Inter-hairpin linker sequences determine the structure of the ββ-solenoid fold: a "bottom-up" study of pneumococcal LytA choline-binding module

Beatriz Maestro [a,b,¥], *Héctor Zamora-Carreras* [c,d,¥], *M. Ángeles Jiménez* [c,*], *Jesús M. Sanz* [a,b,e,*]

[a] Centro de Investigaciones Biológicas Margarita Salas. Spanish National Research Council (CSIC), Madrid, Spain.

[b] Instituto de Investigación, Desarrollo e Innovación en Biotecnología Sanitaria de Elche (IDiBE), Universidad Miguel Hernández, Elche, Spain.

[c] Instituto de Química-Física "Rocasolano", Spanish National Research Council (CSIC), Madrid, Spain.

[d] Department of Immunology, Ophthalmology and ENT, School of Medicine, Universidad Complutense de Madrid, Madrid, Spain

[e] Centro de Investigación Biomédica en Red de Enfermedades Respiratorias (CIBERES), Madrid, Spain

[¥] Equally contributing authors

[*]Equally contributing and corresponding authors

JMS: Centro de Investigaciones Biológicas Margarita Salas, c/ Ramiro de Maeztu, 9, 28040-Madrid, Spain; E-mail: jmsanz@cib.csic.es
MAJ: Instituto de Química-Física "Rocasolano", c/ Serrano 119, 28006-Madrid, Spain; E-mail: majimenez@iqfr.csic.es

Abstract

The ββ-solenoid structures are part of many proteins involved in the recognition of bacterial cell wall. They are elongated polypeptides consisting of repeated β-hairpins connected by linker sequences and disposed around a superhelical axis stabilised by short-range interactions. Among the most studied ββ-solenoids are those belonging to the family of choline-binding modules (CBMs) from the respiratory pathogen *Streptococcus pneumoniae* (pneumococcus) and its bacteriophages, and their properties have been employed to develop several biotechnological and biomedical tools. We have carried out a theoretical, spectroscopic and thermodynamic study of the ββ-solenoid structure of the CBM from the pneumococcal LytA autolysin using peptides of increasing length containing 1-3 repeats of this structure. Our results show that hints of native-like tertiary structure are only observed with a minimum of three β-hairpins, corresponding to one turn of the solenoid superhelix, and identify the linker sequences between hairpins as the major directors of the solenoid folding. This study paves the way for the rational structural engineering of ββ-solenoids aimed to find novel applications.

Keywords: β-hairpin; repeat proteins; *Streptococcus pneumoniae*; NMR; tryptophan peptides; choline-binding proteins

1. Introduction

The ββ-solenoid spiral staircase is a subset of the β-solenoid protein fold, a superhelical arrangement of repeated β-strands [1] . The distinctive trait of the ββ-solenoid is that the repeating unit is a β-hairpin instead of an isolated β-strand. The term was first coined to describe the structure of the choline-binding module of the LytA autolysin from the respiratory pathogen *Streptococcus pneumoniae* (pneumococcus) (CBM-LytA hereafter) [2]. Since then, and with variations on superhelical curvature and repeats per turn, this supersecondary motif has been found in the structure of several other proteins such as the DSR-M and DSR-V dextransucrases from *Leuconostoc citreum* [3], an alternansucrase from *L. mesenteroides* [4] and toxins A and B from *Clostridium difficile* [5,6]. However, most of the characterized ββ-solenoids have been found in the choline-binding protein (CBP) family of bacterial surface proteins coded by *S. pneumoniae* and many of its bacteriophages, such as the LytA autolysin [7].

The CBPs display a characteristic organisation with a functional module responsible for their diverse biological roles (ranging from daughter cell separation upon division to adhesion to the host cells), and a common choline-binding module (CBM) that helps properly place the CBP on the cell wall through the recognition of the choline groups present in the teichoic and lipoteichoic acids [7]. In turn, CBMs are composed of a variable number (3-18) of aromatic-rich sequences termed choline-binding repeats (CBRs, Pfam PF01473 sequence set, https://pfam.xfam.org/family/PF01473) which are built from a 12-residue β-hairpin plus a 8-10 residue linker sequence [7]. The standard ββ-solenoid of the CBMs is configured by the repetition of CBRs with a 120º left-handed rotation leading to a superhelical arrangement with exactly 3 CBRs per turn [2,8,9] (Fig. 1A). Binding of choline occurs between two consecutive repeats and is mainly driven by hydrophobic and cation-π interactions with the methyl substituents and the quaternary

amine group of the ligand, respectively. The binding site is configured by two aromatic residues from one hairpin and another one from the next (Fig. 1A) [2,7–9], while a fourth aromatic would contribute to keep the cited side chains in place through a stabilising T-shaped π- π stacking interaction [7,10] .

CBM-LytA is the best characterized ββ-solenoid polypeptide in terms of structure, folding and stability [7,10–17]. It contains 6 CBRs plus a C-terminal β-hairpin tail that deviates from the superhelical axis (Fig. 1A) and that is involved in dimerization upon choline binding [2,9]. As a consequence of the elongated shape of the module, the central CBRs only establish contacts with their neighbouring motifs in the sequence, without any further long-range interactions with the rest of the protein. A truncated version of CBM-LytA (C-LytA protein, lacking the β-hairpin of CBR1) presents the accumulation of partially folded intermediates when subjected to thermal [12,16,18] or chemical denaturation either by guanidine hydrochloride [16] or submicellar concentrations of SDS [13]. Moreover, peptides containing the β-hairpin sequences of CBRs 2 and 4 constitute autonomous folding units able to acquire a native-like structure in solution that is still competent to bind choline, although with poor affinity as only a half binding site is provided [19–21].

Although the affinity of CBMs for free choline in solution is only marginal - in the millimolar range [10,12,18,22–24]- , the solenoid arrangement of tandem CBRs leads to a high avidity efficiency in the recognition of the polydentate, choline-containing cell wall macromolecular substrate [10]. This has led to several biomedical and biotechnological applications of these peptides. For instance, we have recently shown that the exogenous addition of CBMs to pneumococcal planktonic cultures blocks the accessibility of host CBPs to the choline residues in the cell wall, preventing daughter cell separation upon division and leading to cell chaining and bacterial aggregates that

are easily phagocytosed [10]. This points to these modules as valuable tools for the development of novel streamlined antimicrobial treatments, selective bacterial diagnostic methods or targeted drug delivery procedures. On another level, this robust multivalent interaction allows the use of the CBM-LytA as an efficient affinity tag for recombinant protein immobilization and purification in solid matrices functionalized with choline or choline analogues [25–28] or in aqueous two-phase systems [29], as well as for tissue imaging using choline dendrimer nanoparticles [30]. All these applications point to the CBMs as attractive tools amenable to be rationally engineered for conformational stability and macromolecular substrate affinity, which requires a thorough study of the structural determinants governing their folding.

Here we have developed a bottom-up basic structural study of the central core of CBM-LytA as a representative polypeptide of the ββ-solenoid fold. Synthetic peptides containing 1-3 CBRs (up to one complete turn of the solenoid staircase) were studied by theoretical analyses, spectroscopy (fluorescence, circular dichroism and NMR) and thermal denaturation techniques. Our findings help to better understand the characteristic organization of ββ-solenoid structures and provide clues for the rational design of novel variants with biotechnological and biomedical interest.

2. Experimental

2.1. Reagents.

Choline chloride and DEAE-cellulose were obtained from Sigma-Aldrich. Owing to the hygroscopic properties of choline, concentrated stock solutions were prepared from a freshly opened bottle and stored in aliquots at -20 ºC. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was purchased from Apollo Scientific (UK). Magnetic nanoparticles (MNPs) functionalized with N,N-diethylethanolamine (DEAE)-starch (200 nm) were purchased from Chemicell.

2.2. *Peptide synthesis*.

Peptides P4, P5, P45 and P345 with free amino and carboxylate ends, were synthesised using Fmoc (fluorenyl-9-methyloxycarbonyl) solid-phase protocols and purified by reverse-phase HPLC up to 95% or more purity by Caslo Aps (Lyngby, Denmark). P4 (TGWKKIADKWYYFN): RP–HPLC: t_R = 8.47 min, purity 99.39 % (buffer A: 0.05 % TFA in H_2O/CH_3CN 98:2; buffer B: 0.05 % TFA in H_2O/CH_3CN 1:9; linear 22-36 % B buffer in 14 min). HRMS: Theoretical molecular mass: 1820.10; found:1821.21 $[M + H]^+$; P5 (TGWVKYKDTWYYLD): RP–HPLC: t_R = 10.35 min, purity 95.02 % (buffer A: 0.05 % TFA in H_2O/CH_3CN 98:2; buffer B: 0.05 % TFA in H_2O/CH_3CN 1:9; linear 25-40 % B buffer in 15 min). HRMS: Theoretical molecular mass: 1838.07; found: 1838.70 $[M + H]^+$; P45 (TGWKKIADKWYYFNEEGAMKTGWVKYKDTWYYLD): RP–HPLC: t_R = 10.79 min, purity 95.05 % (buffer A: 0.05 % TFA in H_2O/CH_3CN 98:2; buffer B: 0.05 % TFA in H_2O/CH_3CN 1:9; linear 28-44 % B buffer in 16 min). HRMS: Theoretical molecular mass: 6939.74; found: 6939.93 $[M + H]^+$.; P345

(DRWRKHTDGNWYWFDNSGEMATGWKKIADKWYYFNEEGAMKTGWVKYK DTWYYLD): RP–HPLC: t_R = 11.61 min, purity 95.28 % (buffer A: 0.05 % TFA in H_2O/CH_3CN 98:2; buffer B: 0.05 % TFA in H_2O/CH_3CN 1:9; linear 25-45 % B buffer in 20 min). HRMS: Theoretical molecular mass: 4285.89; found: 4286.56 $[M + H]^+$.

2.3. Data representation and calculations.

Estimations of contacting residues and interaction areas were carried out with UCSF Chimera 1.10 [31]. Figures presenting protein structures were generated using the same tool. Graphs were rendered with SigmaPlot 10.0 (Systat Software, Germany).

2.4. Protein purification

The C-LytA protein was overexpressed in *Escherichia coli* RB791 [pCE17] and purified by affinity chromatography on DEAE-cellulose as previously described [15]. The protein was selectively eluted from the DEAE-cellulose column with 20 mM sodium phosphate buffer pH 7.0 plus 143 mM choline, dialysed at 4 ºC against 20 mM sodium phosphate buffer pH 7.0 to remove the ligand used for elution and stored at -20 ºC.

Peptide and protein concentrations were determined spectrophotometrically (Evolution 201, Thermo Scientific) using an estimated molar extinction coefficient at 280 nm (ε_{280}) of 62 540 M^{-1} cm^{-1} (C-LytA), 13 980 M^{-1} cm^{-1} (P4), 15 470 M^{-1} cm^{-1} (P5), 29 450 M^{-1} cm^{-1} (P45) and 47 440 M^{-1} cm^{-1} (P345) as calculated by the ProtParam tool in the ExPASy server (https://web.expasy.org/protparam/).

2.5. Electronic Circular Dichroism Spectroscopy (ECD).

ECD experiments were carried out in a Jasco J-815 spectropolarimeter (Tokyo, Japan) equipped with a Peltier PTC-423S. Peptide concentration was approximately 30 μM and the cuvette path lengths were 0.1 cm for far-UV region (250-195 nm) and 1.0 cm for near-UV region (320-250 nm). Samples were centrifuged 5 min prior to ECD measuring, although no visible precipitate was detected. All measurements were carried out in triplicate at 5 °C in the presence of 20 mM glycine buffer at pH 3.0 (peptides) or of 20 mM sodium phosphate buffer at pH 7.0 (C-LytA, since this protein precipitates at acidic pH [29]. Isothermal wavelength spectra for these samples were acquired at a scan speed of 50 nm·min^{-1} with a response time of 2 s and averaged over at least six scans. After subtracting the baseline from the sample spectra, ECD data were processed with the adaptive smoothing method provided by the Jasco Spectra Analysis software. Molar ellipticities ([θ]) were expressed in deg cm^2 (dmol of residues)$^{-1}$.

For ECD-monitored thermal denaturation, the sample was layered with mineral oil to avoid evaporation, and the heating rate was set to 50 ºC h^{-1}. The thermal scans were fitted by least squares to eqn. (S1) (see Supplementary Material for a detailed description). The free energy value was estimated with the Gibbs-Helmholtz equation (eqn. (1))

$$\Delta G_{\mathrm{UF}}^{\mathrm{app}}(T) = \Delta H_m \left(1 - {}^{T}/_{T_m}\right) - \Delta C_p \left[(T_m - T) + T ln \left({}^{T}/_{T_m}\right)\right] \qquad (1)$$

where $\Delta G_{\mathrm{UF}}^{\mathrm{app}}$ (T) is apparent free energy of unfolding at a temperature T and a given ligand concentration [L], ΔH_m is the van't Hoff enthalpy, T_m is the midpoint of denaturation (in Kelvin) and ΔC_p is the difference in heat capacity between the native and denatured states.

Due to the flexible nature of the peptides, ΔC_p was assumed to have a negligible value (see text). In the absence of ligand, $\Delta G_{UF}^{app}(T) = \Delta G_{UF}(T)$, *i.e*. the intrinsic free energy of unfolding. Therefore, the additional stabilisation produced upon ligand binding $[\Delta G_{stab}(T, [L])]$ can be calculated as:

$$\Delta G_{stab}(T, [L]) = \Delta G_{UF}^{app}(T, [L]) - \Delta G_{UF}(T) \quad (2)$$

2.6. Fluorescence spectroscopy.

Fluorescence measurements were performed on a PTI-fluorimeter TI-Quanta Maser QM-62003SE (Birmingham, NJ, USA) using a 5×5 mm path-length cuvette and a protein concentration of 1 μM in 20 mM glycine buffer at pH 3.0 and 25 ºC. For choline or NaCl titrations, aliquots from a 2.7 M choline or 2 M NaCl stock solution in glycine buffer were added stepwise and incubated approximately for 5 min prior to measurement. Tryptophan emission spectra were obtained using an excitation wavelength of 280 nm and the emission was registered between 300 and 400 nm, using excitation and emission slits of 1 nm and a scan rate of 60 nm·min^{-1}.

The average emission wavelength ($\langle\lambda\rangle$) of fluorescence spectra was calculated as depicted in eqn. (3):

$$\langle\lambda\rangle = \left.\sum_1^n (I_i\lambda_i)\right/\sum_1^n I_i \quad (3)$$

where I_i is the fluorescence intensity measured at wavelength λ_i .

The choline titration data were fitted to a simple ligand binding model (eqn. (4)), considering independent binding sites of the same type:

$$|\Delta\langle\lambda\rangle| = \frac{|\Delta\langle\lambda\rangle|_{max}\ [L]}{K_d + [L]} \qquad (4)$$

where $|\Delta\langle\lambda\rangle|$ is the absolute change in average emission wavelength at each ligand concentration, $|\Delta\langle\lambda\rangle|_{max}$ is the maximum change in average emission wavelength, $[L]$ is the choline concentration and K_d is the dissociation constant.

2.7 Solution NMR

NMR samples were prepared by dissolving the amount of lyophilised peptides in 0.5 mL of H_2O:D_2O 9:1 v/v or pure D_2O (99.96%) required for a 0.5-1.0 mM concentration. Sodium 2,2 [20] [20]-dimethyl-2-silapentane-5-sulfonate (DSS) was used as internal reference for ^1H chemical shifts. pH was measured employing a glass micro-electrode and adjusted to 3.0 by adding minimal volumes of NaOD or DCl, isotopic effects were not corrected. Samples were placed in 5 mm NMR tubes.

NMR spectra were acquired with a Bruker Avance-600 spectrometer operating at a proton frequency of 600.1 MHz and equipped with a cryoprobe. Cryoprobe temperature was calibrated using a methanol sample. Using previously reported standard techniques [20], 1D ^1H NMR and 2D phase-sensitive two-dimensional correlated spectroscopy (COSY), total correlated spectroscopy (TOCSY), nuclear Overhauser enhancement spectroscopy (NOESY), and ^{13}C natural abundance ^1H-^{13}C heteronuclear single quantum coherence (HSQC) spectra were recorded at 25 ºC (unless otherwise indicated). TOCSY and NOESY mixing times were set to 60 ms and 150 ms, respectively. Data processing was performed with the TOPSPIN program (Bruker Biospin, Karlsruhe, Germany), as

described in previous publications [20]. ^{13}C δ-values were indirectly referenced against the IUPAC-IUB recommended $^{1}H/^{13}C$ chemical shift ratio (0.25144953) [32].

Assignment of ^{1}H NMR signals was performed by analyses of the 2D NMR spectra using the SPARKY software (T. D. Goddard and D.G. Kneller, SPARKY 3, University of California, San Francisco) and following the standard sequential assignment strategy [33,34]. The ^{13}C resonances were straightforwardly identified from the cross-peaks between the protons and the bound carbons observed in the $^{1}H,^{13}C$-HSQC spectra. Tables S1 and S2 list the ^{1}H and ^{13}C chemical shifts.

The $^{1}H_{\alpha}$ chemical shifts were used to estimate structure populations by following previously reported procedures [20,35] assuming a two-state folded/unfolded structural transition, with the folded state being the β-hairpin containing structures. β-hairpin percentages were obtained from the average of the positive ΔδHα values at the strand residues divided by +0.40 ppm [36], which is the mean ΔδHα at protein β-strands, and multiplied by 100. Assuming an experimental error of 0.01 ppm in the measurement of ^{1}H δ-values, the errors in the estimated populations are 3%.

2.8. Structure calculation.

Structures were calculated as previously reported for peptide P4 [20] using the iterative procedure for automatic NOE assignment of the CYANA 2.1 program [37]. The experimental input data for the structure calculations are the lists of: assigned chemical shifts, integrated NOE cross-peaks present in the NOESY spectra, and phi and psi dihedral angles derived from ^{1}H and ^{13}C chemical shifts using TALOS+ webserver [38]. The automatic integration subroutine of SPARKY software (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco) was used to get NOE cross-

peak integrals. The final ensembles of the 20 lowest target function conformers were visualized and examined using the MOLMOL program [39].

2.9. Binding of CBR peptides to DEAE-coated magnetic nanoparticles.

FluidMAG magnetic nanoparticles (1 mg) functionalized with 2-diethylaminoethyl-starch groups (DEAE-MNPs, 200 nm diameter, Chemicell, Germany) were washed and equilibrated in 20 mM glycine buffer, pH 3.0. The particles were then resuspended in 350 μL glycine buffer (approx. 1.9×10^{11} particles mL^{-1}) containing the peptides as described in the text. The incubation was maintained for 30 min at room temperature in a rotary mixer (Ovan). At this point, bound protein was separated from the solution using a magnet (MagnetoPURE). The supernatant was subsequently centrifuged 3 min at 5000 × *g* in a miniSpin centrifuge (Eppendorf) to ensure the removal of residual MNPs, and the amount of unbound peptide was measured spectrophotometrically.

3. Results

3.1. Dissection of the ββ-solenoid structure of the LytA choline-binding module.

A scheme of the internal contacts between the CBRs of CBM-LytA is depicted in Fig. 1B. The border repeats CBR1 and CBR6 display additional stabilising interactions outside the ββ-solenoid with the N-terminal functional module and the C-terminal dimerization hairpins of the full parental protein, respectively [2,9]. While these repeats fulfil an essential role in determining the limits of the solenoid (as will be discussed below) we will focus on the central core of the structure, which is structurally similar among most CBPs [7]. Table 1 breaks down the inter-CBR contacts into hairpin- (*h*) or linker- (*k*) mediated interactions. A close inspection of these data reflects, in the first instance, a symmetrical distribution (~50%) in the total number of interactions between any given CBR and the previous or the following CBR (around 18 contacts), irrespective of their nature. However, there are variations on the relative contribution of hairpins or linkers as some packings heavily rely on hairpin-hairpin contacts (e.g. CBR4-CBR5 or CBR5-CBR6, with 8 of such contacts), but others are more based in linker-hairpin and linker-linker components (e.g. CBR1-CBR2 or CBR2-CBR3) (Fig. 1B and Table 1). In fact, there is an inverse linear correlation between the number of interactions involving the hairpins and those involving the linkers for each CBR in the central region of the solenoid (r^2 = 0.970, Fig. S1A). On the other hand, the analysis of the interaction area between consecutive CBRs reveal remarkable traits. The contact surface of any two consecutive CBRs (Table 1) shows no apparent correlation with the number of interactions established between the hairpin of the first repeat and the linker of the second one (Fig. S1B), or only a slight, negative correspondence with the number of contacts involving their respective hairpins (Fig. S1C). In contrast, there are strong positive correlations between this interaction area and the contacts established by the linker

sequence of the first CBR of the pair, be it with the hairpin (r^2 = 0.970, Fig. S1D) or with the linker (r^2 = 0.986, Fig. S1E) of the second repeat. This points to the linker sequences as major determinants of the stability of any CBR-CBR packing in the central core of CBM-LytA.

3.2. Rationale of peptide synthesis

To gain a deeper insight into the structural determinants of the CBM-LytA conformation, we designed a bottom-up experimental study of peptides encompassing the central repeats of LytA, in particular CBRs 3, 4 and 5 (Fig. 1C). Several reasons led us to choose this particular set: *i*) the selected CBRs are not involved in extra-CBR interactions; *ii*) the structure and stability of the hairpin belonging to CBR4 (contained in the P4 peptide, see Fig. 1C) has already been thoroughly studied [20,21]; *iii*) the set contains the CBR3-CBR4 packing, with a majority of contacts involving a linker sequence, together with the CBR4-CBR5 pair, mainly stabilized by hairpin-based interactions (Table 1); and finally, *iv*), three consecutive CBRs is the minimal configuration necessary to complete a single turn of the ββ-solenoid CBM staircase [2,9]. A scheme of the synthetized peptides is depicted in Fig. 1C. Peptides P4 (residues 239-252 of full-length LytA) and P5 (residues 259-272) represent the β-hairpins from the CBR4 and CBR5 repeats, respectively, while P45 (residues 239-272) contains both hairpins and their linking sequence, and thus represents the minimal structure theoretically able to shape a standard, full choline-binding site (Fig. 1A). Finally, peptide P345 (residues 218-272) represents one turn of the solenoid superhelical structure and contains three hairpins (belonging to CBRs 3, 4 and 5) and their two internal linker sequences (Fig. 1C). This peptide supplies the linker sequence k_3, which might provide a number of stabilising interactions for the packing between the following CBRs 4 and 5 (Fig. 1B, Table 1).

3.3. Structure of CBR peptides in aqueous solution

The structures of the P4, P5, P45 and P345 peptides were firstly assessed by far-UV electronic circular dichroism (ECD) at 5 ºC in 20 mM glycine-HCl buffer (Gly buffer) at pH 3.0 (Fig. 2A). All spectra display a prominent, positive band in the 220 - 240 nm region, with maxima centred at 227-228 nm that is likely to arise from the presence of rigid structures around the tryptophan side chains leading to exciton coupling of π-π* electronic transitions in the ^1B_b absorption band [40]. Aromatic contributions are also observed in the near-UV ECD spectra (except P345, which was not soluble at the required concentrations), that display a prominent positive band around 261-263 nm and two minima at 286 and 294 nm (Fig. 2B). All these CD features have been previously described for the isolated CBR2 and CBR4 hairpins [19,20] as well as for the full-length CBMs and CBPs [10,15,24], suggesting that all peptides spontaneously acquire a native-like secondary structure in solution. It is also noteworthy that the intensity of the P345 far-UV ellipticity is appreciably lower than that of the rest of peptides, resembling more the spectrum of the parental module C-LytA (Fig. 2A).

Following up our previous studies on the CBR2 [19] and CBR4 hairpins [20] (the latter corresponding to the P4 peptide in this work), we carried out a detailed NMR study of the P5 and P45 peptides in aqueous solution and in the absence of choline. The study of the P345 peptide was hampered by its insolubility at the concentrations required for the NMR experiments. A standard strategy was followed to assign the ^{1}H and ^{13}C chemical shifts of the peptides in aqueous solution (Tables S1 and S2). Then, their structural behaviour was deduced from the NMR parameters, *i.e*. the conformational shifts of $^1H_\alpha$, $^{13}C_\alpha$ and $^{13}C_\beta$ atoms ($\Delta\delta = \delta^{observed} - \delta^{RC}$, ppm; Fig. 3), and the sets of

observed NOE cross-peaks (Fig. S2). To obtain further details of the spatial arrangement of these peptides, structure calculations were performed on the basis of the distance restraints derived from the complete sets of observed NOEs and the dihedral angle restraints obtained from the $^{1}H_{\alpha}$, $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts by using the program TALOS+ (Table 2).

The calculated structures for the P5 hairpin are well defined (Fig. 4A), as shown by the low pairwise RMSDs of the backbone atoms (Table 2), and very similar to its native structures in the CBM-LytA module (PDB code: 4X36) (Fig. 5), as deduced from the average of the backbone RMSD values calculated by fitting each conformer to the native structure (0.799 ± 0.003 Å). The side chains are also well ordered in nearly all the residues (except for the N– and C–terminal ones) since their χ_1 dihedral angles show very little variation among the 20 calculated structures (χ_1 variation range < 4º for most residues), and display orientations quite similar in general to those in the native protein (RMSD including side chains is 1.73 ± 0.04 Å; Fig. 5). On the other hand, in the long peptide P45, the characteristic patterns of β-hairpins in the profiles of $\Delta\delta_{H\alpha}$, $\Delta\delta_{C\alpha}$ and $\Delta\delta_{C\beta}$ were observed in two segments, residues 239-252 and 259-272, whereas the central linker segment displayed values within the random coil range (Fig. 3). Moreover, non–sequential NOE cross–peaks typical of β-hairpin structures were only observed for the segments 239-252 and 259-272 (see the characteristic Hα-Hα NOEs in Fig. S2). The absence of non-sequential NOEs for residues 253-258 confirms a mostly unstructured linker except for the presence of significant local contacts of Glu254 and Gly255 with Asn252 (which belongs to hairpin 4). In addition, numerous non-sequential NOEs involving side chain protons in the short peptide P4 (Table S3), and in regions 239-252 and 259-527 of the long P45 sequence are in accordance with the formation of native-like β-hairpins. This evidences that the P45 peptide forms two native-like β-hairpins along its

chain and that the linker shows no regular structure. In fact, the ensemble of the 20 lowest target function conformers is in general poorly defined (see the very large RMSD values in Table 2), although the regions corresponding to the two β-hairpin cores are well defined (Fig. 4), as indicated by the small pairwise RMSDs presented by the backbone atoms (Table 2). They are also very similar to the native structures, as shown by the small backbone RMSD values (1.33 ± 0.01 and 1.18 ± 0.08 Å for 239–252 and 259–272 regions, respectively) and depicted in Fig. 5. Hence, the poor definition of the complete P45 structure can be explained by the relative position of the two β–hairpins not being fixed (Fig. 4E-F), as there are no conformational restrictions around the linker segment.

Based on the averaged $\Delta\delta_{H\alpha}$ values at the strand residues (see Experimental section), the population of natively folded hairpins 4 and 5 in the P45 peptide (measured at 35 ºC to avoid an otherwise unacceptable signal broadening) is approximately 69 and 80 %, respectively (averaged $\Delta\delta_{H\alpha}$ values are +0.27 and +0.33, respectively). These values are comparable to that of the P5 peptide at 25 ºC, which is estimated to be approximately 73 % ($\Delta\delta_{H\alpha}$ averaged for strand residues is +0.29 ppm). Regarding the P4 peptide, a value of *c.a.* 100% was previously reported by NMR at 25 ºC, although ECD-monitored thermal denaturation revealed a loss of around 20% of ellipticity signal at 35 ºC (the temperature of the P45 NMR experiments) [20]. These data altogether therefore confirm that the structural propensity of the β–hairpins 4 and 5 is similar either isolated or when being part of the same peptide sequence in P45, although they cannot establish a well-packed native-like interaction between them.

A detailed comparison of the side-chain conformations of the aromatic residues in the P4, P5 and P45 peptides with the native, choline-ligated LytA structure [9] reveal interesting features (Fig. 6 and Table S4). In the short peptides, the aromatic residues that configure the choline-binding half-sites (Trp241-Trp248 in hairpin-4, Trp261-Trp268 in

hairpin 5) present an appreciably more open conformation than in the full-length protein, where they are disposed in an almost perpendicular fashion (105º and 82º, respectively). However, when the two hairpins are connected in the P45 peptide, the hairpin 4 subsite approaches that of the native protein, while the hairpin 5 subsite deviates even more (Fig. 6 and Table S4). This trend is also observed with the additional aromatic residue (Tyr250) that establishes a π-π stacking interaction in hairpin 4, as it acquires an intermediate conformation between that of the short P4 peptide and the LytA full structure (Fig. 6). In the absence of packing interactions between the hairpins (Fig. 4E-F), this suggests that the addition of the linker sequence to the hairpin-4 restricts the conformational flexibility of the aromatic side chains of the latter, that configure the choline-binding half-site (maybe through local contacts with the neighbour linker residues Glu254 and Gly255, as mentioned above), forcing them to acquire a native-like conformation more able to recognise the ligand.

3.4. Peptide stability

The thermodynamic stability of the CBR peptides was examined by monitoring the far-UV ECD signal as a function of temperature (Fig. 7). Heating the samples up to 95 ºC led to featureless spectra devoid of the aromatic positive contributions at ~ 227 nm, indicating an extensively unfolded conformation (Fig. S3) [41]. Since all changes were virtually reversible upon cooling (Fig. S3), we analysed the unfolding data by equilibrium thermodynamics (eqn. (S1)), making use of the Gibbs-Helmholtz expression (eqn. (1)). All ECD-monitored thermal transitions displayed a sigmoidal, extended shape without hints of multiple steps (Fig, 7) that reflects a relatively small degree of cooperativity upon unfolding compatible with the lack of a sizeable hydrophobic core. For this reason, we assumed the ΔC_p value to be negligible, in line with the previous analysis of the LytA

CBR4 hairpin [20] and other similar β-hairpin peptides [42–44] including the tryptophan-stabilised *trpzip* peptide [45]. Results of fittings are shown in Table 3. In the absence of choline, both the change in enthalpy, entropy and free energy values for the four peptides are similar and of small magnitude, indicating a marginal stability at 25 ºC. Of note, we observed a tendency of the melting temperatures (T_m) and apparent stabilisation energies (ΔG_{UF}^{app}) to decrease with the length of the peptide (Table 3). In fact, the T_m's for P45 and P345 are even lower than those of the P4 and P5 peptides separately. Since there are no hints of packing interactions in the folded states of at least P45 (Fig. 4), this decrease in stability might be explained in terms of residual, non-specific interactions established among unfolded molecules of the long peptides that would shift the thermodynamic equilibrium towards the unfolded state.

3.5. Functionality of peptides: binding of choline

Although the configuration of standard choline-binding sites in CBPs require the concerted participation of two consecutive β-hairpins (Fig. 1A), the P45 peptide does not have sufficient information to ensure a native-like supersecondary fold in the absence of the ligand, but instead displays an extensive conformational freedom that prevents the hairpins to be fixed in close proximity (Fig. 4). Remarkably, however, as described above, the side chains of the residues that are part of the half-site provided by hairpin-4 (Trp241 and Trp248) acquire in P45 a near-native conformation that might provide a better capacity for ligand recognition compared to the short P4 or P5 peptides (Fig. 6). In this context, we aimed to investigate whether choline binding to these "improved" half-site might be enhanced and even reduce peptide flexibility by promoting the approach of the β-hairpins in the longer P45 and P345 peptides. We first attempted to elucidate the structure of the peptides by NMR in the presence of choline or choline analogues such as

tetramethylammonium chloride. However, since the affinity of all these peptides for the ligands is extremely low (see below), requiring the addition of ligand to concentrations >100 mM to saturate the binding sites. Despite the use of highly deuterated compounds (> 98%), the residual ^{1}H ligand signals were strong enough to conceal the peptide signals, making any analysis by this technique unfeasible. Therefore, we carried out the binding analysis by ECD and fluorescence spectroscopies.

Figure 8 (upper panels) displays the far- and near-UV ECD spectra of the P4, P5, P45 and P345 peptides at 5 ºC in the presence of a high concentration of choline (500 mM). Again, the near-UV spectrum of P345 could not be acquired because of the insolubility of the sample. As an overall feature, the ligand induces an increase in the intensities of the ~ 227-228 nm positive bands and in the minima at ~ 286 and ~ 294 nm. These characteristics are shared with the full-length C-LytA module (Fig. 8) as well as other CBPs and CBMs [15,24], and have been explained as a consequence of the increase in rigidity in the environment of the aromatic side chains. A detailed inspection of the changes in the far-UV region is depicted in Fig. 8 (lower panels). Two behaviours can be discerned upon binding of choline: on one side, the spectra of the P4 and P5 peptides, each providing one imperfect binding half-site, are affected to a minor extent (Fig. 8; see insets for difference spectra), suggesting minimal, if any, conformational changes in the hairpins upon ligand binding. The small increase in ellipticity for P5, without any concomitant bandshift, might be explained at least in part in terms of ligand-induced refolding of the denatured population of the peptide (around 27% as estimated by NMR, see above), since the difference spectrum approximately reproduces, at a small scale, the same shape of the native spectrum. On the other hand, changes in the spectra of P45 and P345 are more complex, as they involve both an increase in ellipticity and a blue-shift of the positive band. The difference spectra show the appearance of a new positive

contribution centred at 222 nm, which was also observed for the parental module C-LytA at pH 7.0 (Fig. 8) and that suggests the establishment of a new interaction involving aromatic residues [40,46].

The addition of 500 mM choline also induces a general stabilisation of the peptides. Analyses of reversible thermal unfolding transitions reveal an increase in T_m and in the apparent free energy ΔG_{UF}^{app} (Fig. 7 and Table 3). However, while the stabilisation of P4 and P5 by the ligand (ΔG_{stab}) is only modest, P45 and, especially, P345, undergo a more evident cooperative transition (Fig. 7) which, in the case of the longer peptide, occurs with a remarkable increase in unfolding enthalpy, and, in consequence, in ΔG_{UF}^{app} and ΔG_{stab} (Table 3). These parametres suggest that packing interactions induced by choline are more complete for the three-hairpin P345 peptide, which encompasses a full turn of the ββ-solenoid structure, and that this is the minimal core capable of adopting a native-like tertiary folding.

Due to the involvement of aromatic residues in the binding of choline [2,9], the affinity of the peptides for the ligand was easily monitored by intrinsic fluorescence spectroscopy (Figs. 9 and S4). This technique was used instead of ECD since the changes in the the far-UV spectra of the isolated hairpins (P4 and P5 peptides) (Fig. 8) were of a too small magnitude to render suitable binding curves, whereas near-UV ECD makes necessary the use of 10-fold higher concentrations of peptide, which might favour possible intermolecular peptide-peptide interactions that could distort the affinity data, especially for the P345 peptide (which is prone to aggregation). All the peptides display a broad fluorescence emission band centred at around 340 nm (Fig. S5), which is compatible with highly solvent exposed tryptophan residues (Fig. 4). Addition of choline chloride induces, in all cases, a blue-shift on the spectrum to around 334 nm concomitantly with an increase in fluorescence intensity (Fig. S5) that might reflect the

partial burial of aromatic side chains upon ligand binding, causing a decrease in the polarity of their surroundings as well as providing some protection from solvent quenching. These spectral changes cannot be ascribed to non-specific ionic strength effects, since addition of similar concentrations of NaCl caused instead a decrease in intensity and without affecting the maximum emission wavelength (Fig. S5). Besides, the salt-induced spectral changes were independent of salt concentration above 75 mM, much below the choline titration endpoints (see Fig. S5, lower panel, for the P4 case as an example). These reductions of fluorescence intensity might be explained *e.g.* by the screening of the positive charges of Lys222, Lys243 and Lys263 sidechains, which establish hairpin-stabilizing cross-strand cation-π interactions with Trp228, Trp248 and Trp268, respectively [2,9,21], so that the peptide would be somewhat destabilized in this region by the salt, promoting the access and subsequent quenching of tryptophan fluorescence by solvent molecules. In any case, these observations suggest that any shifts in emission maxima are specific reporters of choline binding, and accordingly, we monitored the changes in average emission wavelength to assess the affinity of the peptides for the ligand (Fig. 9). In all cases a hyperbolic-type titration plot was obtained, and data were readily fitted to a simple binding equation assuming one type of independent sites with a macroscopic dissociation constant (K_d) (eqn.(3)). Attempts to unveil two or more types of sites were unsuccessful (data not shown), although this possibility cannot be ruled out, in particular if the sites have affinities that differ in less than one order of magnitude. The resulting fits are reported in Fig. 9 and Table 3, and show in all cases only marginal affinities, with macroscopic dissociation constants on the scale of hundreds of millimolar. However, a clear trend is observed since the affinity increases as more hairpins are added to the sequence, so that the P345 peptide has a 2.5 to 3-fold higher affinity (K_d = 117 mM) than the P5 (K_d = 290 mM) or P4 (K_d = 360 mM)

peptides, with P45 in an intermediate level (K_d = 170 mM). This supports the hypothesis of a conformational change affecting the relative disposition of the hairpins in the P45 and, especially, the P345 peptide, leading to some hairpin-hairpin packing, as mentioned above, and the conformation of native-like choline binding sites.

To this point, it should be recalled that the actual biological ligand of CBRs is the polydentate structure of the cell wall, where multiple choline residues decorate the teichoic and lipoteichoic acids attached to the peptidoglycan and cell membrane, respectively [47]. As a consequence of chemical multivalency, it has been recently shown that the affinity of choline-binding modules, such as C-LytA, for diethylaminoethyl (DEAE)- functionalized magnetic nanoparticles (DEAE-MNPs), a polydentate cell wall mimic, increases up to 100 times compared to monomeric, soluble choline [10] Therefore, to assess the interaction of the CBR peptides on a macromolecular substrate, these were incubated with DEAE-MNPs and their adsorption on the particles was monitored by measuring the remaining peptide in the supernatant by UV absorption spectroscopy. To facilitate the comparison between peptides, their concentration was chosen so that the number of hairpin equivalents in the assay was the same (12.9 μM). Therefore, the concentration of P45 and P345 peptides was set to 1/2 (6.45 μM) and 1/3 (4.3 μM) of the concentration of P4 and P5 (12.9 μM), respectively. While the single hairpins (P4 and P5 peptides) did not present a detectable affinity for the polydentate substrate, both P45 and P345 were totally adsorbed (data not shown). This confirms that the presence of at least a second hairpin in the same peptide chain appreciably increases the binding ability of CBR peptides, and that this leap in avidity becomes more evident with cell-wall macromolecular mimics.

4. Discussion

Secondary structure propensities often dictate the global stability of proteins, and in the case of the repeating, largely monotonic ββ-solenoid structure, without long-distance interactions characteristic of the typical globular folds, β-hairpin stability might be predicted to be a crucial factor. In this sense, the β-hairpins belonging to CBRs 4 and 5 have proved to be very robust structures, as they acquire a native-like conformation in aqueous solution at 25 ºC without the assistance of any cosolvents or ligands, as indicated by ECD (Figs. 2, 4 and S4) and NMR (Figs. 3-5 and Tables 2, S1 and S2), and they also show an appreciable thermodynamic stability (Fig. 7 and Table 3). Moreover, they are able to specifically bind choline, albeit with marginal affinity (Fig. 9 and Table 3). This complements similar, previous observations on the CBR2 hairpin [19]. On the other hand, although the P45 peptide represents the minimal structure theoretically capable of acquiring a native-like tertiary fold with a standard choline-binding site (Fig. 1A), its elucidated 3D structure (Fig. 4) and the thermodynamic stability data (Table 3) indicate an extensive flexibility in the unligated structure without any stable interaction between the hairpins. Nevertheless, compared with the P4 peptide, the presence of the *k*4 linker sequence (Fig. 1C) in the P45 peptide induces a detectable change in the geometry of the hairpin-4 half-site, which acquires a more native-like configuration (Fig. 6 and Table S4). This is reflected in an significant increase in affinity for choline (Fig. 9) and in unfolding enthalpy (Table 3) upon ligand binding, that suggest that a small population of this peptide is able to adopt a full standard choline-binding site with the participation of aromatic side chains belonging to the two hairpins and leading to an incipient hydrophobic core. Furthermore, the addition of a third CBR in the P345 peptide (CBR3), containing the linker sequence k_3 that might contribute to P45 stability with up to 7 new interactions (Table 1, Fig. 1B), remarkably augments both binding affinity and stability.

While, in the absence of choline, the P345 peptide reproduces the same independent unfolding behaviour of the hairpins observed in P45, with a similar marginal stability (Table 3), ligand binding results in an appreciable increase in cooperativity on P345 thermal denaturation (Fig. 7) as well as in its unfolding enthalpy, entropy and overall free unfolding energy (ΔG_{UF}^{app}) (Table 3). Moreover, this is accompanied by a notable increase (around 3-fold) in affinity for choline (Fig. 9 and Table 3). In this sense, the far-UV ECD results (Fig. 8) show that, on ligand addition, the P45 and P345 ECD spectra not only enhance their intensities but are also blue-shifted as a consequence of the appearance of new positive contributions to the spectrum (Fig. 8, insets) that might arise from the coupling of the two aromatic residues from the subsite of one hairpin with the third aromatic residue supplied by the next. Therefore, these results suggest that the presence of three hairpins is the minimum requirement for the ββ-solenoid to constitute the nucleus of a compact, tertiary fold, in the presence of the ligand and in agreement with the structuration of this superhelix in three steps per turn [2,9]. This is also in accordance with the observation that no CBP has been described so far with less than three CBRs in their sequence [7]. Moreover, our theoretical analysis of CBR-CBR interactions in CBM-LytA show that linkers also determine the interaction area between consecutive CBRs, leading to inhomogeneities (Table 1) that might be at the basis of the modulation of substrate recognition, for instance the origin of high- and low-affinity choline-binding sites in LytA [18].

The low affinity for choline displayed by the P45 and P345 peptides posed technical drawbacks that precluded the elucidation of their 3D structure by NMR. Nevertheless, contrary to peptides P4 and P5 alone, both P45 and P345 show evident binding to cell-wall-mimic DEAE-MNPs similar to full CBMs ([10], see above), which

reinforces the hypothesis of the configuration of choline-binding sites with native-like characteristics.

The ββ-solenoid fold of CBMs, of which CBM-LytA is an example, is not a particularly stable structure in the absence of a ligand. The relatively low thermal stability of CBMs such as those belonging to LytA, LytB, CbpD and Pce, with T_m's similar to that of the P345 peptide (around 40 ºC, Table 4), reveals that these important modules are on the edge of unfolding at physiological temperatures (or even completely unfolded in the case of LytC). This is a striking but common trait for these proteins despite the differences in their sequences, and for this reason it might play a biological role, for instance by facilitating interactions with the Sec membrane translocation machinery [48] and/or by providing flexibility to CBMs to establish a dynamic on-off equilibrium with the cell wall until they reach their final topological location in the peptidoglycan [7]. Furthermore, a loose structure also facilitates high-order conformational changes such as the unusual β-sheet → α-helix transition undergone by the C-LytA module as induced by detergent micelles and phospholipid vesicles, a process that might facilitate translocation across the cell membrane in the absence of a signal peptide [20]. These structural peculiarities arise from the fact that, with the normal exceptions of the N- and C-terminal flanking CBR sequences, there are no longer-range interactions for any CBR beyond the previous and the following repeats (Fig. 1).

The major differences between P345 and the full-length CBMs are therefore found in terms of choline binding. Table 4 shows that the ligand induces a much higher thermal stabilisation on the full-length CBMs than on the peptide, concomitant with a higher binding affinity (between 20 and 1000-fold). To this point, it should be recalled that, unlike P345, the ββ-solenoid structures of full-length CBMs are "capped" by external agents such as the cognate functional modules within the CBP, as well as with other short

hairpin sequences that deviate from the standard CBR configuration [7]. These short "caps" or "staples" represent termination signals of the CBMs that, in the presence of choline, confer much stability to the proteins by reducing the intrinsic conformational freedom of the repeats, either by constituting dimerization motifs (*e.g.* LytA), or containing a closing, non-standard ligand binding site (e.g. CbpF [49], CbpJ [50], CbpM [51] or Pce[52]). The ubiquitous presence of these capping sequences highlights their roles as essential regulators of CBP stability and activity. This is illustrated by the case of LytA, where removal of the C-terminal dimerization hairpin decreases by 5-fold the affinity for choline (K_d = 25 mM), and by 90% the hydrolytic activity of this amidase [11].

5. Conclusions

Three CBRs are the minimal requirement for a choline-binding module to acquire the ββ-solenoid conformation of pneumococcal LytA autolysin with native-like three-dimensional characteristics. While the constituting β-hairpins spontaneously acquire a folded conformation in solution, linker sequences are pivotal in determining the superhelical fold. Future experimental work on different ββ-solenoids needs to take into account the connecting stretches between hairpins, as subtle variations in these sequences, rather than in the hairpins themselves (which are consistently robust in structure) might lead to distinct biophysical properties and conformations of binding sites. For the pneumococcal CBM system, this consideration would also be beneficial for the rational design of novel variants that may find their application as improved specific pneumococcal recognition devices [10] or in biotechnological applications as tunable affinity tags for the immobilization and purification of recombinant proteins.

Acknowledgments

We are indebted to Dr. Douglas V. Laurents for his extensive revision of the manuscript and his very useful comments. Grants: BIO2016-79323-R, CTQ2017-84371P and PID2019-105126RB-I00 (Agencia Estatal de Investigacion AEI/FEDER-EU-10.13039/501100011033-, and Ministerio de Ciencia e Innovación, Spain); Intramural incorporation project #201820I132 (CSIC, Spain). The CIBER de Enfermedades Respiratorias (CIBERES) is an initiative of the Spanish Instituto de Salud Carlos III. The NMR experiments were performed in the "Manuel Rico" NMR laboratory, LMR, CSIC, a node of the Spanish Large-Scale National Facility ICTS R-LRB.

References

[1] A. V. Kajava, A.C. Steven, β-Rolls, β-helices, and other β-solenoid proteins, Adv. Protein Chem. 73 (2006) 55–96. https://doi.org/10.1016/S0065-3233(06)73003-0.

[2] C. Fernández-Tornero, R. López, E. García, G. Giménez-Gallego, A. Romero, A novel solenoid fold in the cell wall anchoring domain of the pneumococcal virulence factor LytA, Nat. Struct. Biol. 8 (2001) 1020–1024. https://doi.org/10.1038/nsb724.

[3] M. Claverie, G. Cioci, M. Vuillemin, N. Monties, P. Roblin, G. Lippens, M. Remaud-Simeon, C. Moulis, Investigations on the determinants responsible for low molar mass dextran formation by DSR-M dextransucrase, ACS Catal. 7 (2017) 7106–7119. https://doi.org/10.1021/acscatal.7b02182.

[4] M. Molina, C. Moulis, N. Monties, S. Pizzut-Serin, D. Guieysse, S. Morel, G. Cioci, M. Remaud-Siméon, Deciphering an undecided enzyme: investigations of the structural determinants involved in the linkage specificity of alternansucrase,

ACS Catal. 9 (2019) 2222–2237. https://doi.org/10.1021/acscatal.8b04510.

[5] T. Murase, L. Eugenio, M. Schorr, G. Hussack, J. Tanha, E.N. Kitova, J.S. Klassen, K.K.S. Ng, Structural basis for antibody recognition in the receptor binding domains of toxins a and B from *Clostridium difficile*, J. Biol. Chem. 289 (2014) 2331–2343. https://doi.org/10.1074/jbc.M113.505917.

[6] P. Chen, K. ho Lam, Z. Liu, F.A. Mindlin, B. Chen, C.B. Gutierrez, L. Huang, Y. Zhang, T. Hamza, H. Feng, T. Matsui, M.E. Bowen, K. Perry, R. Jin, Structure of the full-length *Clostridium difficile* toxin B, Nat. Struct. Mol. Biol. 26 (2019) 712–719. https://doi.org/10.1038/s41594-019-0268-0.

[7] B. Maestro, J.M. Sanz, Choline binding proteins from *Streptococcus pneumoniae*: A dual role as enzybiotics and targets for the design of new antimicrobials, Antibiotics. 5 (2016) 21. https://doi.org/10.3390/antibiotics5020021.

[8] C. Fernández-Tornero, E. García, R. López, G. Giménez-Gallego, A. Romero, Two new crystal forms of the choline-binding domain of the major pneumococcal autolysin: Insights into the dynamics of the active homodimer, J. Mol. Biol. 321 (2002) 163–173. https://doi.org/10.1016/S0022-2836(02)00596-X.

[9] Q. Li, W. Cheng, C. Morlot, X.H. Bai, Y.L. Jiang, W. Wang, D.I. Roper, T. Vernet, Y.H. Dong, Y. Chen, C.Z. Zhou, Full-length structure of the major autolysin LytA, Acta Crystallogr. Sect. D Biol. Crystallogr. 71 (2015) 1373–1381. https://doi.org/10.1107/S1399004715007403.

[10] E. Roig-Molina, M. Sánchez-Angulo, J. Seele, F. García-Asencio, R. Nau, J.M. Sanz, B. Maestro, Searching for antipneumococcal targets: choline-binding modules as phagocytosis enhancers, ACS Infect. Dis. 6 (2020) 954–974. https://doi.org/10.1021/acsinfecdis.9b00344.

[11] J. Varea, J.L. Saiz, C. López-Zumel, B. Monterroso, F.J. Medrano, J.L.R. Arrondo, I. Iloro, J. Laynez, J.L. García, M. Menéndez, Do sequence repeats play an equivalent role in the choline-binding module of pneumococcal LytA amidase?, J. Biol. Chem. 275 (2000) 26842–26855. https://doi.org/10.1074/jbc.M004379200.

[12] P. Usobiaga, F.J. Medrano, M. Gasset, J.L. García, J.L. Saiz, G. Rivas, J. Laynez, M. Menéndez, Structural organization of the major autolysin from *Streptococcus pneumoniae*, J. Biol. Chem. 271 (1996) 6832–6838. https://doi.org/10.1074/jbc.271.12.6832.

[13] B. Maestro, J.M. Sanz, Extensive unfolding of the C-LytA choline-binding module by submicellar concentrations of sodium dodecyl sulphate, FEBS Lett. 581 (2007) 375–381. https://doi.org/10.1016/j.febslet.2006.12.042.

[14] J.M. Sanz, R. Lopez, J.L. Garcia, Structural requirements of choline derivatives for “conversion” of pneumococcal amidase A new single-step procedure for purification of this autolysin, FEBS Lett. 232 (1988) 308–312. https://doi.org/10.1016/0014-5793(88)80759-2.

[15] J.M. Sánchez-Puelles, J.M. Sanz, J.L. García, E. García, Cloning and expression of gene fragments encoding the choline-binding domain of pneumococcal murein hydrolases, Gene. 89 (1990) 69–75. https://doi.org/10.1016/0378-1119(90)90207-8.

[16] B. Maestro, J.M. Sanz, Accumulation of partly folded in the equilibrium unfolding of the pneumococcal choline-binding module C-LytA, Biochem. J. 387 (2005) 479–488. https://doi.org/10.1042/BJ20041194.

[17] B. Maestro, A. González, P. García, J.M. Sanz, Inhibition of pneumococcal choline-binding proteins and cell growth by esters of bicyclic amines, FEBS J. 274

(2007) 364–376. https://doi.org/10.1111/j.1742-4658.2006.05584.x.

[18] F.J. Medrano, M. Gasset, C. López-Zúmel, P. Usobiaga, J.L. García, M. Menéndez, Structural characterization of the unligated and choline-bound forms of the major pneumococcal autolysin LytA amidase: Conformational transitions induced by temperature, J. Biol. Chem. 271 (1996) 29152–29161. https://doi.org/10.1074/jbc.271.46.29152.

[19] B. Maestro, C.M. Santiveri, M.A. Jiménez, J.M. Sanz, Structural autonomy of a β-hairpin peptide derived from the pneumococcal choline-binding protein LytA, Protein Eng. Des. Sel. 24 (2011) 113–122. https://doi.org/10.1093/protein/gzq087.

[20] H. Zamora-Carreras, B. Maestro, E. Strandberg, A.S. Ulrich, J.M. Sanz, M.Á. Jiménez, Micelle-triggered β-hairpin to α-helix transition in a 14-residue peptide from a choline-binding repeat of the pneumococcal autolysin LytA, Chem. - A Eur. J. 21 (2015) 8076–8089. https://doi.org/10.1002/chem.201500447.

[21] H. Zamora-Carreras, B. Maestro, E. Strandberg, A.S. Ulrich, J.M. Sanz, M.Á. Jiménez, Roles of amphipathicity and hydrophobicity in the micelle-driven structural switch of a 14-mer peptide core from a choline-binding repeat, Chem. - A Eur. J. 24 (2018) 5825–5839. https://doi.org/10.1002/chem.201704802.

[22] L. Lagartera, Estructura y función de la fosforilcolín esterasa de *Streptococcus pneumoniae*, Doctoral thesis. Universidad Complutense de Madrid, Spain, 2006.

[23] P. Rico-Lastres, Caracterización estructural y funcional de la glucosaminidasa LytB de *Streptococcus pneumoniae*, Doctoral Thesis. Universidad Complutense de Madrid, Spain, 2015. https://dialnet.unirioja.es/servlet/tesis?codigo=129144 (accessed March 7, 2021).

[24] B.A. Monterroso, J.L. Sáiz, P. García, J.L. García, M. Menéndez, Insights into the

structure-function relationships of pneumococcal cell wall lysozymes, LytC and Cpl-1, J. Biol. Chem. 283 (2008) 28618–28628. https://doi.org/10.1074/jbc.M802808200.

[25] J. Caubín, H. Martín, A. Roa, I. Cosano, M. Pozuelo, J.M. De La Fuente, J.M. Sánchez-Puelles, M. Molina, C. Nombela, Choline-binding domain as a novel affinity tag for purification of fusion proteins produced in *Pichia pastoris*, Biotechnol. Bioeng. 74 (2001) 164–171. https://doi.org/10.1002/bit.1106.

[26] J.M. Sánchez-Puelles, J.M. Sanz, J.L. García, E. García, Immobilization and single-step purification of fusion proteins using DEAE-cellulose, Eur. J. Biochem. 203 (1992) 153–159. https://doi.org/10.1111/j.1432-1033.1992.tb19840.x.

[27] S. Ortega, J.L. García, M. Zazo, J. Varela, I. Muñoz-Willery, P. Cuevas, G. Giménez-Gallego, Single-step purification on DEAE-sephacel of recombinant polypeptides produced in *Escherichia coli*, Bio/Technology. 10 (1992) 795–798. https://doi.org/10.1038/nbt0792-795.

[28] I. Campos-Pinto, E. V. Capela, A.R. Silva-Santos, M.A. Rodríguez, P.R. Gavara, M. Fernandez-Lahore, M.R. Aires-Barros, A.M. Azevedo, LYTAG-driven purification strategies for monoclonal antibodies using quaternary amine ligands as affinity matrices, J. Chem. Technol. Biotechnol. 93 (2018) 1966–1974. https://doi.org/10.1002/jctb.5460.

[29] B. Maestro, I. Velasco, I. Castillejo, M. Arévalo-Rodríguez, Á. Cebolla, J.M. Sanz, Affinity partitioning of proteins tagged with choline-binding modules in aqueous two-phase systems, J. Chromatogr. A. 1208 (2008) 189–196. https://doi.org/10.1016/j.chroma.2008.08.106.

[30] V.M. Hernández-Rocamora, S.W.A. Reulen, B. de Waal, E.W. Meijer, J.M. Sanz,

M. Merkx, Choline dendrimers as generic scaffolds for the non-covalent synthesis of multivalent protein assemblies, Chem. Commun. 47 (2011) 5997–5999. https://doi.org/10.1039/c0cc05605g.

[31] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera - A visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612. https://doi.org/10.1002/jcc.20084.

[32] J.L. Markley, A. Bax, Y. Arata, C.W. Hilbers, R. Kaptein, B.D. Sykes, P.E. Wright, K. Wüthrich, Recommendations for the presentation of NMR structures of proteins and nucleic acids, J. Mol. Biol. 280 (1998) 933–952. https://doi.org/10.1006/jmbi.1998.1852.

[33] K. Wüthrich, NMR with proteins and nucleic Acids, Europhys. News. 17 (1986) 11–13. https://doi.org/10.1051/epn/19861701011.

[34] K. Wüthrich, M. Billeter, W. Braun, Polypeptide secondary structure determination by nuclear magnetic resonance observation of short proton-proton distances, J. Mol. Biol. 180 (1984) 715–740. https://doi.org/10.1016/0022-2836(84)90034-2.

[35] R. Vila, I. Ponte, M.A. Jiménez, M. Rico, P. Suau, An inducible helix-Gly-Gly-helix motif in the N-terminal domain of histone H1e: A CD and NMR study, Protein Sci. 11 (2009) 214–220. https://doi.org/10.1110/ps.29602.

[36] D.S. Wishart, B.D. Sykes, F.M. Richards, Relationship between nuclear magnetic resonance chemical shift and protein secondary structure, J. Mol. Biol. 222 (1991) 311–333. https://doi.org/10.1016/0022-2836(91)90214-Q.

[37] P. Güntert, Automated NMR structure calculation with CYANA., Methods Mol.

Biol. 278 (2004) 353–378. https://doi.org/10.1385/1-59259-809-9:353.

[38] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, J. Biomol. NMR. 44 (2009) 213–223. https://doi.org/10.1007/s10858-009-9333-z.

[39] R. Koradi, M. Billeter, K. Wüthrich, MOLMOL: A program for display and analysis of macromolecular structures, J. Mol. Graph. 14 (1996) 51–55. https://doi.org/10.1016/0263-7855(96)00009-4.

[40] H.E. Auer, Far-ultraviolet absorption and circular dichroism spectra of L-tryptophan and some derivatives, J. Am. Chem. Soc. 95 (1973) 3003–3011. https://doi.org/10.1021/ja00790a046.

[41] N.J. Greenfield, Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions, Nat. Protoc. 1 (2006) 2527–2535. https://doi.org/10.1038/nprot.2006.204.

[42] Y. Xu, R. Oyola, F. Gai, Infrared study of the stability and folding kinetics of a 15-residue β-hairpin, J. Am. Chem. Soc. 125 (2003) 15388–15394. https://doi.org/10.1021/ja037053b.

[43] S. Honda, N. Kobayashi, E. Munekata, Thermodynamics of a β-hairpin structure: Evidence for cooperative formation of folding nucleus, J. Mol. Biol. 295 (2000) 269–278. https://doi.org/10.1006/jmbi.1999.3346.

[44] C.M. Santiveri, J. Santoro, M. Rico, M.A. Jiménez, Thermodynamic analysis of β-hairpin-forming peptides from the thermal dependence of ^{1}H NMR chemical shifts, J. Am. Chem. Soc. 124 (2002) 14903–14909. https://doi.org/10.1021/ja0278537.

[45] A.G. Cochran, N.J. Skelton, M.A. Starovasnik, Tryptophan zippers: Stable, monomeric β-hairpins, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 5578–5583.

https://doi.org/10.1073/pnas.091100898.

[46] A. Roy, P. Bouř, T.A. Keiderling, TD-DFT modeling of the circular dichroism for a tryptophan zipper peptide with coupled aromatic residues, in: Chirality, 21 Suppl 1:E163-71, 2009. https://doi.org/10.1002/chir.20792.

[47] J.L. Mosser, A. Tomasz, Choline-containing teichoic acid as a structural component of pneumococcal cell wall and its role in sensitivity to lysis by an autolytic enzyme., J. Biol. Chem. 245 (1970) 287–298. https://doi.org/10.1016/s0021-9258(18)63393-9.

[48] J.A. Lycklama a Nijeholt, A.J.M. Driessen, The bacterial Sec-translocase: Structure and mechanism, Philos. Trans. R. Soc. B Biol. Sci. 367 (2012) 1016–1028. https://doi.org/10.1098/rstb.2011.0201.

[49] R. Molina, A. González, M. Stelter, I. Pérez-Dorado, R. Kahn, M. Morales, S. Campuzano, N.E. Campillo, S. Mobashery, J.L. García, P. García, J.A. Hermoso, Crystal structure of CbpF, a bifunctional choline-binding protein and autolysis regulator from *Streptococcus pneumoniae*, EMBO Rep. 10 (2009) 246–251. https://doi.org/10.1038/embor.2008.245.

[50] Q. Xu, J.W. Zhang, Y. Chen, Q. Li, Y.L. Jiang, Crystal structure of the choline-binding protein CbpJ from *Streptococcus pneumoniae*, Biochem. Biophys. Res. Commun. 514 (2019) 1192–1197. https://doi.org/10.1016/j.bbrc.2019.05.053.

[51] Z. Zhang, W. Li, C. Frolet, R. Bao, A.M. Di Guilmi, T. Vernet, Y. Chen, Structure of the choline-binding domain of Spr1274 in *Streptococcus pneumoniae*, Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 65 (2009) 757–761. https://doi.org/10.1107/S1744309109025329.

[52] I. Pérez-Dorado, A. González, M. Morales, R. Sanles, W. Striker, W. Vollmer, S.

Mobashery, J.L. García, M. Martínez-Ripoll, P. García, J.A. Hermoso, Insights into pneumococcal fratricide from the crystal structures of the modular killing factor LytC, Nat. Struct. Mol. Biol. 17 (2010) 576–581. https://doi.org/10.1038/nsmb.1817.

Table 1. Summary of inter-CBR interactions. h_i mean hairpin sequence of CBR*i* and k_i denotes linker sequence between CBR*i* and CBR*i+1*. For instance, the value "3" at the intersection between h_1 and h_2 means 3 interactions between the hairpin of CBR1 and the hairpin of CBR2

	CBR1		**CBR2**		**CBR3**		**CBR4**		**CBR5**		**CBR6**	
CBR substructure	h_1	k_1	h_2	k_2	h_3	k_3	h_4	k_4	h_5	k_5	h_6	k_6
h_1			3	6								
k_1			5	5								
h_2	3	5			5	4						
k_2	6	5			5	4						
h_3			5	5			5	6				
k_3			4	4			4	3				
h_4					5	4			8	6		
k_4					6	3			2	1		
h_5							8	2			8	2
k_5							6	1			4	4
h_6									8	4		
k_6									2	4		
TOTAL	9	10	17	20	21	15	23	12	20	15	12	6
	19		37		36		35		35		18	
Contacts with neighbour sequences	Prev	Next	Prev	Next	Prev	Next	Prev	Next	Prev	Next	Prev	Next
	9[a]	19	19	18	18	18	18	17	17	18	18	*3*[b]*+10*[c]
Interaction area (Å^2)	Nt-CBR1 [a]	CBR1-CBR2		CBR2-CBR3		CBR3-CBR4		CBR4-CBR5		CBR5-CBR6		CBR6-Ct [b+c]
	534	955		924		718		524		771		866

[a] Interactions involving CBR1 with the N-terminal domain

[b] Interactions involving CBR6 with the C-terminal dimerization hairpin within the same monomer

[c] Interactions involving CBR6 with the C-terminal dimerization hairpin of the cognate monomer

Table 2. Main structural statistical parameters for the ensemble of the 20 lowest target function conformers calculated for P5 and P45 peptides in aqueous solution (20 mM Gly buffer, pH 3.0)

	P5	P45
Number of upper limit distance restraints [a]		
Intraresidue (i – j = 0)	73	162
Sequential (|i – j| = 1)	41	36
Medium range (1 < |i – j| < 5)	32	39
Long range (|i – j| ≥ 5)	57	41
Total number	203	278
Averaged total number per residue	14.5	8.2
Number of dihedral angle constraints [b]		
ϕ angles	12	32
ψ angles	12	32
Total number	24	64
Pairwise RMSD (Å)		
Backbone atoms		
Full-length		6 ± 2
Residues 241-250	--------	0.4 ± 0.2
Residues 261-270	0.02 ± 0.01	0.7 ± 0.4
All heavy atoms		
Full-length		7 ± 2
Residues 241-250	--------	1.5 ± 0.4
Residues 261-270	0.7 ± 0.2	1.5 ± 0.5

[a]Obtained from the observed sets of NOEs (see Methods)
[b]Derived from $^{1}H\alpha$, $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts using TALOS+ (see Methods)

Table 3. Thermodynamic quantities of peptides.

Parameter		Peptide P4	P5	P45	P345
T_m (K)	No choline	321 ± 1	316 ± 1	311 ± 1	311 ± 1
	500 mM choline	326 ± 2	325 ± 2	320 ± 1	325 ± 1
ΔH_m (kJ/mol)	No choline	38 ± 3	45 ± 5	45 ± 4	46 ± 2
	500 mM choline	50 ± 5	37 ± 4	62 ± 5	150 ± 10
ΔS_m (J/mol K)	No choline	120 ± 10	130 ± 20	140 ± 10	140 ± 10
	500 mM choline	150 ± 20	110 ± 10	180 ± 20	420 ± 40
ΔG_{UF}^{app} (25 °C) (kJ/mol)	No choline	2.5 ± 0.2	2.5 ± 0.3	1.9 ± 0.2	1.9 ± 0.1
	500 mM choline	4.2 ± 0.4	3.1 ± 0.4	4.2 ± 0.3	12.2 ± 0.6
ΔG_{stab} (25 °C) (kJ/mol)[c]		1.7 ± 0.4	0.6 ± 0.5	2.3 ± 0.4	10.3 ± 0.6
K_d (mM)[b]		360 ± 10	290 ± 20	170 ± 10	117 ± 6

[a] Calculated as ΔG_{UF}^{app} (in 500 mM choline) - ΔG_{UF}^{app} (no choline) using eqn(2)
[b] Macroscopic dissociation constant for choline binding (eqn(4))

Table 4. Midpoint unfolding temperatures for choline-binding modules belonging to CBPs and for P345 peptide.

CBP	Midpoint temperature (ºC)		K_d (mM)
	- Choline	+ Choline (mM)	
LytA [12,18]	40/61[a,b]	71/75 (140)[a,b]	0.1/5.3
LytB [23]	38	61/72 (140)[a]	0.7
LytC [24]	< 4	37 (140)	2.7
CbpD [10]	39	61 (140)[c]	0.3
Pce (CbpE) [22]	40	65 (41)	0.4/2.2
P345[d]	38	52 (500)	117

[a] Multiple thermal transitions.
[b] Partially dimeric at pH 7.0 in the absence of choline, and dimeric in the presence of the ligand.
[c] Marin, V., Maestro, B. and Sanz, J.M., unpublished.
[d] This work

Figure Captions

Figure 1. (A) Cartoon representation of the structure of pneumococcal LytA N-acetylmuramoyl-L-alanine amidase (PDB code 4X36). CBRs 3, 4 and 5 are depicted in magenta, cyan and blue, respectively. A close-up view shows the choline binding sites between CBRs 3-4 and CBRs 4-5, respectively, highlighting the bound choline and the aromatic residues establishing direct contacts with the ligand (in stick representation). (B) Scheme of the inter-CBR interaction network. Wide blue arrows denote the β-strands that configure the β-hairpins within each CBR, whereas orange boxes indicate linker loops between CBRs; solid black lines indicate hairpin-loop interactions; dashed black lines show hairpin-hairpin interactions; red lines correspond to linker-linker contacts; blue lines denote interactions between CBRs and residues outside the ββ-solenoid within the same monomer; green lines indicate interactions between CBR6 and the C-terminal dimerization hairpin of the partner monomer in the dimer. (C) Scheme of the peptides studied in this work (P4, P5, P45 and P345); *h* indicates β-hairpins, and *k* denotes linking loop regions. Numbering scheme of full-length LytA protein.

Figure 2. Far-UV (A) and near-UV (B) ECD spectra of CBM-LytA-derived peptides in Gly buffer (5 ºC).

Figure 3. NMR data for peptides P5 and P45 in aqueous solution. Bar plots of (A) $\Delta\delta_{H\alpha}$ values ($\Delta\delta_{H\alpha} = \delta_{H\alpha}^{observed} - \delta_{H\alpha}^{RC}$, ppm) and (B) $\Delta\delta_{C\alpha}$ values ($\Delta\delta_{C\alpha} = \delta_{C\alpha}^{observed} - \delta_{C\alpha}^{RC}$, ppm) (RC = random coil) as a function of sequence for peptides P45 (black bars), P4 (red bars; data taken from Zamora-Carreras *et al* [20], and P5 (blue bars) in D_2O at pH 3.0. $\delta_{H\alpha}^{RC}$ values were estimated as described in the Experimental section. The N- and C-terminal residues are not shown. The dashed lines indicate the random coil (RC) range. Data for peptide P45 are at 35 °C, and for peptides P4 and P5 at 25 °C.

Figure 4. NMR solution structures of P5 and P45 in aqueous solution: (A) Ensemble of the 20 lowest target function structures of P5 peptide overlaid onto the backbone atoms (black). (B) Ensemble of the 20 lowest target function structures for region 239-252 of P45 peptide overlaid onto the backbone atoms (black). (C) Ensemble of the 20 lowest target function structures for region 259-272 of P45 peptide overlaid onto the backbone atoms (black). (D) A representative conformer from the 20 calculated structures for the P45 peptide. (E-F) Ensemble of the 20 lowest target function structures of P45 peptide

overlaid onto the backbone atoms of residues 241-251 (E) and of residues 261-271 (F). "N" and "C" indicates amino and carboxylate termini, respectively. In panels A-C, side chains colour code is: blue for positively charged; red for negatively charged; cyan for polar; green and magenta for aromatic and hydrophobic side chains pointing downwards and upwards, respectively. The colour code in panels D-F is: hairpin 4 region in red, hairpin 5 region in blue, and linker segment in grey in D and black in E-F.

Figure 5. Comparison of β-hairpin structures formed by peptides P5 and P45 in aqueous solution with those in the native full-length LytA. Ensemble of the 20 lowest target function structures overlaid onto the corresponding atoms in the crystalline LytA structure (PDB code: 4X36, in thick grey) for the regions 239-252 (A-B) and 259-272 (C-D) of the long peptide P45 and for the short peptide P5 (E-F). Backbone in panels B, D and F is horizontally rotated 180° relative to panels A, C and E. Side chains of residues at the turn region and at hydrogen-bonded sites are shown in A, C and E, and those of residues at non-hydrogen-bonded sites are in B, D and F. For the protein, backbone atoms are in thick grey, and side chains in thick gold. For the peptides, backbone atoms are in black, and side chains colour code is: blue for positively charged; red for negatively charged; cyan for polar; green and magenta for aromatic and hydrophobic side chains pointing downwards and upwards, respectively. The amino and carboxylate termini are labelled by "N" and "C", respectively.

Figure 6. Disposition of aromatic residues in the choline-binding subsites of the hairpins corresponding to CBRs 4 and 5: green, isolated P4 and P5 peptides; blue, P45 peptide; grey, full-length LytA amidase (PDB 4X36). For presentation purposes, representative peptide conformations were selected among the ensemble of 20 lowest target function NMR structures showing aromatic-aromatic dihedral angles closest to the average of the NMR ensemble (see Table S4). For P4, angles of depicted aromatic side chains are: 136° (Trp241-Trp248) and 29° (Trp248-Tyr250); for P5, angles are: 115° (Trp261-Trp268) and 80° (Trp268-Tyr270); for P45, angles are 105° (Trp241-Trp248), 71° (Trp248-Tyr250), 127° (Trp261-Trp268) and 68° (Trp268-Tyr270).

Figure 7. Thermal unfolding scans of peptides monitored by far-UV ECD in Gly buffer in the absence (black circles) and presence (red circles) of 500 mM choline. Solid lines

display fittings to the Gibbs-Helmholtz equation (eqn (1)) and correspond to the thermodynamic quantities shown in Table 3.

Figure 8. *Upper panels*: Far-UV and near-UV ECD spectra of peptides and the C-LytA module in the presence of 500 mM choline; *Lower panels*, individual comparisons in the absence (black line) and the presence (red line) of 500 mM choline. The C-LytA spectra were recorded in 20 mM sodium phosphate buffer, pH 7.0, since the protein precipitates at pH 3.0. *Insets*, difference spectra between choline-bound and choline-free peptides and protein.

Figure 9. Titration of peptides with choline monitored by intrinsic fluorescence. Data are the average of three experiments. Solid lines represent fittings to a simple binding model (eqn (4)).

Figure 1

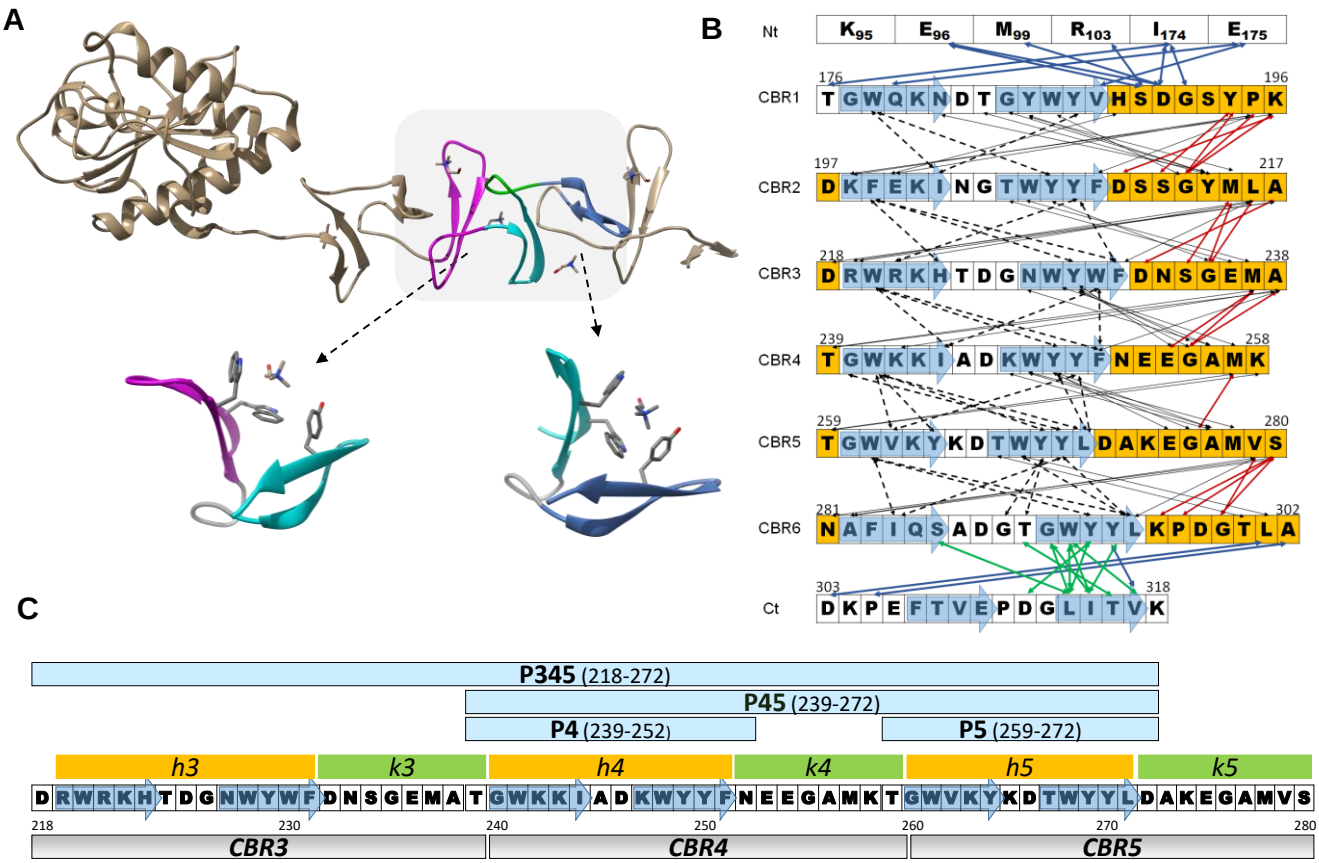

A
B
Nt
K95
E96
M99
R103
I174
E175
CBR1
176
196
T G W Q K N D T G Y W Y V H S D G S Y P K
CBR2
197
217
D K F E K I N G T W Y Y F D S S G Y M L A
CBR3
218
238
D R W R K H T D G N W Y W F D N S G E M A
CBR4
239
258
T G W K K I A D K W Y Y F N E E G A M K
CBR5
259
280
T G W V K Y K D T W Y Y L D A K E G A M V S
CBR6
281
302
N A F I Q S A D G T G W Y Y L K P D G T L A
Ct
303
318
D K P E F T V E P D G L I T V K
C
P345 (218-272)
P45 (239-272)
P4 (239-252)
P5 (259-272)
h3
k3
h4
k4
h5
k5
D R W R K H T D G N W Y W F D N S G E M A T G W K K I A D K W Y Y F N E E G A M K T G W V K Y K D T W Y Y L D A K E G A M V S
218
230
240
250
260
270
280
CBR3
CBR4
CBR5

Figure 2

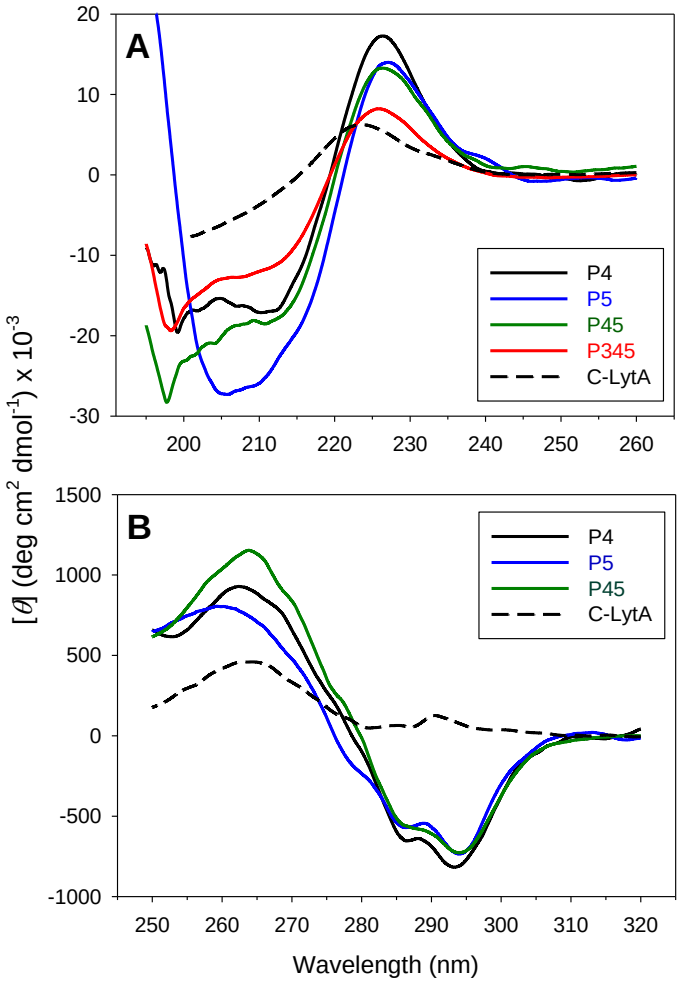


Figure 3

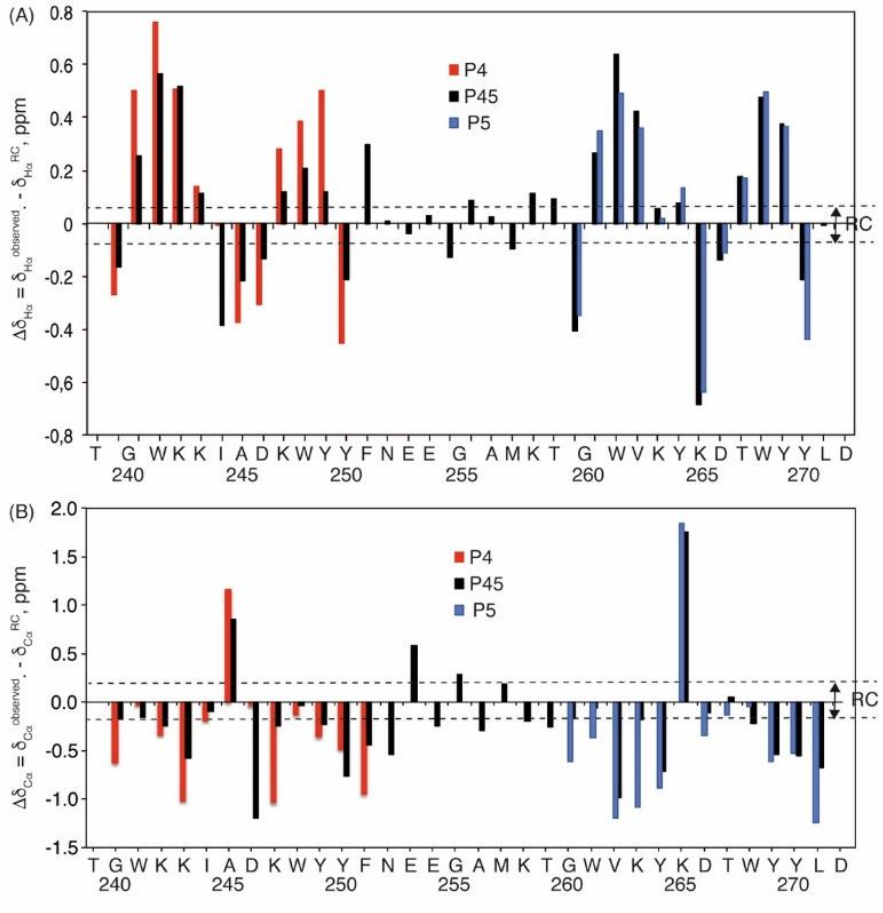

(A)
P4
P45
P5
RC
T G W K K I A D K W Y Y F N E E G A M K T G W V K Y K D T W Y Y L D
240 245 250 255 260 265 270
(B)
P4
P45
P5
RC
T G W K K I A D K W Y Y F N E E G A M K T G W V K Y K D T W Y Y L D
240 245 250 255 260 265 270

Figure 4

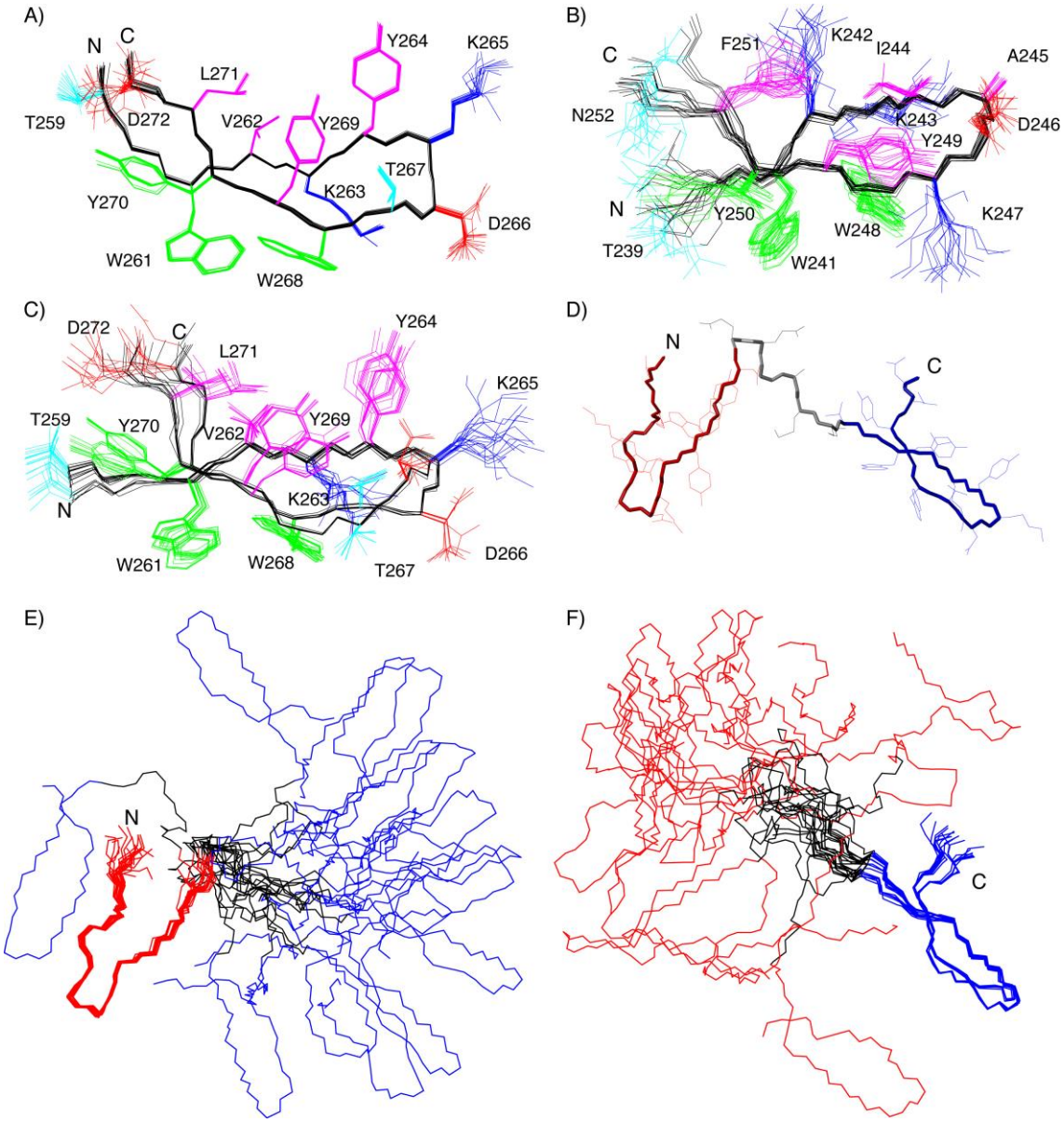

A)
N
C
L271
T259
D272
V262
Y264
Y269
K265
T267
K263
Y270
D266
W261
W268
B)
C
F251
K242
I244
A245
N252
K243
D246
Y249
N
Y250
W248
K247
T239
W241
C)
D272
C
L271
Y264
K265
T259
Y270
V262
Y269
K263
N
W261
W268
T267
D266
D)
N
C
E)
N
F)
N
C

Figure 5

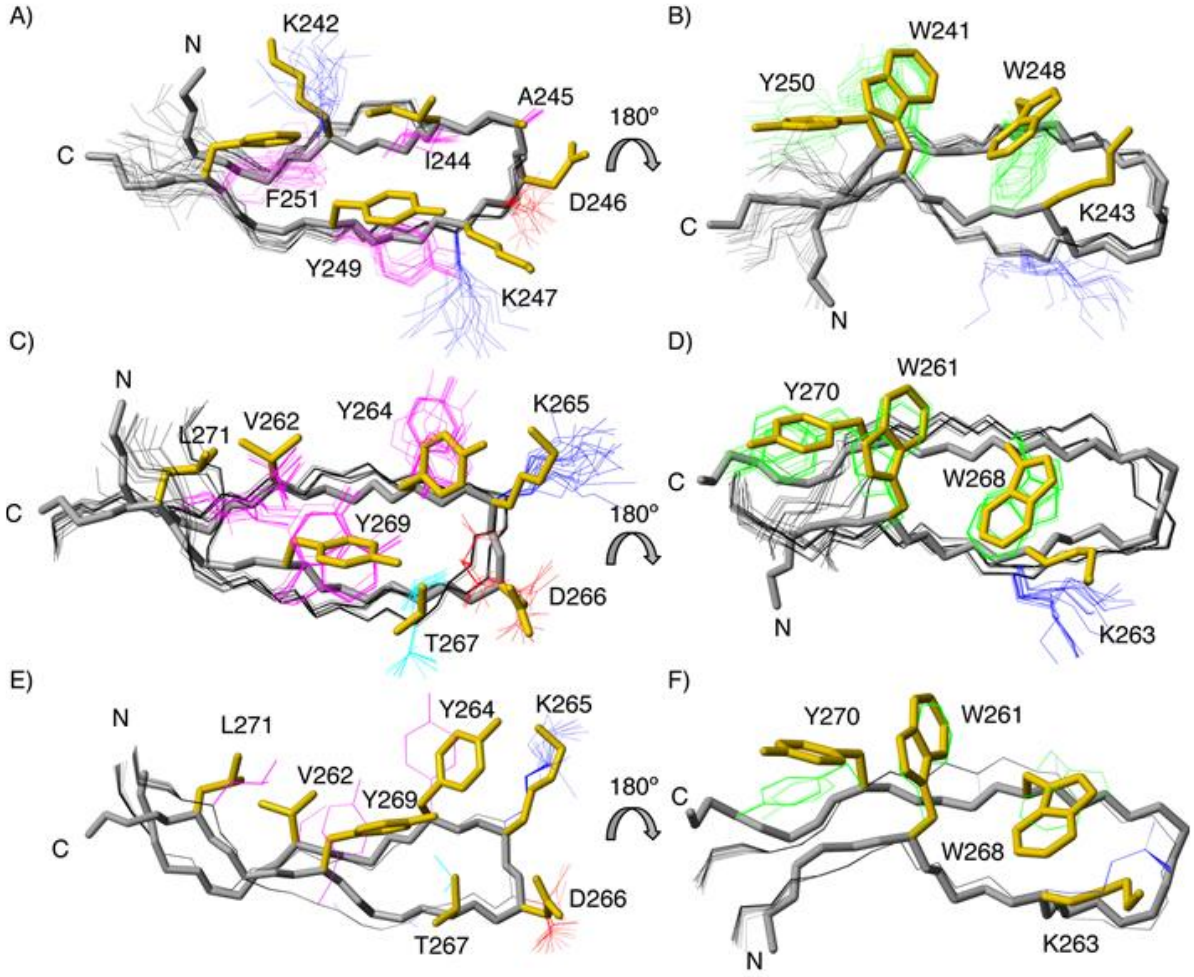

A)
K242
N
A245
C
I244
F251
D246
Y249
K247
180°
B)
W241
Y250
W248
K243
C
N
C)
N
L271
V262
Y264
K265
C
Y269
D266
T267
180°
D)
W261
Y270
C
W268
N
K263
E)
N
L271
Y264
K265
V262
Y269
C
D266
T267
180°
F)
Y270
W261
C
W268
N
K263

Figure 6

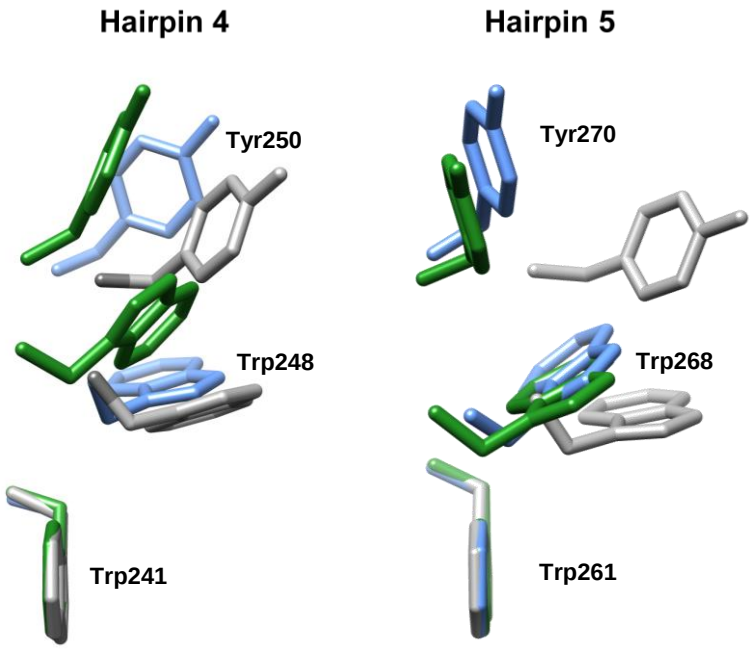

Hairpin 4
Hairpin 5
Tyr250
Tyr270
Trp248
Trp268
Trp241
Trp261

Figure 7

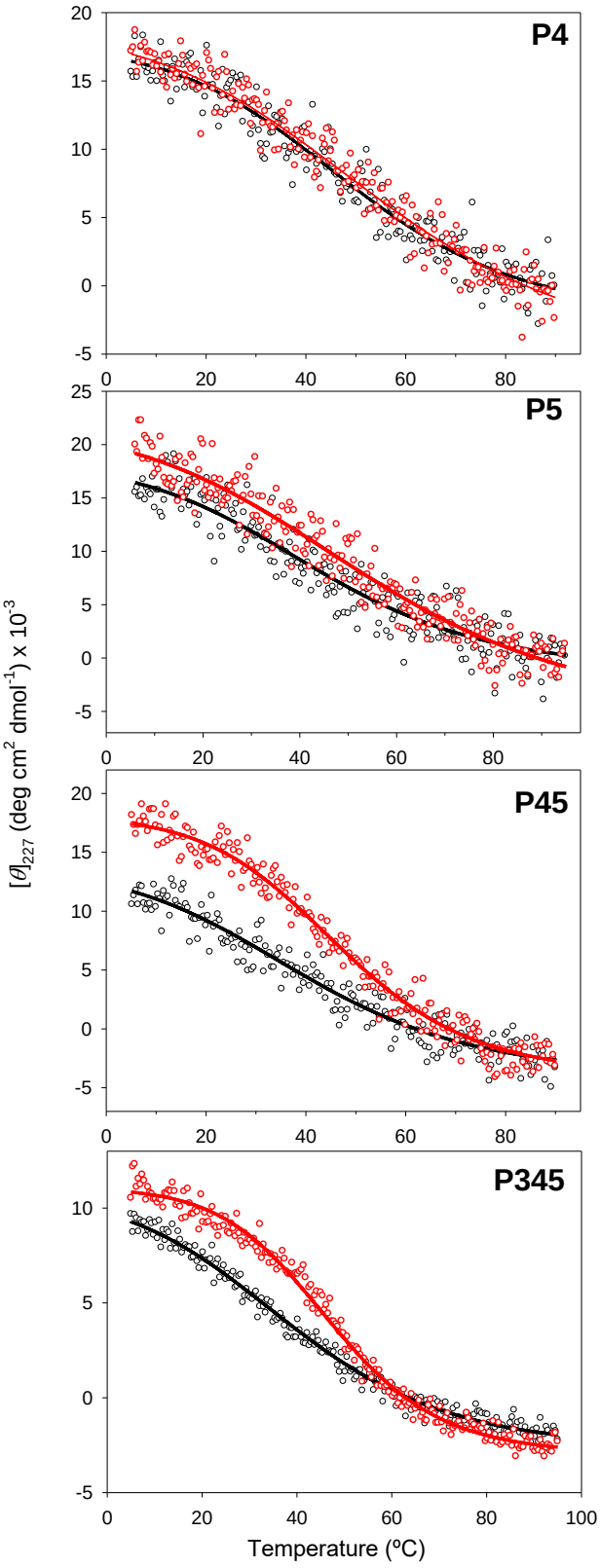


Figure 8

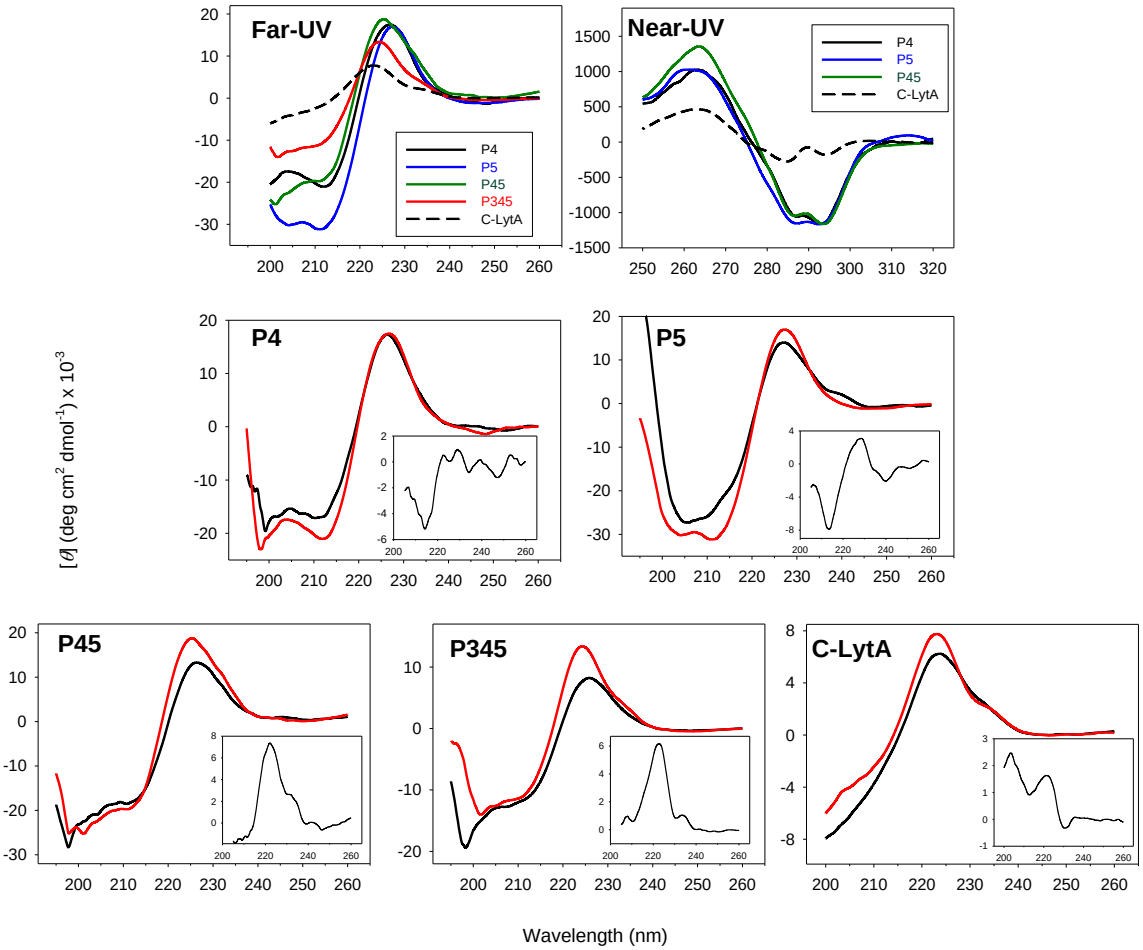
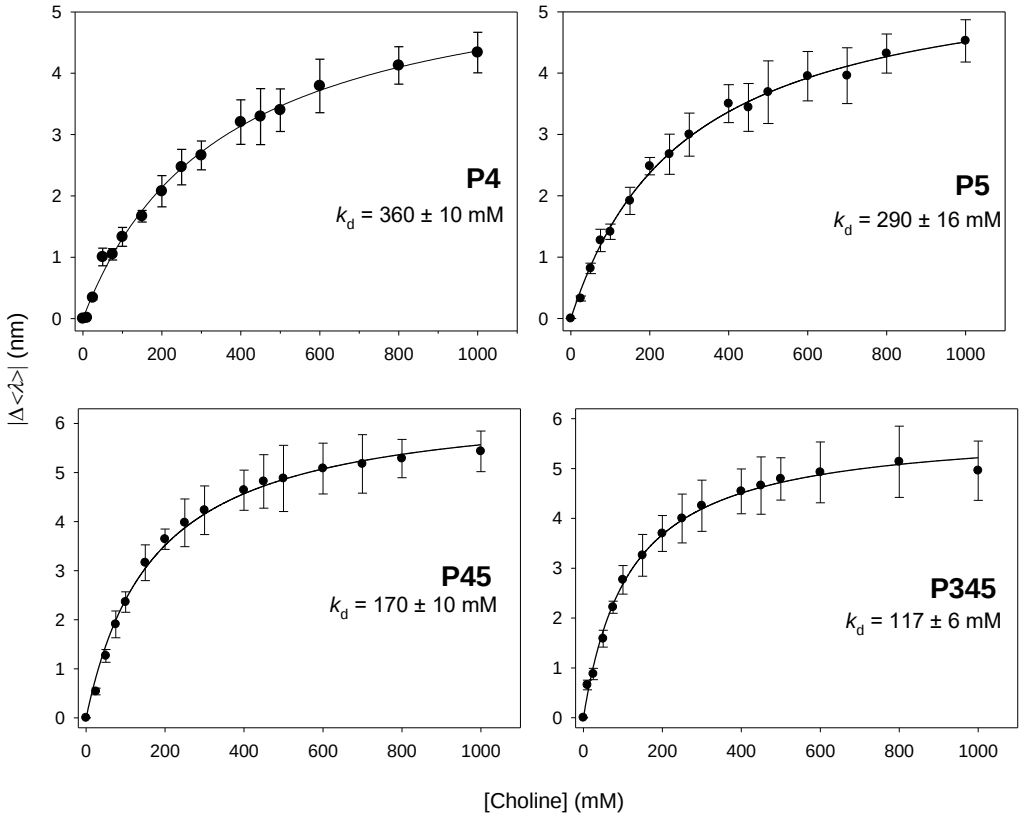

Far-UV
Near-UV
P4
P5
P45
P345
C-LytA
[Θ] (deg cm2 dmol-1) x 10-3
Wavelength (nm)

Figure 9


P4
k_d = 360 ± 10 mM
P5
k_d = 290 ± 16 mM
P45
k_d = 170 ± 10 mM
P345
k_d = 117 ± 6 mM
|Δ<λ>| (nm)
[Choline] (mM)