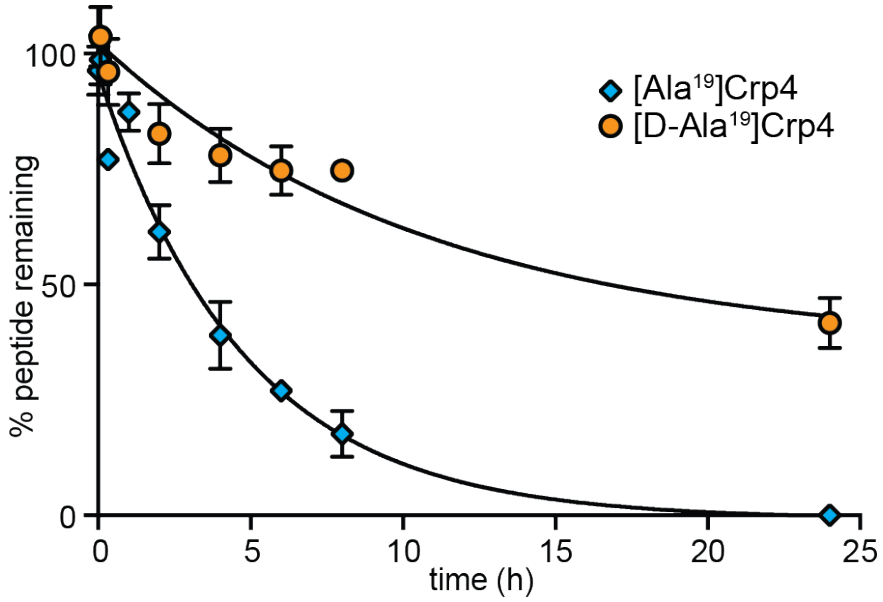


Figure 6. *Stabilities of [Ala^{19}]Crp4 (blue diamonds) and [D-Ala^{19}]Crp4 (orange circles) in simulated intestinal fluid.*

In summary, we have used solid phase peptide synthesis, detailed structural analysis by NMR spectroscopy and proteolytic stability assays to compare the folding, structures and stabilities of two Crp4 analogues in which the conserved Gly^{19} is mutated to either L-Ala^{19} or D-Ala^{19}. Replacement of the conformationally unrestricted Gly^{19} by L-Ala^{19} is tolerated, however, folding efficiency, structural stability and proteolytic

stability are compromised in comparison to the D-Ala[19] analogue. These results concur with previous analyses of other structural features in providing evidence that structural stability is key to persistence of defensin peptides in their biological contexts and is a feature that might have evolved to support their antibacterial activities. From a broader perspective, it is interesting to note that while Gly residues are the smallest proteinogenic amino acid residues and are unique in their additional flexibility in terms of adopting different conformations, in some cases Gly residues can serve to reduce strain and increase protein rigidity.

Acknowledgements

We are grateful to Colton D. Payne for computational assistance. ACC is supported by a University of Queensland Development Fellowship (Project 613982) and an Early Career Researcher grant (Project 616535) from The University of Queensland.

Conflict of Interest

The authors declare no conflict of interest

ORCIDs

Richard J. Clark: 0000-0002-6807-5426

Anne C. Conibear: 0000-0002-5482-6225

K. Johan Rosengren: 0000-0002-5007-8434

André J. Ouellette: 0000-0002-7987-5779

References

[1] Haney, E. F., Mansour, S. C., and Hancock, R. E. (2017) Antimicrobial Peptides: An Introduction, *Methods Mol. Biol.1548* , 3-22.

[2] Avila, E. E. (2017) Functions of Antimicrobial Peptides in Vertebrates, *Curr. Protein Pept. Sci. 18* , 1098-1119.

[3] van der Does, A. M., Hiemstra, P. S., and Mookherjee, N. (2019) Antimicrobial Host Defence Peptides: Immunomodulatory Functions and Translational Prospects, *Adv. Exp. Med. Biol. 1117* , 149-171.

[4] Solanki, S. S., Singh, P., Kashyap, P., Sansi, M. S., and Ali, S. A. (2021) Promising role of defensins peptides as therapeutics to combat against viral infection, *Microb. Pathog. 155* , 104930.

[5] Selsted, M. E., and Ouellette, A. J. (2005) Mammalian defensins in the antimicrobial immune response, *Nat. Immunol. 6* , 551-557.

[6] Conibear, A. C., and Craik, D. J. (2014) The chemistry and biology of theta defensins, *Angew. Chem. Int. Ed. Engl.53* , 10612-10623.

[7] Maemoto, A., Qu, X., Rosengren, K. J., Tanabe, H., Henschen-Edman, A., Craik, D. J., and Ouellette, A. J. (2004) Functional analysis of the alpha-defensin disulfide array in mouse cryptdin-4,*J. Biol. Chem. 279* , 44188-44196.

[8] Conibear, A. C., Rosengren, K. J., Daly, N. L., Henriques, S. T., and Craik, D. J. (2013) The cyclic cystine ladder in theta-defensins is important for structure and stability, but not antibacterial activity, *J. Biol. Chem. 288* , 10830-10840.

[9] Tongaonkar, P., Punj, V., Subramanian, A., Tran, D. Q., Trinh, K. K., Schaal, J. B., Laragione, T., Ouellette, A. J., Gulko, P. S., and Selsted, M. E. (2019) RTD-1 therapeutically normalizes synovial gene signatures in rat autoimmune arthritis and suppresses proinflammatory mediators in RA synovial fibroblasts, *Physiol. Genomics51* , 657-667.

[10] Wu, Z., Hoover, D. M., Yang, D., Boulègue, C., Santamaria, F., Oppenheim, J. J., Lubkowski, J., and Lu, W. (2003) Engineering disulfide bridges to dissect antimicrobial and chemotactic activities of human beta-defensin 3, *Proc. Natl. Acad. Sci. USA 100* , 8880-8885.

[11] Ayabe, T., Satchell, D. P., Wilson, C. L., Parks, W. C., Selsted, M. E., and Ouellette, A. J. (2000) Secretion of microbicidal alpha-defensins by intestinal Paneth cells in response to bacteria,*Nat. Immunol. 1* , 113-118.

[12] Ouellette, A. J., Hsieh, M. M., Nosek, M. T., Cano-Gauci, D. F., Huttner, K. M., Buick, R. N., and Selsted, M. E. (1994) Mouse Paneth cell defensins: primary structures and antibacterial activities of numerous cryptdin isoforms, *Infect. Immun. 62* , 5040-5047.

[13] Rosengren, K. J., Daly, N. L., Fornander, L. M., Jönsson, L. M., Shirafuji, Y., Qu, X., Vogel, H. J., Ouellette, A. J., and Craik, D. J. (2006) Structural and functional characterization of the conserved salt bridge in mammalian paneth cell alpha-defensins: solution structures of mouse CRYPTDIN-4 and (E15D)-CRYPTDIN-4, *J. Biol. Chem. 281* , 28068-28078.

[14] Wilson, C. L., Ouellette, A. J., Satchell, D. P., Ayabe, T., López-Boado, Y. S., Stratman, J. L., Hultgren, S. J., Matrisian, L. M., and Parks, W. C. (1999) Regulation of intestinal alpha-defensin activation by the metalloproteinase matrilysin in innate host defense,*Science 286* , 113-117.

[15] Shirafuji, Y., Tanabe, H., Satchell, D. P., Henschen-Edman, A., Wilson, C. L., and Ouellette, A. J. (2003) Structural determinants of procryptdin recognition and cleavage by matrix metalloproteinase-7,*J. Biol. Chem. 278* , 7910-7919.

[16] Llenado, R. A., Weeks, C. S., Cocco, M. J., and Ouellette, A. J. (2009) Electropositive charge in alpha-defensin bactericidal activity: functional effects of Lys-for-Arg substitutions vary with the peptide primary structure, *Infect. Immun. 77* , 5035-5043.

[17] Figueredo, S. M., Weeks, C. S., Young, S. K., and Ouellette, A. J. (2009) Anionic amino acids near the pro-alpha-defensin N terminus mediate inhibition of bactericidal activity in mouse pro-cryptdin-4,*J. Biol. Chem. 284* , 6826-6831.

[18] Figueredo, S. M., and Ouellette, A. J. (2010) Inhibition of bactericidal activity is maintained in a mouse alpha-defensin precursor with proregion truncations, *Peptides 31* , 9-15.

[19] Jing, W., Hunter, H. N., Tanabe, H., Ouellette, A. J., and Vogel, H. J. (2004) Solution Structure of Cryptdin-4, a Mouse Paneth Cell α-Defensin, *Biochemistry 43* , 15759-15766.

[20] Andersson, H. S., Figueredo, S. M., Haugaard-Kedström, L. M., Bengtsson, E., Daly, N. L., Qu, X., Craik, D. J., Ouellette, A. J., and Rosengren, K. J. (2012) The α-defensin salt-bridge induces backbone stability to facilitate folding and confer proteolytic resistance,*Amino Acids 43* , 1471-1483.

[21] Craveur, P., Joseph, A. P., Rebehmed, J., and de Brevern, A. G. (2013) β-Bulges: extensive structural analyses of β-sheets irregularities, *Protein Sci. 22* , 1366-1378.

[22] Leader, D. P., and Milner-White, E. J. (2021) Identification and characterization of two classes of G1 β-bulge, *Acta Crystallogr. D Struct. Biol. 77* , 217-223.

[23] Xie, C., Prahl, A., Ericksen, B., Wu, Z., Zeng, P., Li, X., Lu, W. Y., Lubkowski, J., and Lu, W. (2005) Reconstruction of the conserved beta-bulge in mammalian defensins using D-amino acids, *J. Biol. Chem. 280* , 32921-32929.

[24] Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) 1H, 13C and 15N chemical shift referencing in biomolecular NMR, *J. Biomol. NMR 6* , 135-140.

[25] Keller, R. (2004) The Computer Aided Resonance Assignment Tutorial, *Cantina Verlag* .

[26] Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects,*J. Biomol. NMR 5* , 67-81.

[27] Güntert, P. (2004) Automated NMR structure calculation with CYANA, *Methods Mol. Biol. 278* , 353-378.

[28] Shen, Y., and Bax, A. (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks, *J. Biomol. NMR 56* , 227-241.

[29] Nederveen, A. J., Doreleijers, J. F., Vranken, W., Miller, Z., Spronk, C. A. E. M., Nabuurs, S. B., Güntert, P., Livny, M., Markley, J. L., Nilges, M., Ulrich, E. L., Kaptein, R., and Bonvin, A. M. J. J. (2005) RECOORD: A recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank, *Proteins59* , 662-672.

[30] Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D Biol. Crystallogr. 54* , 905-921.

[31] Dominguez, C., Boelens, R., and Bonvin, A. M. (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information, *J. Am. Chem. Soc. 125* , 1731-1737.

[32] Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr. D Biol. Crystallogr. 66* , 12-21.

[33] Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures, *J. Mol. Graphics 14* , 51-55.

[34] PyMOL. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. .

[35] Schroeder, C. I., and Rosengren, K. J. (2020) Three-Dimensional Structure Determination of Peptides Using Solution Nuclear Magnetic Resonance Spectroscopy, *Methods Mol. Biol. 2068* , 129-162.

[36] Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, *J. Biomol. NMR 44* , 213-223.

[37] Schroeder, B. O., Wu, Z., Nuding, S., Groscurth, S., Marcinowski, M., Beisner, J., Buchner, J., Schaller, M., Stange, E. F., and Wehkamp, J. (2011) Reduction of disulphide bonds unmasks potent antimicrobial activity of human β-defensin 1, *Nature 469* , 419-423.

[38] Kong, X. D., Moriya, J., Carle, V., Pojer, F., Abriata, L. A., Deyle, K., and Heinis, C. (2020) De novo development of proteolytically resistant therapeutic peptides for oral administration, *Nat. Biomed. Eng. 4* , 560-571.