Received: 1 March 2023 | Revised: 9 September 2023 | Accepted: 25 September 2023

DOI: 10.1002/prot.26611

RESEARCH ARTICLE

PROTEINS WILEY

Lessons on protein structure from interleukin-4: All disulfides are not created equal

Daniela C. Vaz[1,2,3,4] | J. Rui Rodrigues[3,4] | Nuno Loureiro-Ferreira[5] | Thomas D. Müller[6] | Walter Sebald[7] | Christina Redfield[8] | Rui M. M. Brito[2]

[1]School of Health Sciences, Polytechnic of Leiria, Leiria, Portugal

[2]Chemistry Department, Faculty of Sciences and Technology, University of Coimbra, Coimbra Chemistry Centre, Institute of Molecular Sciences, Coimbra, Portugal

[3]Laboratory of Separation and Reaction Engineering—Laboratory of Catalysis and Materials (LSRE-LCM), School of Technology and Management, Polytechnic of Leiria, Leiria, Portugal

[4]Associate Laboratory in Chemical Engineering (ALiCE), University of Porto, Porto, Portugal

[5]SURF, Amsterdam, The Netherlands

[6]Department of Molecular Plant Physiology and Biophysics, Julius-von-Sachs-Institute, University of Würzburg, Würzburg, Germany

[7]Department of Physiological Chemistry II, Theodor-Boveri-Institute (Biocentre), University of Würzburg, Würzburg, Germany

[8]Department of Biochemistry, University of Oxford, Oxford, United Kingdom

Correspondence
Daniela C. Vaz and Rui M. M. Brito, Chemistry Department, Faculty of Sciences and Technology, University of Coimbra, Coimbra Chemistry Centre, Institute of Molecular Sciences, 3004-535 Coimbra, Portugal.
Email: dvaz@uc.pt and rbrito@uc.pt

Funding information
European Regional Development Fund; Fundação para a Ciência e a Tecnologia, Grant/Award Numbers: UIDB/00313/2020, UIDP/00313/2020; FEDER/COMPETE, Grant/Award Numbers: CENTRO-07-CT62-FEDER-002012, RECI/QEQ-QFI/0168/2012; Rede Nacional de Ressonância Magnética Nuclear (RNRMN/PTNMR); ALiCE, Grant/Award Number: LA/P/0045/2020; LSRE-LCM, Grant/Award Numbers: UIDP/50020/2020, UIDB/50020/2020; FCT/MCTES (PIDDAC)

Abstract

Interleukin-4 (IL-4) is a hematopoietic cytokine composed by a four-helix bundle stabilized by an antiparallel beta-sheet and three disulfide bonds: Cys3-Cys127, Cys24-Cys65, and Cys46-Cys99. IL-4 is involved in several immune responses associated to infection, allergy, autoimmunity, and cancer. Besides its physiological relevance, IL-4 is often used as a "model" for protein design and engineering. Hence, to understand the role of each disulfide in the structure and dynamics of IL-4, we carried out several spectroscopic analyses (circular dichroism [CD], fluorescence, nuclear magnetic resonance [NMR]), and molecular dynamics (MD) simulations on wild-type IL-4 and four IL-4 disulfide mutants. All disulfide mutants showed loss of structure, altered interhelical angles, and looser core packings, showing that all disulfides are relevant for maintaining the overall fold and stability of the four-helix bundle motif, even at very low pH. In the absence of the disulfide connecting both protein termini Cys3-Cys127, C3T-IL4 showed a less packed protein core, loss of secondary structure (∼9%) and fast motions on the sub-nanosecond time scale (lower S^2 order parameters and larger τ_c correlation time), especially at the two protein termini, loops, beginning of helix A and end of helix D. In the absence of Cys24-Cys65, C24T-IL4 presented shorter alpha-helices (14% loss in helical content), altered interhelical angles, less propensity to form the small anti-parallel beta-sheet and increased dynamics. Simultaneously deprived of two disulfides (Cys3-Cys127 and Cys24-Cys65), IL-4 formed a partially folded "molten globule" with high 8-anilino-1-naphtalenesulphonic acid-binding affinity and considerable loss of secondary structure (∼50%decrease), as shown by the far UV-CD, NMR, and MD data.

KEYWORDS
^{15}N NMR relaxation, disulfide bonds, four-helix bundle, interleukin-4, molecular dynamics simulations, RDC-refined NMR solution structures

1 | INTRODUCTION

Disulfide bridges are highly conserved elements of protein structure.[1–3] Disulfides bring together two non-adjacent cysteine residues (Cys-Cys), and covalently link distinct segments or domains, of the same, or of different polypeptide chains. When not functionally relevant, disulfides act as structural scaffolds, either by imposing geometrical restraints and reducing the entropy decrease upon protein folding, or by providing additional conformational stability under extreme conditions.[4–8] Disulfide bonds also increase the half-life of proteins by protecting the polypeptide chain from oxidation and proteolytic enzymes,[9] and/or by thermodynamically stabilizing native conformations.[10–19] Proteins containing stable disulfides are primarily present in the oxidizing extracellular milieu of most living organisms. The extracellular domain of the Light-Density-Lipoprotein receptor[20] for instance, presents one of the highest numbers of intra-chain disulfide bonds (26 disulfides) found in proteins. Moreover, extracellular proteins that contain disulfides rarely present more than one free sulfhydryl (–SH) which normally has a functional role.[2,21] Conversely, intracellular proteins generally do not present disulfides, given the reducing environment of the cytosol.[22,23] Therefore, it is paramount to characterize the relevance of these structural elements to better understand how they influence protein structure, stability, and dynamics as well as to provide new insights in protein design and engineering.

Interestingly, a pattern of disulfide bond conservation has been found among cytokines.[2,24,25] The preservation of these structural elements through evolution proves their relevance for cytokine function and stability, being most of the conserved disulfides local (< 50 residues apart). In contrast, non-conservation has been associated with the loss of both cysteines involved in the disulfide, since the presence of a free reactive —SH group is extremely unfavorable.[1,2] Short-chain cytokines (~130 residues) generally present one to two disulfides per protein, with short or local connections.[2,24] However, interleukin-4 (IL-4) is an exception, with three native disulfides. IL-4 is a 129-residue (14.5 kDa) short-chain cytokine composed by an *up-up-down-down* four-helix bundle, an antiparallel β-sheet and three disulfide bonds, two local (Cys24-Cys65 and Cys46-Cys99), and an extra non-local (Cys3-Cys127).[26–29] IL-4 is involved in the mediation of several immune responses associated to infection, allergy, autoimmunity, and cancer, among others.[30–33] Beyond its functional relevance, IL-4 presents a well-defined core,[26–28] high conformational stability,[34] and is often used as a "model structure" for protein studies. Nonetheless, even if IL-4 is a thermodynamically stable protein ($\Delta G(H_2O)_{25^\circ C}$ 9.3 kcal/mol, T_m 83.8°C),[34] removing disulfides Cys3-Cys127 and Cys24-Cys65 by site-directed mutagenesis[35] leads to the formation of much less conformationally stable proteins: C3T-IL-4 ($\Delta G(H_2O)_{25^\circ C}$ 4.3 kcal/mol, T_m 70.0°C) and C24T-IL-4 ($\Delta G(H_2O)_{25^\circ C}$ 6.0 kcal/mol, T_m 73.5°C).[34] Hence, to further understand the role of each disulfide in the overall protein structure, packing, and dynamics of IL-4, we have carried out spectroscopic studies (nuclear magnetic resonance [NMR], circular dichroism [CD], intrinsic and extrinsic fluorescence) and molecular dynamics (MD) simulations on four disulfide mutants of IL-4: C3T-IL4, C24T-IL4, C99T-IL4 and 1SS-IL4. C3T-IL-4 has Cys 3 replaced by threonine and lacks disulfide Cys3-Cys127 (nonlocal extra disulfide, cysteines >100 residues apart), which brings together the two protein termini in *wild-type* IL-4 (WT-IL-4). C24T-IL4 has Cys 24 mutated by threonine and therefore lacks disulfide Cys24-Cys65 (local disulfide, <50 residues apart) that connects the two long interhelical loops AB and CD. In turn, 1SS-IL4 lacks simultaneously both disulfides Cys3-Cys127 and Cys24-Cys65. Given the relevance of disulfide Cys46-Cys99 for protein folding (since it is formed between a core structural element in α-helix [Helix B] and loop CD), it was not possible to purify a stable mutant protein without this disulfide.[35] Nonetheless, we have also carried out MD simulations on a C99T-IL4 mutant, lacking disulfide Cys46-Cys99. The design of these mutant proteins allowed us to, both independently and simultaneously (1SS-IL4), study the effect of IL-4 native disulfides (local and non-local) on the overall 3D structure, dynamics and packing of the protein.

2 | MATERIALS AND METHODS

All chemicals were of the highest purity commercially available and were purchased from Sigma-Aldrich (Merck KGaA, Germany). Isotope-labeled compounds were acquired from Cambridge Isotope Laboratories Inc. (MA, USA).

2.1 | Protein expression and purification

Recombinant proteins WT-IL4, C3T-IL4, C24T-IL4, and C3AC24AC65SC127S-IL4 (in short 1SS-IL4) were engineered by site-directed mutagenesis using WT-IL-4 cDNA (British Bio-technology Ltd., Oxford, UK) and produced in *Escherichia coli* K12 JM103 recA−. Constructs were produced by the integration of the respective DNA sequence between the XhoI and BamHI sites of the expression vector RTSpRC109.[36] Proteins were refolded from inclusion bodies and purified as described previously.[36] Purified samples were freeze-dried and stored at −20°C. Unlabeled and uniformly ^{15}N-single-labeled and $^{13}C/^{15}N$-double-labeled human interleukin-4 (IL-4) samples were produced and purified as described previously.[26–28] Recombinant proteins were tested for activity by surface plasmon resonance in a BIAcore 2000 system, using immobilized α-chain domains of the IL-4 receptor (IL-4Rα) at a density of 15–20 pg/mm^2 in a CM5 biosensor. Protein concentrations were determined spectrophotometrically at 278 nm (extinction coefficient of 9370 M^{-1} $cm^{-1}/\epsilon_{278}^{0.1\%} = 0.625$ mg^{-1} cm^2).[37] Quality control of protein samples was performed by analytical size exclusion chromatography and polyacrylamide gel electrophoresis, to confirm the absence of dimers. Uniformly labeled samples were prepared in 25 mM $NaCD_3COO/CD_3COOD$, 9:1 H_2O/D_2O, 0.05% NaN_3, pH 5.0. Deuterated solutions of 0.1 M DCl and NaOD were used for pH adjustments.

2.2 | NMR spectroscopy

NMR experiments were carried out on uniformly labeled IL-4 samples, in 25 mM $NaCD_3COO/CD_3COOD$, 9:1 H_2O/D_2O, 0.05% NaN_3, pH 5.0, 298 K. Experiments were run on home-built spectrometers

equipped with triple resonance probes operating at ^{1}H-frequencies of 500, 600, and 750 MHz and Oxford Instruments Company magnets (Oxford, UK), OMEGA software, digital control equipment (Bruker Instruments), home-built triple resonance pulsed field gradient probe heads and home-built linear amplifiers. Data processing was performed with FELIX (Accelrys, Inc, USA). Zero filling and phase-shifted sinebell squared apodization functions were applied during data processing.

The restraints used to calculate the Residual Dipolar Coupling (RDC)-refined NMR solutions were obtained as previously described,[28,38] along with additional RDC restraints and dihedral angle restraints derived from the Hα, ^{13}C, and ^{15}N chemical shifts (Table S1). Dipolar couplings were measured at 35°C using a 2D in-phase and anti-phase [^{1}H—^{15}N] Heteronuclear Single Quantum Coherence (HSQC) correlation experiment.[39,40] Pairs of downfield and upfield spectra were produced in isotropic and anisotropic media (in the presence of a compressed 5% polyacrylamide gel). Parameters such as D_a and R (rhombicity ($R = D_r/D_a$, where D_a is the axial component and D_r is the rhombic component of the alignment tensor)[40,41] were estimated via the Xplor-NIH python script *CalTensor.py*[42] by making use of the observed RDCs and of the previous NMR solution structure of WT-IL4 (1ITM.pdb). Computed values for D_a and R were of 3.74 Hz and 0.49 for WT-IL4 at pH 5.6; of 3.53 Hz and 0.66 for WT-IL4 at pH 2.4, and of −1.78 Hz and 0.25 for C3T-IL4. Additional dihedral angle restraints for the torsion angles ϕ, ψ, ω, and χ1, derived from the Hα, ^{13}C, and ^{15}N chemical shifts, by PREDICTOR (PREDIction of TORsion angles from chemical shift and homology),[43] were also used in the calculations.

2.3 | NMR structure calculation and analysis

Structures were calculated with Xplor-NIH.[42] Computations started from a fully extended conformation with the correct stereochemistry and included all experimental restraints (Table S1). The process was initiated with a simulated annealing (SA) protocol with r^{-6} averaging for non-stereospecifically assigned protons.[44] Calculations included a pseudopotential for the protein radius of gyration (R_{gyr}) to improve protein packing and a conformational database potential for the optimization of dihedral angle distributions.[45] Initial annealing consisted of 10 000 steps starting at 5000 K, with slow cooling steps of 6.25–25 K. Annealing was followed by a refinement stage consisting of 10 000 steps at 3000 K, and 10 000 steps during cooling to 25 K. Stereochemical quality of the final structures was analyzed with PROCHECK-NMR.[46] Regions of secondary structure were identified with STRIDE[47] whereas solvent accessible surface areas were obtained using NACCESS[48] and inter-helical angles measured with Q-helix.[49] Hydrogen bonds were defined for donor (D) ⋯ acceptor (A) distances shorter than 3.2 Å and D ⋯ H ⋯ A, with angles deviating less than 35° from linearity, and were evaluated using visual molecular dynamics (VMD).[50] Protein packing was evaluated according to the occluded surface area (OSA) algorithm.[51]

2.4 | NMR relaxation experiments and backbone dynamics

^{15}N NMR relaxation measurements were performed on ^{15}N-labeled IL-4 samples, 25 mM $NaCD_3COO/CD_3COOD$, 9:1 H_2O/D_2O, 0.05% NaN_3, at pH 5.0, 308 K. Relaxation measurements were carried out using 2D sensitivity-enhanced heteronuclear NMR spectroscopy to measure T_1, T_2, and the heteronuclear nuclear Overhauser effect (NOEs),[52–54] at a ^{15}N frequency of 50.7 MHz, as reported previously for WT-IL-4,[38] for comparison. Heteronuclear NOE experiments were carried out at both 50.7 and 60.8 MHz,[53] with and without NOE.

The inertia tensor calculated from the average NMR solution structure of C3T-IL-4, pointed towards an almost spherical shape (1.39:1.25:1.00), as seen previously for WT-IL-4 (inertia tensor 1.81:1.46:1.00, 1ITM.pdb). Thus, the ^{15}N relaxation data for C3T-IL-4 were analyzed according to the formalisms proposed by Lipari and Szabo,[55,56] with extensions by Clore et al.,[57,58] as applied previously to WT-IL-4.[38] The theoretical values for the T_1 and T_2 relaxation times and the NOE were calculated using the appropriate relaxation models,[55,56] with the respective order and time parameters S^2, S^2_s, S^2_f, τ_R, τ_e and τ_s,[55–58] and compared with the values extracted from the experimental data. Procedures used to extract order parameters and correlation times for internal motions from the relaxation parameters are summarized in Table S2. The mean T_1/T_2 ratio per residue in C3T-IL-4, with a NOE greater than 0.75 at 50.7 MHz, that did not show evidence of significant fast motional averaging or exchange broadening was found to be 8.11 ± 0.73, ranging from 6.30 for L7 to 15.61 for T63. Therefore, for residues with a T_1/T_2 ratio <7.38 (mean value minus one standard deviation), τ_e (effective correlation time for internal motion) is making a significant contribution, while for residues with a T_1/T_2 ratio >8.84 (mean value plus one standard deviation) chemical exchange broadening is producing a significant decrease in T_2. Since the overall rotational correlation time (τ_R) must be the same for all residues, the eight residues that had T_1/T_2 ratios outside one standard deviation of the mean were excluded from the calculation of τ_R, and the remaining 24 residue (located in helical regions, with a mean T_1/T_2 ratio of 8.15 ± 0.35) were used to obtain a single value of τ_R of 10.3 ± 0.1 ns, which was used in all subsequent analysis.

2.5 | CD spectroscopy

CD experiments were performed on an Olis DSM20 CD spectropolarimeter calibrated with d-10-camphorsulfonic acid. Far-UV CD spectra were recorded between 260 and 182 nm using a 1-mm pathlength cuvette. Protein concentrations were about 10 μM, in 10 mM sodium phosphate buffer, at pH 6.0, and in 10 mM sodium citrate buffer, pH 3.0. CD spectra were run with a step-resolution of 0.2 nm, a scan speed of 50 nm/min, and using a bandwidth of 1.0 nm. Spectra were averaged over five scans and corrected by subtraction of the buffer signal. Results were expressed as mean residue ellipticities $[\Theta]_{MRW}$, where Θ_{obs} is the observed ellipticity in millidegrees, MRW is the

mean residue weight, c is the concentration (mg/mL), and l is the length of the light path (cm).

$$[\Theta]_{MRW} = \frac{\Theta_{obs}(0.1MRW)}{l \times c}.$$

Secondary structure was estimated from the far-UV CD spectra by CONTIN.[59]

2.6 | Urea-induced unfolding experiments

Fluorescence experiments were performed on a Varian Cary Eclipse fluorescence spectrophotometer equipped with a thermostated cell. Intrinsic fluorescence emission spectra were obtained at controlled temperature, using an excitation wavelength of 295 nm. Urea-induced unfolding experiments were performed by dilution of stock solutions of IL-4 to a final concentration of 10 μM in the presence of increasing concentrations of denaturant, either in 10 mM of sodium phosphate buffer at pH 6.0 and pH 7.5 or in 10 mM citrate–phosphate buffer at pH 3.0, pH 4.0, and pH 5.0. Fresh stock solutions of urea were prepared gravimetrically, with final concentrations calculated according to Pace and Scholtz.[60] Protein samples with urea were incubated at a controlled temperature of 298 K for up to 24 h. Reversibility of the unfolding process was confirmed by extensive dialysis of urea denatured samples against 10 mM sodium phosphate buffer (pH 6.0) and compared with native samples by gel filtration chromatography and fluorescence. IL-4 urea-unfolding profiles were constructed plotting fluorescence intensity at 380 nm against urea concentration.

Equilibrium chemical unfolding curves were analyzed using a two-state model (native [N] to unfolded [U]). The experimentally observed spectroscopic property (y) as a function of urea concentration is the result of contributions from both native (N) and unfolded (U) populations, and may be directly related to the equilibrium constant and free energy change [$\Delta G(H_2O)$] of unfolding, according the extrapolation model described previously[35]:

$$y = \frac{(Y_N + m_N[\text{urea}]) + (Y_U + m_U[\text{urea}]).e^{\frac{(m[\text{urea}] - \Delta G(H_2O))}{RT}}}{1 + e^{\frac{(m[\text{urea}] - \Delta G(H_2O))}{RT}}}.$$

The free energy change in the absence of denaturant [$\Delta G(H_2O)$] and m, the dependence of the free energy (ΔG) on the concentration of denaturant, were determined by nonlinear least squares fits to the unfolding data, using the program Origin (OriginLab Corporation).

2.7 | 8-Anilino-1-naphtalenesulphonic acid-binding experiments

8-Anilino-1-naphtalenesulphonic acid (ANS)-binding experiments were performed by making small additions of concentrated ANS stock solutions to the protein samples. Protein concentrations were 13 μM for WT-IL4 and 16 μM for 1SS-IL4, in 10 mM sodium phosphate buffer, pH 6.0 and in 10 mM citrate–phosphate buffer, pH 2.4. Final concentrations of ANS varied between 1 and 50 μM. Total sample dilutions never exceed 2%. ANS fluorescence spectra were monitored between 400 and 600 nm, at 298 K, using an excitation wavelength of 370 nm. Apparent dissociation constants (K_d) were determined by nonlinear least square fitting to the titration data (excitation at 370 nm and emission at 450 nm), assuming a simple protein-ligand binding model:

$$y = \left(1 + [L]_{total}/[P]_{total} + K_d/[P]_{total}\right) - \frac{\sqrt{\left(1 + [L]_{total}/[P]_{total} + K_d/[P]_{total}\right)^2 - 4[L]_{total}/[P]_{total}}}{2}$$

where [P] and [L] represent the concentration of protein and ligand (ANS), respectively, and K_d is the apparent dissociation constant.

2.8 | MD simulations

The crystal structure, and not the NMR structure to avoid bias, of human WT-IL-4 obtained by Müller and collaborators (2B8U.pdb) with 1.8 Å resolution was used as a starting point for the MD simulations. In addition to the protein molecule, the models contained 131 water molecules and 3 sulphate ions. Except for one sulphate ion and one water molecule, all the remaining atoms in the model had a 1.0 occupancy. Sulphate ions were removed from the model and the protonation state of histidine residues was assessed with WHAT IF.[61] H1 and H59 were found to be protonated at Nδ1 and H58, H74 and H76 at Nϵ2. The structures of C3T-IL4, C24T-IL4, C99T-IL4, and 1SS-IL4 were generated by mutation of WT-IL4 (2B8U.pdb) with Modeller 9.19.[62] All IL-4 variants, with water molecules retained, were placed in rhombic dodecahedron cells, keeping a minimum distance of 15 Å between solute atoms and cell boundaries and were solvated with TIP3P water molecules. System charge neutrality was achieved by adding chloride ions to each system. All energy minimizations and MD simulations were performed with GROMACS 2018.6[63] using the Amber ff99SB-ILDN force field 2018.6.[64] Each system was energy-minimized for 2000 steps using the steepest descent algorithm. Equilibration was performed in three steps: (1) 0.2 ns of MD at constant temperature (298 K) with protein and solvent coupled separately to two thermostats using the V-rescale algorithm ($\tau = 0.1$ ps)[65] and with position restraints (1000 kJ mol^{-1} nm^{-1}) placed in protein heavy atoms; (2) followed by 0.2 ns at constant temperature (V-rescale algorithm) and pressure (Berendsen algorithm, 1 bar, $\tau = 2$ ps),[66] while maintaining the position restraints; and (3) finally by 0.2 ns without position restraints and changing the pressure control to the Parrinello-Rahman algorithm.[67] Runs (1000 ns) started from this point. Equilibration and production procedures were repeated three times for each system. A time step of 2 fs was used throughout the simulations. The particle-mesh Ewald algorithm was used for computing electrostatic interactions with a cut-off of 10 Å and a grid spacing of 1.2 Å with

fourth-order B-spline interpolation. Van der Waals interactions were calculated with a cut-off of 10 Å, with long range dispersion corrections for energy and pressure. The LINCS algorithm[68] was used to constrain hydrogen-heavy atom bond lengths. Trajectory coordinates were recorded at every 10 ps. Systems details are summarized in Table S3. VMD and GROMACS were used to analyze the resulting MD trajectories. Principal Components Analyses used C-alpha carbons coordinates, extracted from the 15 trajectories (5 variants × 3 runs) pooled together as a single "trajectory." GROMACS *covar* was used to calculate and diagonalize the covariance matrix of the concatenated trajectory using the protein coordinates prior to energy minimization, as a reference structure. Protein "tails" (residues 1–2 and 128–129) were excluded from the analysis. This originated a set of 375 eigenvalues (125 atoms times 3D Cartesian coordinates). GROMACS *gmx anaeig* was used to calculate the projection of the individual trajectories onto the first and second eigenvectors. GROMACS *rotacf* was used to calculate the rotational correlation functions of the NH atom pairs from the MD trajectories. Correlation functions were calculated for all residues of the protein, except for the first one (His) and for proline residues. Trajectories were previously aligned to a same reference structure to remove protein translation and rotations. Order parameters S^2 were computed from the MD trajectories by averaging the second half of the rotational correlation curves. AlphaFold2[69] was also used to produce models of the 3D structure of the four IL-4 mutants (C3T-IL4, C24T-IL4, C99T-IL4, and 1SS-IL4) from their amino acid sequence.

3 | RESULTS AND DISCUSSION

3.1 | Loss of structural integrity in the absence of Cys24-Cys65 or Cys3-Cys127

Comparative analysis of the NMR solution structures of C3T-IL4 (IL-4 mutant lacking disulfide Cys3-Cys127) and C24T-IL4 (IL-4 mutant lacking disulfide Cys24-Cys65) showed that in the absence of one, Cys3-Cys127 or Cys24-Cys65, out of the three native disulfides, both mutant proteins presented differences in protein structure and packing, relatively to WT-IL4 (Figure 1). Loss of secondary structure, along with changes in interhelical angles, and looser and less packed protein structural cores, especially within the helical and loop regions, were observed. NMR solution structures of C3T-IL4 (PBD ID 8A4F) and

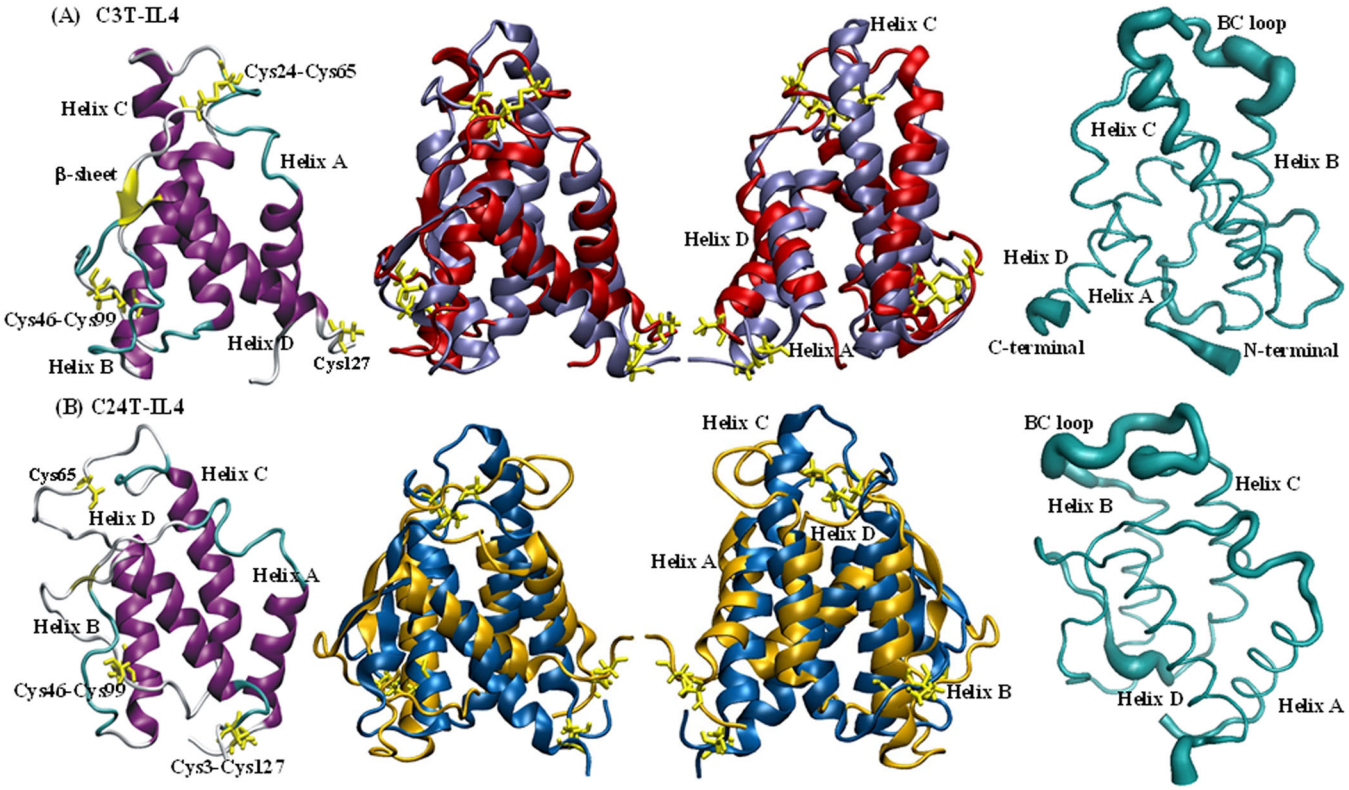

FIGURE 1 Nuclear magnetic resonance (NMR) solution structures of the interleukin-4 (IL-4) disulfide mutants. Ribbon representations of the three-dimensional structures of C3T-IL4 (A) and C24T-IL4 (B). Left diagrams show the three-dimensional structures of the lowest energy structures of C3T-IL4 and C24T-IL4, at pH 5.0, 298 K, with the four α-helices (purple) labeled A–D and a small anti-parallel β-sheet (yellow arrows). Disulfides (yellow bonds) (Cys3-Cys127, Cys24-Cys65, and/or Cys46-Cys99) and cysteines are also shown. Center diagrams (two-side view images) show the structural comparison between WT-IL4 (blue) and C3T-IL4 (red) and WT-IL4 (blue) and C24T-IL4 (dark yellow). Right diagrams show backbone atoms root mean square deviations (RMSDs) for the ensembles of 10 lowest energy NMR structures of C3T-IL4 and C24T-IL4. Ribbon thickness is proportional to RMSD per residue. Images produced with visual molecular dynamics (VMD) and PyMOL.

C24T-IL4 were obtained at pH 5.0, 298 K (Figure 1 and Table S1). Statistics for the final NMR structure ensembles (10 lowest energy structures) are shown in Table 1. Quality assessment by PROCHECK-NMR indicated that 70.0% of the residues of C24T-IL4 have phi (ϕ) and psi (ψ) dihedral angles in the most favored regions of the Ramachandran plot, while 30.0% of the residues have phi and psi angles in allowed regions (equivalent resolution of 2.8 Å). In turn, the ensemble of structures of C3T-IL4 presented 71.8% of the residues with phi (ϕ) and psi (ψ) angles in the most favored regions of the Ramachandran plot, and 28.2% in allowed regions (equivalent resolution of 2.2 Å). Protein packing was estimated by the OSA algorithm, to obtain occluded surface packing values (OSPs). OSPs generally range from 0.0 to 0.7, with buried residues presenting OSPs around 0.6, and solvent-exposed residues around 0.20.[70] WT-IL4 is a well-packed protein with an average OSP of 0.300, followed by C3T-IL4 (average OSP 0.291) and C24T-IL4 (average OSP 0.277). These OSP values are in good agreement with the solvent accessible surface area (SASA) of C24T-IL4 and C3T-IL4 (Figure 3). Residues with higher OSPs (packed more tightly) are less solvent exposed than residues with lower OSPs, located at helices A and D and at the BC loop, in the case of C24T-IL4. Relatively to WT-IL4, C3T-IL4 showed a particular increase in SASA from residues 60 to 100 (located at helix C and BC loop).

TABLE 1 Structural statistics for the ensembles of NMR structures obtained for the IL-4 variants, 298 K.

Restraint	WT-IL4, pH 5.6	WT-IL4, pH 2.4	C3T-IL4, pH 5.0	C24T-IL4, pH 5.0
RMS deviations from ideal covalent geometry[a]				
Bonds (Å)	0.00375 ± 0.00049	0.00310 ± 0.00029	0.00320 ± 0.00066	0.00295 ± 0.00022
Angles (°)	0.493 ± 0.025	0.432 ± 0.025	0.423 ± 0.030	0.411 ± 0.014
Impropers (°)	0.415 ± 0.020	0.362 ± 0.036	0.382 ± 0.023	0.269 ± 0.013
RMS deviations from experimental restraints[a]				
NOE (Å) total	0.0852 ± 0.0036	0.064 ± 0.007	0.098 ± 0.009	0.092 ± 0.011
Intra-residue [(*i*, *j*), $i = j$]	0.061 ± 0.007	0.067 ± 0.010	0.069 ± 0.012	0.073 ± 0.008
Sequential [(*i*, *j*), $(i - j) = 1$]	0.071 ± 0.003	0.074 ± 0.014	0.085 ± 0.008	0.091 ± 0.013
Short-range [(*i*, *j*), $(2 \le (i - j) \le 4]$	0.127 ± 0.004	0.050 ± 0.019	0.113 ± 0.023	0.117 ± 0.039
Long-range [(*i*, *j*), $(i - j) \ge 5]$	0.068 ± 0.011	0.037 ± 0.010	0.211 ± 0.067	0.123 ± 0.071
Torsion angles (°)	0.689 ± 0.075	0.878 ± 0.421	0.634 ± 0.208[a]	0.649 ± 0.172
Hydrogen bonds (Å)	0.067 ± 0.006	0.087 ± 0.012	0.098 ± 0.012	0.057 ± 0.016
Dipolar (Hz)	0.990 ± 0.091	0.256 ± 0.099	0.620 ± 0.079	—
Coordinate precision (Å)[b]				
Backbone RMSD to mean res. 1–129	0.71 ± 0.18	1.68 ± 0.25	2.23 ± 0.91	4.22 ± 0.83
Backbone RMSD to mean res. 3–127	0.58 ± 0.09	1.50 ± 0.27	2.11 ± 0.96	4.15 ± 0.83
Helices[b]				
Helix A	0.31 ± 0.05	0.39 ± 0.10	0.78 ± 0.27	0.28 ± 0.05
Helix B	0.23 ± 0.06	0.50 ± 0.13	0.52 ± 0.12	0.74 ± 0.06
Helix C	0.28 ± 0.06	0.61 ± 0.23	0.74 ± 0.23	1.53 ± 0.06
Helix D	0.32 ± 0.14	0.35 ± 0.09	0.45 ± 0.19	0.64 ± 0.11
Structural parameters—interhelical angles[c]				
AB	44.4°	47.7°	52.2°	38.6°
CA	24.8° (−155.2°)	24.6 ° (−155.4°)	13.3° (−166.7°)	43.8° (−136.2°)
DA	34.7° (−145.3°)	25.7° (−154.3°)	76.5° (−103.5°)	53.4° (−126.6°)
CB	25.7° (−154.3°)	34.7° (−145.3°)	43.7° (−136.3°)	19.4° (−160.6°)
DB	34.0° (−146.0°)	24.4° (−155.6°)	48.3° (−131.7°)	37.8° (−142.2°)
CD	39.2°	26.7°	77.7°	18.4°
Structural quality—Ramachandran statistics[d]				
Most favored regions	81.4%	72.8%	71.8%	70.0%
Allowed regions	18.6%	26.7%	28.2%	30.0%

Abbreviations: IL-4, interleukin-4; NMR, nuclear magnetic resonance; RMSD, root mean square deviation; VMD, visual molecular dynamics.
[a]Parameters determined by XPLOR-NIH.
[b]Parameters determined with VMD.
[c]Parameters determined with Q-helix.
[d]Ramachandran statistics determined by PROCHECK-NMR.

TABLE 2 Regions of secondary structure of WT-IL4, C3T-IL4, and C24T-IL4.

Type of structure[a]		WT-IL4, pH 5.6	WT-IL4, pH 2.4	C3T-IL4	C24T-IL4
β-sheet		28–31/105–108	29–30/106–107	28–29/107–108	30/107
3_{10}-Helix		33–35	__	35–37	60–62
α-Helix	A	5–20	6–19	4–16	6–16
	B	41–60	41–58	41–58	41–57
	C	70–94	74–93	70–90	76–93
	D	109–125	109–125	109–127	109–124

Abbreviations: IL-4, interleukin-4; NMR, nuclear magnetic resonance.
[a]Secondary structure according to the STRIDE algorithm present in the lowest energy structures of the ensembles of NMR structures of each IL-4 variant.

When compared to WT-IL4, the NMR solution structure of C24T-IL4 exhibited a significant decrease in α-helical content (∼14%), associated with a reduction in length of all α-helices along with a lower propensity to form the short anti-parallel β-sheet (that connects the two loops AB and CD) (∼19% overall loss in secondary structure) (Figure 1 and Table 2). Major differences in root mean square deviations (RMSDs) (Figure 1) are located at terminal regions and protein loops (especially the long AB loop and short BC loop). Moreover, this mutant protein showed a less packed core in agreement with its lower conformational stability (C24T-IL4: $T_m = 73.5°C$, $\Delta G(H_2O)_{25°C} = 6.0 \pm 0.4$ Kcal.mol^{-1} WT-IL4: $T_m = 83.8°C$, $\Delta G(H_2O)_{25°C} = 9.3 \pm 0.4$ Kcal.mol^{-1}); and lower affinity for the alpha-chain IL-4 (IL-4α) receptor ($K_{d\,(C24T\text{-}IL4)} = 7.5 \times 10^{-11}$ M; $K_{d(WT\text{-}IL4)} = 5.6 \times 10^{-11}$ M) as reported previously.[34] Nonetheless, the presence of a lower content in secondary structure and more flexible loops, having the disulfide Cys3-Cys127 still present in C24T-IL4 seems to assure the relative orientation of helix A towards helices C and D, with interhelical angles (24.8° and 34.7°, Table 1), which are within the range of angles found among hematopoietic cytokines (20° to 40°), which is important for the high affinity interaction with the receptor alpha chain (first step of the two-step receptor interaction and heterodimerization).[35] Thus, absence of disulfide Cys24-Cys65 in C24T-IL4, primarily destabilizes the regions close to the missing disulfide, but also affects the overall structure of the four-helix bundle motif. Thus, losing this stabilizing disulfide linkage not only introduces more flexibility and disrupts helix-to-helix interactions (as stated by the loss of some long-range NOE restraints) but also interferes with the formation of the short β-sheet, and with the regular length of helices A, B, and C. Also, MD results on C24T-IL4 also pointed towards higher RMSDs, higher flexibility, looser packings, and loss of native interactions in all tested runs (Figure 5B).

Similarly, C3T-IL4 (lacking disulfide Cys3-Cys127) also formed a less packed protein core (Figure 3), with changes in secondary structure (shorter α-helices A, B, and C and lower propensity to form the anti-parallel β-sheet) and in the relative orientation of the α-helices (Figure 1), also affecting protein dynamics (Figure 2). Even if most of the more ordered core of WT-IL4 is still preserved in C3T-IL4 (average S^2 of 0.81 ± 0.17, 77% residues with $S^2 > 0.75$) the slower millisecond to microsecond time-scale dynamics, as well as the fast time-scale dynamics of the two proteins are different. As a result, a higher overall correlation time (τ_R) of 10.3 ± 0.1 ns was obtained for C3T-IL4. Thus, C3T-IL4 showed a much more flexible structure than WT-IL4, exhibiting particularly fast motions on the sub-nanosecond time scale, especially at the two protein termini, interhelical loops, beginning of helix A and end of helix D. In sum, C3T-IL4 showed altered backbone dynamics, loss of secondary structure (∼9%) (α-helices, 3_{10}-helix, and β-sheet), along with changes in interhelical angles, that led to looser and less packed protein conformations, especially affecting the helical core and loop regions, in agreement with a higher affinity for ANS, lower conformational stability (C3T-IL4: $T_m = 70.0°C$, $\Delta G(H_2O)_{25°C} = 4.3 \pm 0.5$ Kcal.mol^{-1}) and lower affinity for the alpha-chain IL-4 (IL-4α) receptor ($K_{d(C3T\text{-}IL4)} = 6.9 \times 10^{-11}$ M), as reported previously.[34] The absence of disulfide Cys3-Cys127, has not only affected the length of the α-helices involved in receptor interaction, but also their relative orientation and dynamics. The interhelical angle between helices A and C decreased from 24.8° (WT-IL4) to 13.3°; whereas between helices A and D increased from 34.7° (WT-IL4) to 76.5°. Therefore, these changes in structure, stability, and dynamics (next section) shown by C3T-IL4, certainly explain the decrease in biological activity shown by this mutant when interacting with the ectodomain of the alpha chain of the IL-4 receptor (IL-4BP).[34]

3.2 | Increased protein dynamics in the absence of Cys3-Cys127

Given the good dispersion of signals in the NMR spectra of C3T-IL4, it was possible to conduct a complete experimental study of protein dynamics (Figure S1). The mean values of 1.56 ± 0.12 s^{-1} for R_1, 10.83 ± 2.86 s^{-1} for R_2 and 0.69 ± 0.20 and 0.69 ± 0.18 obtained for the ^{1}H-^{15}N NOEs at 50.7 and 60.8 MHz, respectively, are within the range of values previously reported for WT-IL4, as well as, for other globular proteins.[57,58,71,72] When compared to WT-IL4, the relaxation rates of C3T-IL4 were decreased in the case of R_1 and increased in the case of R_2, producing an overall increase in the T_1/T_2 ratios (8.04 ± 2.29) (Figure 2). Thus, a higher overall rotational correlation time (τ_R) of 10.3 ± 0.1 ns was obtained for C3T-IL4, when compared to WT-IL4 ($\tau_R = 7.6$ ns).[38]

Order parameters (S^2) obtained from the free-model analysis (Table S2) are plotted against amino acid sequence in Figure 2C. An

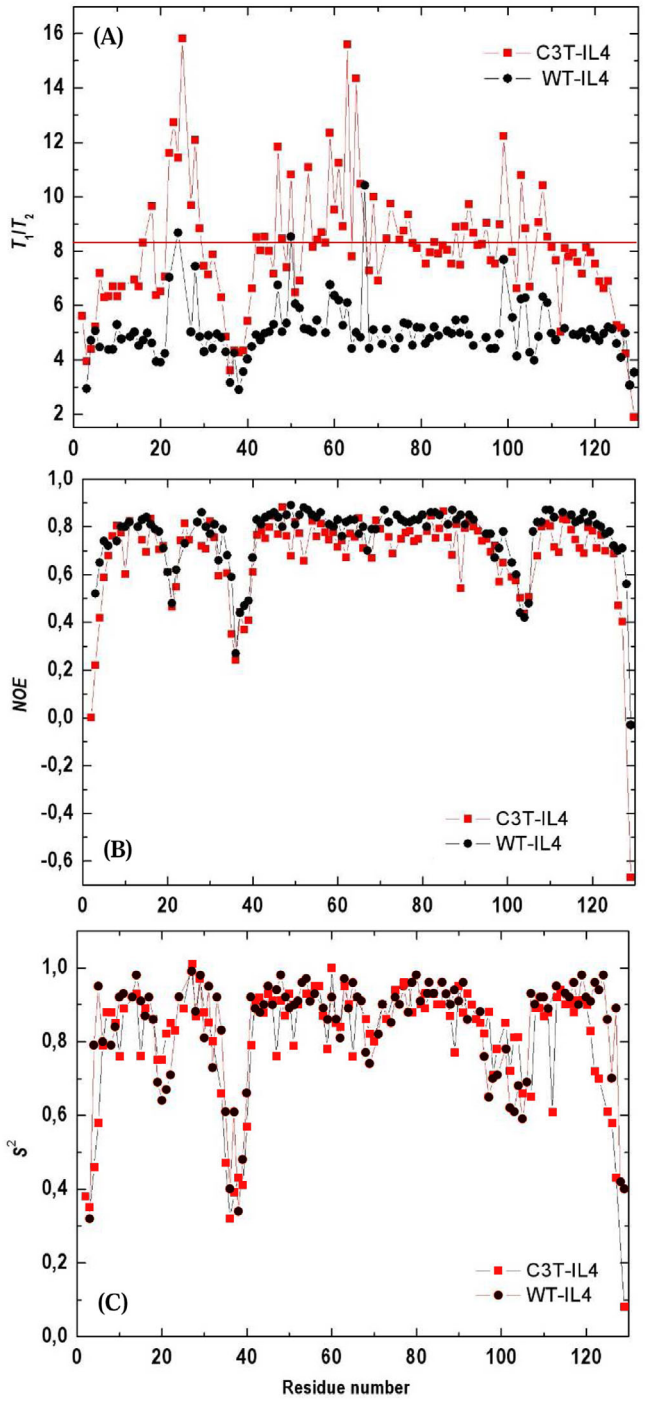

FIGURE 2 Comparison of the relaxation of C3T-IL4 with WT-IL4. (A) Plot of the observed T_1/T_2 ratio versus amino acid sequence. The horizontal line at 8.15 represents the average ratio used for the determination of the rotational correlation time (τ_R). (B) Plot of the observed ^{15}N NOE at 50.7 MHz versus amino acid sequence. (C) Plot of the calculated S^2 values versus amino acid sequence.

average order parameter (S^2) of 0.81 ± 0.17 was obtained for the backbone of C3T-IL4, with 77% of the residues having a S^2 greater than 0.75. Table S4 and Figure S2 show the distribution of the average S^2 according to secondary structure and RMSD. Most of the high S^2 values (S^2 > 0.75) are assigned to residues located in α-helices, with an average S^2 of 0.86 ± 0.09. Among the four helices, Helix A presented the lowest average S^2 of 0.79 ± 0.15 and helix C the highest S^2 of 0.90 ± 0.05, when compared to 0.81 for the whole protein. These order parameters are consistent with previously obtained data for WT-IL4, that showed a highest S^2 of 0.92 for helix C and a S^2 of 0.87 for the entire protein.[38] This indicates that the more ordered backbone, characteristic of the four-helix bundle of WT-IL4, was preserved to some extent. A relatively high average order parameter (S^2) of 0.86 ± 0.11 was also found for the β-sheet and for the short BC loop between helices B and C. Nonetheless, some residues showed significantly smaller S^2 (≤0.75), which are indicative of motions of larger amplitude. Akin to WT-IL4, C3T-IL4 presented low S^2 for residues located at the long AB and CD loops, beginning of helix A and end of helix D, as well as at both N- and C-termini (Figure 2C and Table S4). The long AB and CD loops showed mean S^2 smaller than those observed for the regions of regular secondary structure. In WT-IL4, these two loops run the length of the protein and interact through the short anti-parallel β-sheet, involving residues 28–31 from loop AB loop and residues 105–108 from loop CD. The β-sheet divides the AB loop in two segments of equal length (AB_1 and AB_2). Interestingly, the second half of the AB loop (AB_2), both in WT-IL4 and C3T-IL4, showed the lowest S^2 (0.51 ± 0.16), whereas loop AB_1 had the highest (0.85 ± 0.08). Low S^2 values for loops AB and CD probe the predominance of motions of significant amplitude on a time scale faster (τ_f and τ_S) than the overall correlation time (τ_R) (Figure S3).

Most residues whose relaxation data were analyzed using the extended form of the spectral density function[57,58] are located at both termini, beginning of helix A, end of helix D and AB and CD loops (Tables S2 and S4 and Figure S3B,D). The extended model included two rapid motions (τ_f and τ_S) faster than τ_R, one very fast motion ($\tau_f \leq 10$ ps) characterized by S_f^2 and a slower motion ($\tau_S \approx 1$ ns), but still faster than τ_R ($\tau_R > \tau_s > \tau_f$) characterized by S_f^2. Interestingly, most of the residues in the aforementioned regions showed motions on the 0.5–2.5 ns (τ_S) time scale. On the other hand, most of the 19 residues, that needed the 2-parameter (S^2 and τ_e) fitting model, were in the AB and CD loops and showed τ_e values on the 25 to 150 ps time scale (Figure S3). In WT-IL4, only eight residues required the two-parameter fitting model for S^2 and τ_e and presented a τ_e in the range of 25–100 ps.[38] Thus, overall, the order parameters determined for C3T-IL4 suggest that, although most of the more ordered backbone of WT-IL4 is maintained, there are some amplitude fluctuations on a fast time scale, especially at the beginning of helix A, end of helix D, AB and CD loops and protein termini. Thus, even if C3T-IL4 maintains most of the native-like elements of secondary structure, it has an altered and substantially less ordered tertiary structure.

The S^2 values are sensitive to motions occurring on a time scale faster than τ_R, while T_2 values are influenced by chemical exchange effects and by events occurring on slower time scales. Conformational fluctuations taking place on the millisecond time scale, for instance, may cause a decrease in T_2. A considerable decrease in T_2 was detected for residues with a high T_1/T_2 ratio. When compared to WT-IL4, only 18 residues in C3T-IL4, with T_1/T_2 ratios significantly above

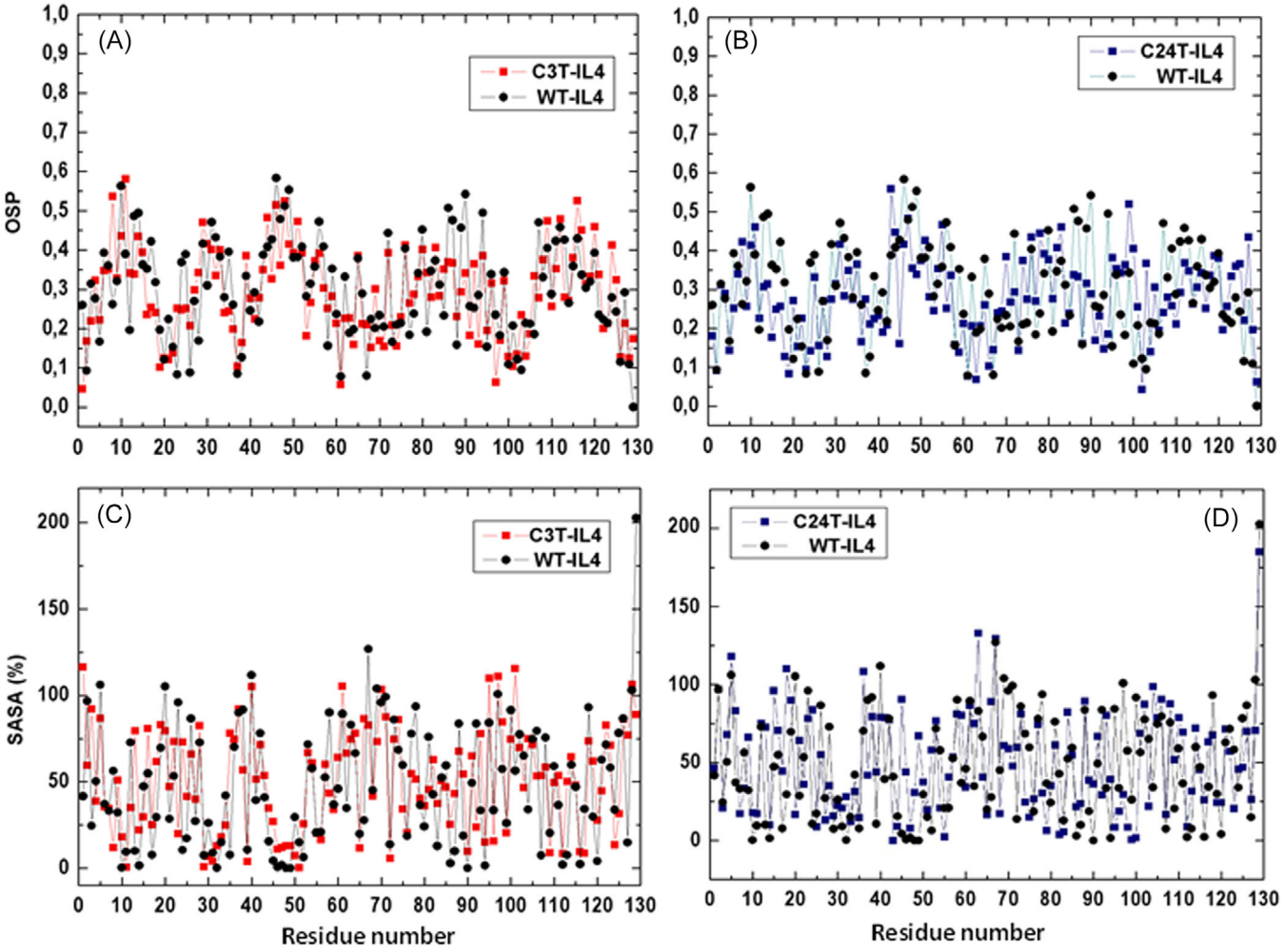

FIGURE 3 Comparison of occluded surface packing (OSP) and solvent accessible surface areas (SASA) of C3T-IL4 and C24T-IL4, with WT-IL4. OSP and SASA of the NMR solution structures of the two interleukin-4 (IL-4) variants missing one native disulfide. OSP values per residue of C3T-IL4 (A) and C24T-IL4 (B), compared to WT-IL4. SASA values of C3T-IL4 (C) and C24T-IL4 (D) compared to WT-IL4.

8.84, required the use of an exchange contribution (Δ_{ex}) of more than 0.5 Hz (Table S2) in the fitting model. Chemical exchange contributions (Δ_{ex} values) are plotted as a function of the amino acid sequence in Figure S3A. High T_1/T_2 ratios suggesting extensive exchange contributions to T_2 (with Δ_{ex} values greater than 1.5 s^{-1}) are observed from residues 20 to 30 and from residues 55 to 65. Most residues in C3T-IL4 presenting an exchange contribution to T_2 are essentially located in helices B and C and in the protein loops, i.e., residues located close to the β-sheet and to the two native disulfides (Cys24-Cys65 and Cys46-Cys99) which are still present in C3T-IL4. Disulfide bonds and β-sheet hydrogen bonds (V29-D31 and S107-L109) are likely to slow down the motions of these residues, bringing them into the time-scale range where exchange effects may be observed. Some of these residues (22, 23, 25, 28, 63, 65, 66, 98, 99, 103, 104, and 108) are in loop regions, where large amplitude sub-nanosecond motions have also been detected. The largest Δ_{ex} values were found in the AB loop, where residue Cys24 from disulfide Cys24-Cys65 is located. Thus, the observation of some weak cross-peaks for these residues in the 2D ^{1}H-^{15}N HSQC spectrum is likely due to shorter ^{15}N T_2 values resulting from chemical exchange broadening. Therefore, although most of the more ordered backbone of WT-IL4 is preserved in C3T-IL4, the slower millisecond to microsecond time-scale dynamics as well as the fast time-scale dynamics of the two proteins are different. In C3T-IL4, the absence of Cys3-Cys127 disulfide that brings together the two protein termini, produces a much more flexible and dynamic protein, that exhibits particularly fast motions on the sub-nanosecond time scale. On the other hand, the altered dynamics detected at the beginning of helix A and end of helix D is affecting helices B and C and disrupting the high structural integrity that characterizes the four-helix bundle of WT-IL4.

In addition, backbone RMSFs obtained from the MD simulations correlate well with the RMSFs obtained from the NMR-derived structures (Figure 5, $r = 0.913$). Even if during the MDs, C3T-IL4 presented no major changes in secondary structure relatively to WT-IL4 (Figure 6), the mutant protein showed markedly increased RMSFs at the two termini, where disulfide Cys3-Cys127 is absent, as well as increased RMSFs in the AB and CD loops, although with differences not as large as those obtained by NMR. Comparison of the main chain

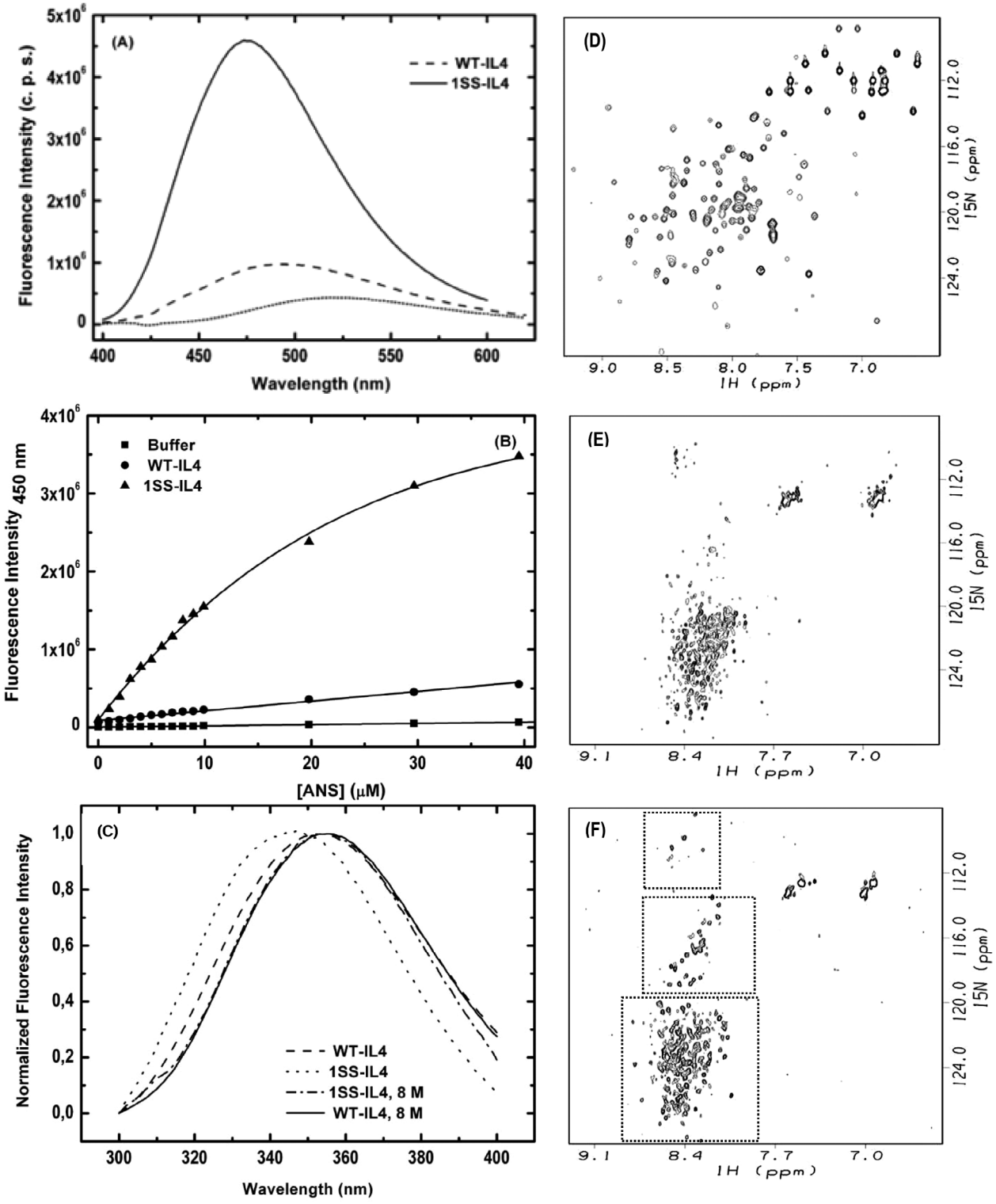

FIGURE 4 Spectroscopic characterization of the disulfide mutant 1SS-IL4. (A) Fluorescence emission spectra of free (dotted line) and bound 8-anilino-1-naphtalenesulphonic acid (ANS) to WT-IL4 (dashed line) and 1SS-IL4 (solid line). (B) ANS titration curves in 10 mM NaPi, pH 6.0, 298 K. Lines are nonlinear least square fits of an equation based on a simple protein-ligand binding model to the experimental data. (C) Intrinsic fluorescence emission spectra of WT-IL4 and 1SS-IL4, in 10 mM NaPi, pH 6.0, 298 K. (D) [^{1}H-^{15}N]-HSQC NMR spectrum of WT-IL4 and of 1SS-IL4 in the absence (E) and presence (F) of 8.0 M urea in 25 mM d_3-acetic acid, 90% H_2O, 10% D_2O, pH 5.1, 298 K. (F) Three clusters of residues identified according to typical chemical shift distributions in the ^{15}N dimension, namely Gly <112.0 ppm, Thr and Ser residues [114.0–120.0] ppm, and remaining residues [120.0–126.0] ppm.

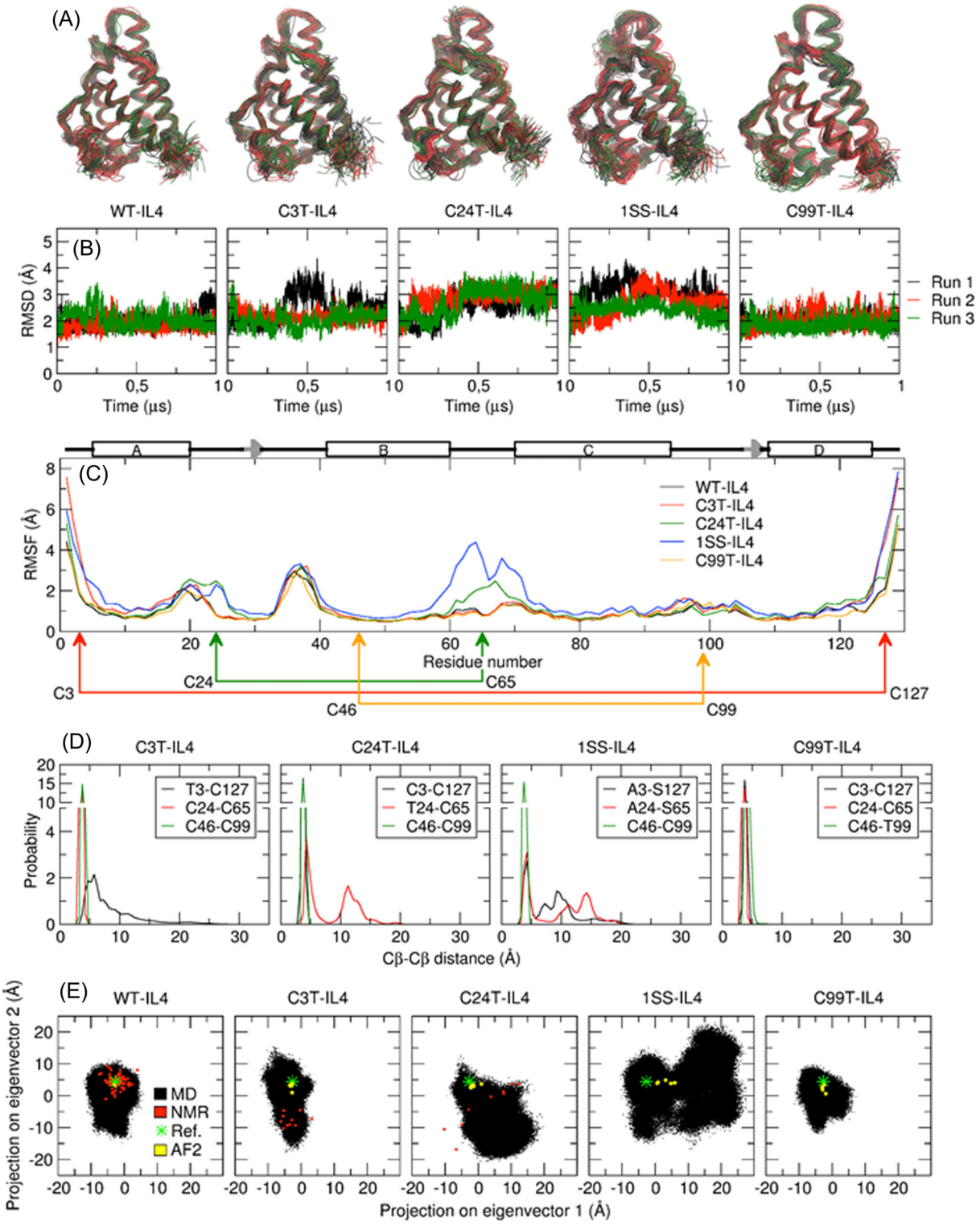

FIGURE 5 Molecular dynamics (MD) simulations of WT-IL4, C3T-IL4, C24T-IL4, C99T-IL4 and 1SS-IL4. (A) Overall fold and secondary of the interleukin-4 (IL-4) variants (snapshots taken at 100-ns intervals). Translation and rotation movements were removed. (B) Root mean square deviation (RMSD) of Cα atoms during MD simulations of IL4 variants. RMSD was computed by using the coordinates of PDB 2B8U as reference. Three 1-μs replicas are shown in different colors for each variant. (C) Root mean square fluctuations (RMSF) of Cα atomic positions along the trajectory for each IL-4 variant replica. Structures before energy minimization were used as reference. Disulfides are identified below the plot. Secondary structure elements in the starting structures (rectangle—alpha-helix and arrow—beta-sheet) are identified above the plot. (D) Cβ-Cβ distance probability of atoms of residue pairs for each IL-4 variant. (E) Projection of the trajectories (MD) along the first two eigenvectors obtained from the principal component analysis (PCA), with the structures derived from NMR and the models predicted by AlphaFold2 (AF2). Ref is the departing structure for the MD simulations.

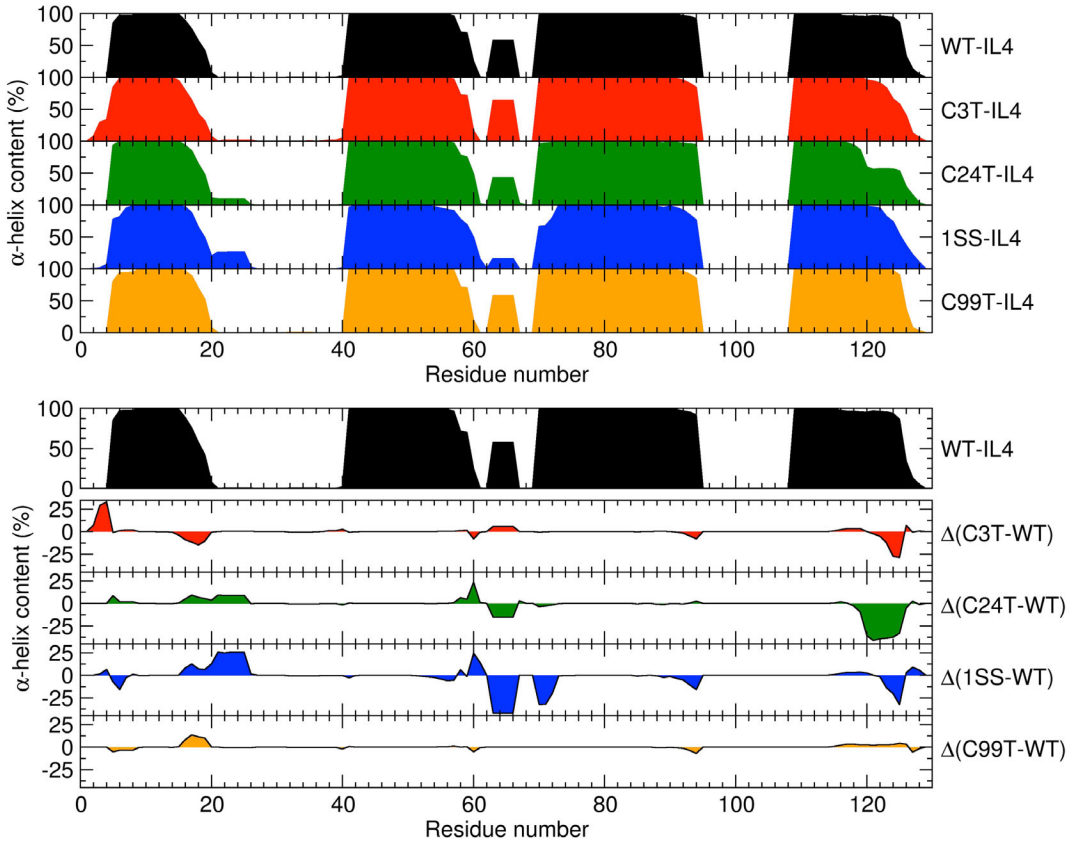

FIGURE 6 Average alpha-helix content of each IL4 variant (molecular dynamics [MD] data). (A) Average alpha-helix content presented by the five interleukin-4 (IL-4) variants along the 3 MD runs of 1000 ns. (B) Average alpha-helix content differences of the IL-4 mutants relatively to WT-IL4, exhibiting regions of major loss of alpha helix.

fluctuations observed in the MDs with the NMR relaxation data showed a moderate correlation between the two ($r = 0.5824$). In fact, fluctuations at the protein termini and end of loop AB were higher in the MDs and generally lower in the remaining regions, when compared to the ensemble of NMR-derived structures, which suggests that while the NMR data are reporting average properties of an ensemble of conformations interconverting rapidly, MD is sensing point singularities, during very short periods (1 μs). Given the higher RMSFs of C3T-IL4 relatively to WT-IL4, a distance analysis of marker atoms was performed (Figure 5). C_β-C_β distances for the residue pair T3/C127 varied widely as function of time in C3T-IL4 and between replicas, reaching peak values as high as 30 Å (highlighting the need of running multiple MD replicas, which in this case, agree with the NMR-derived structures, as seen by the principal component analysis (PCA) analysis depicted in Figure 5E). Order parameters (S^2) were also computed from the MD trajectories (Figure S4A). In agreement with the NMR data, regions of higher mobility were also detected at the protein termini and loop regions in all MD runs. Loops AB and CD proved to be particularly flexible, except for the cysteine residues Cys24 and Cys99 involved in the native disulfides still present (i.e., Cys-24-Cys65 and Cys46-Cys99), which clearly restrict the dynamic behavior of the protein in these regions and also stabilize the short β-sheet (Thr28-Val29/Ser107-Thr108). Interestingly, when comparing the two sets of order parameters (S^2 MD and S^2 NMR values) (Figure S4C), the MD trajectories presented, on average, lower S^2 values, predicting global higher flexibility, especially in the first half of the AB loop (from residue 20 to 24) and CD loop. This reinforces the role of the native disulfides present (Cys-24-Cys65 and Cys46-Cys99) in diminishing loop flexibility and contributing to the overall dynamics of the C3T-IL4 protein in solution, as stated by the NMR data.

3.3 | IL-4 forms a quasi-molten globule in the absence of Cys24-Cys65 and Cys3-Cys127

Deprived of both disulfides Cys3-Cys127 and Cys24-Cys65 (mutant 1SS-IL4), IL-4 formed a collapsed partially unfolded conformation as judged by the marked changes in the NMR spectra (Figure 4E), decrease in secondary structure (Figure S5), loss of tertiary interactions and higher ANS-binding tendency (Figure 4A). Far-UV CD experiments were performed to evaluate the secondary structure of WT-IL4 and 1SS-IL4 (Figure S5), as we reported previously for C3T-IL4 and C24T-IL4.[34] A significant decrease in ellipticity as evidenced by a decrease in signal intensity at 193, 208 and 222 nm, was detected in the spectrum of 1SS-IL4, pointing towards a considerable loss in

secondary structure. CONTIN estimated an α-helical content of only 29% for 1SS-IL4 (~38 residues in α-helix), that could either result from the loss of nearly two helices, or to the shortening to a half of all four helices of the bundle.

ANS binding experiments were also carried out on WT-IL4 and 1SS-IL4 (Figure 4A), as we reported for C3T-IL4 and C24T-IL4.[34] ANS titration of WT-IL4 led to a 30 nm blue-shift in the emission maximum (λ_{max}) of the probe from 520 nm (free probe) to 490 nm (bound). In the case of 1SS-IL4, a larger blue-shift of λ_{max} to 474 nm (46 nm) was observed, along with a 5-fold increase in fluorescence intensity (Figure 4A). Apparent dissociation constants (K_d) were determined by non-linear least square fitting to the titration data (Figure 4B), assuming a simple protein-ligand one-site binding model (Materials and Methods). Dissociation constants (K_d) of 35.0 μM and 8.8 μM were obtained for WT-IL4 and 1SS-IL4, respectively. Hence, according to the λ_{max} and K_d values obtained, 1SS-IL4 presented higher tendency than WT-IL4 to bind to the hydrophobic probe. This indicates that 1SS-IL4 presents solvent-accessible non-polar patches, a typical feature shown by partially folded or "molten globule" conformations.

In addition, isothermal urea-induced unfolding experiments monitored by fluorescence spectroscopy (Figure S5) were also conducted for WT-IL4 and 1SS-IL4. Fluorescence emission spectra of all IL-4 variants are essentially dominated by the fluorescence of the single tryptophan residue (Trp 91, ~50% solvent accessibility in WT-IL4 by NACCESS), located at the end of helix C.[34] At pH 7.5, WT-IL4 exhibited a fluorescence emission maximum (λ_{max}) at 350 nm, that suffered a blue-shift as the pH decreased. Nonetheless, upon chemical unfolding, an increase in fluorescence intensity along with a red shift of λ_{max} to 356 nm was always observed (Table S5). Figure S5C shows the equilibrium unfolding curves of WT-IL4, at different pH values, 298 K. All unfolding curves fitted well to a two-state model (Materials and Methods). The free energy change in the absence of denaturant ΔG (H_2O)), and the cooperativity of the unfolding transitions (m) were obtained as fitting parameters for all curves (Figure S5D). The conformational stability ($\Delta G(H_2O)$) of WT-IL4 gradually decreased, as the pH decreased, with values in the range of typical values for small monomeric proteins (5 to 15 kcal.mol^{-1}).[60] As the pH decreased, unfolding transitions became less cooperative, suggesting that, even in the absence of denaturant, an altered conformation was adopted by WT-IL4, supporting the 3-nm blue-shift in λ_{max}, observed when the pH was lowered from 7.5 to 3.0. In turn, the fluorescence emission spectra of 1SS-IL4 (simultaneously lacking both native disulfides Cys3-Cys127 and Cys24-Cys65) in the absence of urea, showed a 6-nm blue-shift in the λ_{max}, relatively to WT-IL4. This blue shift suggests a more hydrophobic or less solvent exposed environment around Trp91. Nonetheless, in the presence of 8.0 M urea, both proteins suffered unfolding and both emission maxima showed a red shift to 356 nm (Figure 4C). As the concentration of denaturant increased, a sharp increase in fluorescence intensity was observed for WT-IL4, while for 1SS-IL4 a gradual increase in intensity with no well-defined pre- or post-transition regions was obtained (data not shown). The absence of a cooperative unfolding transition may indicate that the protein mutant 1SS-IL4 is an ensemble of partially unfolded collapsed conformations (a "molten globule"), even under non-denaturing conditions.

Molten globule conformations are known to give poorly resolved and broadened ^{1}H NMR signals and weak or no signals in [^{1}H-^{15}N] HSQC spectra, given the conformational fluctuations of the polypeptide chain on a millisecond to microsecond time scale.[76,77] Thus, to evaluate the structural changes presented by 1SS-IL4, heteronuclear 2D [^{1}H-^{15}N]-HSQC experiments were carried out with uniformly ^{15}N-labeled 1SS-IL4 in the absence and presence of 8.0 M urea (Figure 4E,F). The 2D [^{1}H-^{15}N] HSQC spectrum of WT-IL4 (Figure 4D) is typical of a well folded protein, as judged by the linewidth of the NMR signals and dispersion of amide resonances. In turn, 1SS-IL4 showed broad peaks, low chemical shift dispersion, and large chemical shift changes (Figure 4E). The increased number of resonances observed for 1SS-IL4 points towards the coexistence of multiple conformations, while the low signal dispersion in the ^{1}H dimension suggests the presence of unfolded or partially folded conformations (Figure 4E). Thus, an ensemble of partially folded or molten globule conformations may be in slow exchange in solution and be responsible for the large number of poorly resolved NMR peaks. In turn, the [^{1}H-^{15}N]-HSQC of 1SS-IL4 in the presence of 8.0 M urea (Figure 4F) presented sharper, fewer, and less dispersed peaks. Nonetheless, the number of resonances identified in the HSQC spectrum is close to the expected number from the total number and type of residues in 1SS-IL4 (3 Gly, 24 Ser, and Thr and ~ 100 others), revealing that different protein conformations are being adopted in the presence and absence of 8.0 M urea. Thus, in the absence of both disulfides (Cys3-Cys127 and Cys24-Cys65), 1SS-IL4 is forming a partially folded protein, fluctuating between conformations, as also stated by all runs of the MD data (Figure 5B).

Poor NMR signal dispersion, low stability, tendency to aggregate, enhanced ANS binding, variable amounts of native-like secondary structure and considerable loss of tertiary contacts are the hallmarks of "molten globule" states.[73–77] Molten globules can be found under a variety of conditions, including at low pH, upon disulfide bond reduction, removal of cofactors, low ionic strength, or at low concentrations of chemical denaturants. Several proteins form molten globules in the absence of their native disulfides[78,79] with α-lactalbumin (α-LA) being the best studied case.[80,81] Thus, this "molten globule" state of 1SS-IL4, may represent an early intermediate in the folding process of IL-4. Hence, if disulfide Cys46-Cys99, as previously shown, is essential for folding IL-4,[35] the simultaneous deletion of both disulfides Cys3-Cys127 and Cys24-Cys65 is also deeply affecting structure acquisition, dynamics, and the overall conformational stability of IL-4.

3.4 | Molecular dynamics simulations

Three independent MD simulation replicas of 1000 nanoseconds each were obtained for WT-IL4, C3T-IL4, C24T-IL4, C99T-IL4 and 1SS-IL4. MD simulations of WT-IL-4 started from the crystal structure (1.8 Å resolution) determined by Müller and collaborators (2B8U.pdb) (and not from the RDC-refined NMR structures, to avoid bias). In turn,

C3T-IL4, C24T-IL4, C99T-IL4 and 1SS-IL4 simulations started from a model chain produced in silico by mutation of the structure of WT-IL-4 (2B8U.pdb). All mutated cysteine residues are in loop regions of IL-4 to minimize the impact of the mutation per se in the secondary structure of the protein. Minimum periodic distances between atoms and their periodic images were calculated. The closest approximation distance was of 19.09 Å for the 2nd run of 1SS-IL4 (large distances). Thus, interactions between proteins and their periodic image were considered negligible. Figure 5B shows the RMSD of backbone Cα atoms during the MD runs. Except for some transient increases in runs 1 and 3, the RMSD for WT-IL4 remained close to 2 Å (1.9 ± 0.3 Å). In turn, C3T-IL4 showed a quite different behavior among the three replicas, with run 1 reaching higher RMSDs than other runs (average RMSD of 2.5 ± 0.5 Å). Mutant proteins C24T-IL4 and 1SS-IL4 showed the greatest altered dynamics, with respect to WT-IL4 (Figure 5). Interestingly, C99T-IL4 presented a similar behavior to WT-IL4 throughout all simulations. This may be indicating that, although the Cys46-Cys99 linkage has shown to be essential for the folding process of IL-4[35] (in particular, formation of helix B and the four-helix bundle), this disulfide may be less important in maintaining the structure and dynamics of the protein in solution.

A root mean square fluctuation (RMSF) analysis was also carried out to identify the regions contributing for the RMSDs differences (Figure 5C). Regions of greater fluctuation for WT-IL4 were at the two protein termini, AB loop and end of helix A, whereas the BC and CD loops and end of helix D showed intermediate mobility. C3T-IL4 and 1SS-IL4 showed increased fluctuations near the N- and C-terminal regions, relatively to WT-IL4 and C24T-IL4, evidencing the effect of removing disulfide Cys3-Cys127. On the other hand, C24T-IL4 and 1SS-IL4 showed higher RMSFs near residues 24 and 65 than WT-IL4 and C3T-IL4, also in agreement with the removal of disulfide Cys24-CysC65, with 1SS-IL4 showing particularly high RMSDs for the segment 60–70 (BC loop). Interatomic distances between Cβ atoms of selected pairs of residues, showed similar distances for Cys24-Cys65 and Cys3-Cys127, for C3T-IL4 and C24T-IL4, respectively, given the presence of the remaining native disulfides. Conversely, interatomic distances between pairs of residues, where disulfides are absent, fluctuated considerably, and were as high as 20 Å. Secondary structure analyses for each run of each IL4 variant were also conducted. Figure 6 shows a "consensus" analysis of α-helix content over the three trajectories of each IL4 variant. For each time frame, a score of 1 was attributed to protein residues in α-helix and a score of 0, if otherwise. Resulting binary matrices of each run were summed for each variant. These results are in agreement with the NMR data presented herein. All disulfide mutants showed lower α-helical contents, especially C24T-IL4 and 1SS-IL4. A PCA was also conducted on a concatenated trajectory of Cα atoms obtained from the fifteen MD trajectories (5 variants x 3 replicas). Figure 5E shows the projection of the trajectories of each IL-4 variant along the first two eigenvectors derived from the concatenated trajectory of Cα atoms. Once again, all runs of all three IL-4 disulfide mutants showed higher variability, dynamics, and structure flexibility, than WT-IL4, particularly mutants C24T-IL4 and 1SS-IL4, also corroborating the NMR data obtained herein.

Although the AlphaFold algorithm[69] shows remarkable accuracy in predicting native state protein 3D structures, it is known to be limited in predicting protein conformational changes upon single-site mutations. Despite this, we have applied AlphaFold2 to the primary sequence of the IL-4 variants studied herein. Given the resemblance of all protein sequences and the similar polarities between cysteine and the replaced residues (threonine and serine), no major differences were found by the algorithm, that would justify all the differences observed experimentally, when the protein misses either one (C3T-IL4, C24-IL4, C99T-IL4), two (1SS-IL4) or even three (0SS-IL4) of its native disulfides (Figure S6). Even though, the five models produced by the algorithm predicted more differences and greater structural destabilization in the case of the 1SS-IL4 variant, as also seen by the spectroscopic (NMR, far-UV CD, fluorescence) and MD data (*see* PCA, Figure 5E).

Protein engineering approaches (experimental and computational) focused on structure-guided stabilization of metastable regions, rather than on the receptor binding interfaces, have been successfully applied to affinity maturation studies on IL-2.[82,83] Thermostable "IL-2 superkine" variants with increased affinity for the IL-2Rβ receptor were produced either by stabilizing protein loops, improving protein packing and/or redesigning the hydrophobic core of the protein. Given the interest in also developing an "IL-4 superkine" as an immunotherapeutic agent that could be tuned to different cell types by modulating receptor affinity, several in silico affinity maturation studies can be carried out towards directed-evolution and global structure optimization of loop mobility, helical-to-helix packing, and protein core interactions. Recent efforts have been also developed by computationally designing a hyperstable IL-4 cytokine mimetics (Neo-4) that is able to signal exclusively via the Type I IL-4 receptor complex and that can be adsorbed to 3D-printed scaffolds.[84] Moreover, fine-tuning approaches based the RoseTTAFold structure prediction network (RFdiffusion), for instance, can be used to improve protein stability and optimize protein–protein interactions toward the development of more sensitive detection methods and more selective therapeutics.[85]

3.5 | Protonation of His76 triggers loss of protein structure and protein stability

pH dependencies also regulate protein structure, stability, and folding/unfolding mechanisms.[86–90] WT-IL4 also showed increased dynamics[38] and loss of secondary structure (Figure S7 and Table 2) at low pH (2.4). Ensembles of the RDC-refined high resolution NMR structures of IL-4 at pH 5.6 (PBD ID 8CH7), and at pH 2.4 (PBD ID 8CGF), are shown in Figure S7. At pH 2.4, WT-IL4 presents shorter A, B and C helices, a shorter antiparallel β-sheet, and lower propensity to form the 3_{10}-helix (Table 2). Moreover, the relative orientation of the helices changed, adopting a more parallel positioning in relation to the main axis of the four-helix bundle. The number of residues present in regular secondary structure at pH 2.4, detected by NMR, agrees

well with the reduction in secondary structure detected by CD (Figure S5). Apparent pK_as for all five His residues were also determined by NMR-monitored pH titrations. His1, His58, and His59 presented pK_as close to reference values (6.04), while His74 and His76 presented lower pK_as, especially His76 (3.94 ± 0.04; Table S6 and Figure S8). Thus, as the pH diminished, protonation of His76 triggered the loss of the first α-helical turn of Helix C (Figure S9), partially destabilizing the structure of the four-helix bundle and the overall conformational stability of the protein, as supported by the far-UV CD, ANS-binding, and NMR data. In addition, differences in protein structure and packing were accompanied by changes in conformational stability ($\Delta G(H_2O)$). As the pH decreases from pH 7.5 to 3.0, $\Delta G(H_2O)$ reduced 0.6 kcal.mol^{-1} per unit of pH (Table S5). For proteins undergoing a two-state unfolding mechanism, *m* values are thought to be proportional to the amount of surface area newly exposed to solvent.[60] Thus, an altered conformation adopted by IL-4, (at pH <5.0), even in the absence of denaturant, is also in agreement with a decrease in the solvent accessibility of Trp 91, as proven by the blue-shift of the fluorescence emission maximum from 350 to 347 nm.

In sum, at low pH, WT-IL4 loses various tertiary contacts and shows a decrease in secondary structure. The antiparallel β-sheet between the two long loops AB and CD becomes one residue shorter. The tendency to form a 3_{10}-helix in the AB loop is diminished, and the overall helical content is reduced in about 9%. Differences in structure and packing of WT-IL4, at pH 2.4, were also accompanied by changes on the conformational stability of the protein, that presented a reduction of 2.8 kcal.mol^{-1}, when the pH is lowered from 7.5 to 3.0. Thus, although WT-IL4 has shown to be a well-structured protein, the protein becomes structurally altered and conformationally less stable as the pH goes below 4.0.

4 | CONCLUSION

Compared to other cytokines, beyond its well-defined four-helix bundle, IL-4 presents extra unique features: a short antiparallel β-sheet (common among short-chain cytokines), tendency to form a supplementary 3_{10}-helix (common among long-chain cytokines) and three disulfide bonds (which are generally only one or two in number and local in topology (<50 residue apart) in the hematopoietic protein family). These structural elements stand out IL-4 as a special cytokine, with high structural and conformational stability, even at very low pH (2.4). Each native disulfide (Cys3-Cys127, Cys24-Cys65, and Cys46-Cys99) has proved to be extremely relevant for maintaining the structure, packing, dynamics, and overall stability of the protein in solution.

AUTHOR CONTRIBUTIONS

Daniela C. Vaz: Conceptualization; investigation; formal analysis; writing – review and editing. **J. Rui Rodrigues:** Investigation; software; formal analysis. **Nuno Loureiro-Ferreira:** Software. **Thomas D. Müller:** Conceptualization; funding acquisition; resources. **Walter Sebald:** Conceptualization; funding acquisition; resources. **Christina Redfield:** Investigation; funding acquisition; methodology; software; formal analysis; resources. **Rui M. M. Brito:** Conceptualization; funding acquisition; project administration; resources.

ACKNOWLEDGMENTS

The authors thank the financial support by FEDER/COMPETE (RECI/QEQ-QFI/0168/2012 and CENTRO-07-CT62-FEDER-002012), by Fundação para a Ciência e Tecnologia (FCT) (UIDB/00313/2020 and UIDP/00313/2020), and by Rede Nacional de Ressonância Magnética Nuclear (RNRMN/PTNMR). J. Rui Rodrigues was financially supported by LA/P/0045/2020 (ALiCE), UIDB/50020/2020 and UIDP/50020/2020 (LSRE-LCM), funded by national funds through FCT/MCTES (PIDDAC).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Daniela C. Vaz https://orcid.org/0000-0001-7562-4676
J. Rui Rodrigues https://orcid.org/0000-0002-9756-1124

REFERENCES

1. Creighton TE. Protein folding coupled to disulphide bond formation. *Biol Chem*. 1997;378:731-744.
2. Thangudu RR, Manoharan M, Srinivasan N, Cadet F, Sowdhamini R, Offmann B. Analysis on conservation of disulphide bonds and their structural features in homologous protein domain families. *BMC Struct Biol*. 2008;8:55.
3. Thornton JM. Disulphide bridges in globular proteins. *J Mol Biol*. 1981;151:261-287.
4. Bönisch H, Schmidt CL, Schäfer G, Ladenstein R. The structure of the soluble domain of an archaeal Rieske iron-sulfur protein at 1.1 A resolution. *J Mol Biol*. 2002;319:791-805.
5. Ladenstein R, Ren B. Protein disulfides and protein disulfide oxidoreductases in hyperthermophiles. *FEBS J*. 2006;273:4170-4185.
6. Ma W, Lu Y, Zuo Y, Wang C, Liu J. Effects of removing a highly conserved disulfide bond in ubiquitin-associated domain of human HOIP on biochemical characteristics. *Protein Expr Purif*. 2022;190:106005.
7. Pace CN, Grimsley GR, Thomson JA, Barnett BJ. Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *J Biol Chem*. 1998;263:11820-11825.
8. Toth EA, Worby C, Dixon JE, Goedken ER, Marqusee S, Yeates TO. The crystal structure of adenylosuccinate lyase from Pyrobaculum aerophilum reveals an intracellular protein with three disulfide bonds. *J Mol Biol*. 2000;301:433-450.
9. Zavodszky M, Chen CW, Huang JK, Zolkiewski M, Wen L, Krishnamoorthi R. Disulfide bond effects on protein stability: designed variants of *Cucurbita maxima* trypsin inhibitor-V. *Protein Sci*. 2001;10:149-160.
10. Arolas JL, Aviles FX, Chang JY, Ventura S. Folding of small disulfide-rich proteins: clarifying the puzzle. *Trends Biochem Sci*. 2006;31:292-301.
11. Bellei M, Bortolotti CA, Di Rocco G, et al. The influence of the Cys46/Cys55 disulfide bond on the redox and spectroscopic properties of human neuroglobin. *J Inorg Biochem*. 2018;178:70-86.

12. Betz SF. Disulfide bonds and the stability of globular proteins. *Protein Sci*. 1993;2:1551-1558.
13. Colombo A, Ribotta PD, León AE. Differential scanning calorimetry (DSC) studies on the thermal properties of peanut proteins. *J Agric Food Chem*. 2010;58:4434-4439.
14. Deplazes E, Chin YK, King GF, Mancera RL. The unusual conformation of cross-strand disulfide bonds is critical to the stability of β-hairpin peptides. *Proteins*. 2020;88(3):485-502.
15. Doig AJ, Williams DH. Is the hydrophobic effect stabilizing or destabilizing in proteins? The contribution of disulphide bonds to protein stability. *J Mol Biol*. 1991;217:389-398.
16. Kuroki R, Inaka K, Taniyama Y, et al. Enthalpic destabilization of a mutant human lysozyme lacking a disulfide bridge between cysteine-77 and cysteine-95. *Biochemistry*. 1992;31:8323-8328.
17. Matsuo K, Watanabe H, Tate S, Tachibana H, Gekko K. Comprehensive secondary-structure analysis of disulfide variants of lysozyme by synchrotron-radiation vacuum-ultraviolet circular dichroism. *Proteins*. 2009;77(1):191-201.
18. Schellman JA. The thermodynamic stability of proteins. *Annu Rev Biophys Biophys Chem*. 1987;16:115-137.
19. Wedemeyer WJ, Welker E, Narayan M, Scheraga HA. Disulfide bonds and protein folding. *Biochemistry*. 2000;39:4207-4216.
20. Rudenko G, Henry L, Henderson K, Ichtchenko K, Brown MS, Goldstein JL. Deisenhofer J structure of the LDL receptor extracellular domain at endosomal pH. *Science*. 2002;298:2353-2358.
21. Petersen MT, Jonson PH, Petersen SB. Amino acid neighbours and detailed conformational analysis of cysteines in proteins. *Protein Eng*. 1999;12:535-548.
22. Jessop CE, Chakravarthi S, Watkins RH, Bulleid NJ. Oxidative protein folding in the mammalian endoplasmic reticulum. *Biochem Soc Trans*. 2004;32:655-658.
23. Heras B, Kurz M, Shouldice SR, Martin JL. The name's bond.....disulfide bond. *Curr Opin Struct Biol*. 2007;17:691-698.
24. Hill EE, Morea V, Chothia C. Sequence conservation in families whose members have little or no sequence similarity: the four-helical cytokines and cytochromes. *J Mol Biol*. 2002;322:205-233.
25. Rozwarski DA, Gronenborn AM, Clore GM, et al. Structural comparisons among the short-chain helical cytokines. *Structure*. 1994;2:159-173.
26. Kraich M, Klein M, Patino E, et al. A modular interface of IL-4 allows for scalable affinity without affecting specificity for the IL-4 receptor. *BMC Biol*. 2006;4:13.
27. Powers R, Garrett DS, March CJ, Frieden EA, Gronenborn AM, Clore GM. The high-resolution, three-dimensional solution structure of human interleukin-4 determined by multidimensional heteronuclear magnetic resonance spectroscopy. *Biochemistry*. 1993;32:6744-6762.
28. Redfield C, Smith LJ, Boyd J, et al. Analysis of the solution structure of human interleukin-4 determined by heteronuclear three-dimensional nuclear magnetic resonance techniques. *J Mol Biol*. 1994;238:23-41.
29. Walter MR, Cook WJ, Zhao BG, et al. Crystal structure of recombinant human interleukin-4. *J Biol Chem*. 1992;267:20371-20376.
30. Lloyd CM, Snelgrove RJ. Type 2 immunity: expanding our view. *Sci Immunol*. 2018;3:eaat1604.
31. Lucey DR, Clerici M, Shearer GM. Type 1 and type 2 cytokine dysregulation in human infectious, neoplastic, and inflammatory diseases. *Clin Microbiol Rev*. 1996;9:532-562.
32. Keegan AD, Leonard WJ, Zhu J. Recent advances in understanding the role of IL-4 signaling. *Fac Rev*. 2021;10:71.
33. Bernstein ZJ, Shenoy A, Chen A, Heller NM, Spangler JB. Engineering the IL-4/IL-13 axis for targeted immune modulation. *Immunol Rev*. 2023;1-29. doi:10.1111/imr.13230
34. Vaz DC, Rodrigues JR, Sebald W, Dobson CM, Brito RM. Enthalpic and entropic contributions mediate the role of disulfide bonds on the conformational stability of interleukin-4. *Protein Sci*. 2006;15:33-44.
35. Kruse N, Lehrnbecher T, Sebald W. Site-directed mutagenesis reveals the importance of disulfide bridges and aromatic residues for structure and proliferative activity of human interleukin-4. *FEBS Lett*. 1991;286:58-60.
36. Weigel U, Meyer M, Sebald W. Mutant proteins of human interleukin 2. Renaturation yield, proliferative activity and receptor binding. *Eur J Biochem*. 1989;180:295-300.
37. Windsor WT, Syto R, Le HV, Trotta PP. Analysis of the conformation and stability of *Escherichia coli* derived recombinant human interleukin-4 by circular dichroism. *Biochemistry*. 1991;30:1259-1264.
38. Redfield C, Boyd J, Smith LJ, Smith RA, Dobson CM. Loop mobility in a four-helix-bundle protein: 15N NMR relaxation measurements on human interleukin-4. *Biochemistry*. 1992;31:10431-10437.
39. Ottiger M, Delaglio F, Bax A. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J Magn Reson*. 1998;131:373-378.
40. Bax A. Weak alignment offers new NMR opportunities to study protein structure and dynamics. *Protein Sci*. 2003;12:1-16.
41. Lipsitz RS, Tjandra N. Residual dipolar couplings in NMR structure analysis. *Annu Rev Biophys Biomol Struct*. 2004;33:387-413.
42. Schwieters CD, Kuszewski JJ, Clore GM. Using Xplor-NIH for NMR molecular structure determination. *Progr NMR Spectrosc*. 2006;48:47-62.
43. Berjanskii M, Wishart DS. NMR: prediction of protein flexibility. *Nat Protoc*. 2006;1:683-688.
44. Brünger AT, Clore GM, Gronenborn AM, Karplus M. Three-dimensional structure of proteins determined by molecular dynamics with interproton distance restraints: application to crambin. *Proc Natl Acad Sci USA*. 1986;83:3801-3805.
45. Kuszewski J, Clore GM. Sources of and solutions to problems in the refinement of protein NMR structures against torsion angle potentials of mean force. *J Magn Reson*. 2000;146:249-254.
46. Laskowski RA, Rullmann JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR*. 1996;8:477-486.
47. Heinig M, Frishman D. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res*. 2004;32:W500-W502.
48. Hubbard SJ, Thornton JM. *"NACCESS," computer program*. Department of Biochemistry and Molecular Biology, University College; 1993.
49. Lee HS, Choi J, Yoon S. QHELIX: a computational tool for the improved measurement of inter-helical angles in proteins. *Protein J*. 2007;26:556-561.
50. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996;14:33-38.
51. Pattabiraman N, Ward KB, Fleming PJ. Occluded molecular surface: analysis of protein packing. *J Mol Recognit*. 1995;8:334-344.
52. Boyd J, Hommel U, Campbell ID. Influence of cross correlation between dipolar and anisotropic chemical shift relaxation mechanisms upon longitudinal relaxation rates of nitrogen-15 in macromolecules. *Chem Phys Lett*. 1990;175:477-482.
53. Kay L, Torchia DA, Bax A. Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*. 1989;28:8972-8979.
54. Kay L, Keifer P, Saarinen T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc*. 1992;114:10663-10665.
55. Lipari G, Szabo A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J Am Chem Soc*. 1982;104:4546-4559.
56. Lipari G, Szabo A. Model-free approach to the interpretation of nuclear. Magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J Am Chem Soc*. 1982;104:4559-4570.

57. Clore GM, Driscoll PC, Wingfield PT, Gronenborn AM. Analysis of the backbone dynamics of interleukin-1 beta using two-dimensional inverse detected heteronuclear 15N-1H NMR spectroscopy. *Biochemistry*. 1990;29:7387-7401.
58. Clore GM, Szabo A, Bax A, Kay L, Driscol P, Gronenborn AM. Deviations from the simple two-parameter model-free approach to the interpretation of ^{15}N nuclear magnetic relaxation of proteins. *J Am Chem Soc*. 1990;112:4989-4991.
59. Provencher SW. CONTIN: a general purpose constrained regularization program for inverting noisy linear algebraic and integral equations. *Comput Phys Commun*. 1982;27:229-242.
60. Pace CN, Scholtz JM. Measuring the conformational stability of a protein. In: Creighton TE, ed. *Protein structure: A practical approach*. 2nd ed. Oxford University Press; 1997:299-321.
61. Vriend G. WHAT IF: a molecular modeling and drug design program. *J Mol Graph*. 1990;8:52-56.
62. Webb B, Sali A. Comparative protein structure modelling using Modeller. *Curr Protoc Bioinformatics*. 2016;54: Chapter 5:Unit-5.6. doi:10.1002/cpbi.3
63. Abraham MJ, Murtola T, Schulz R, et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015;1:19-25.
64. Lindorff-Larsen K, Piana S, Palmo K, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct Funct Genet*. 2010;78:1950-1958.
65. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys*. 2007;126:014101.
66. Berendsen HJC, Postma JPM, Van Gunsteren WF, Dinola A, Haak JR. Molecular dynamics with coupling to an external bath. *J Chem Phys*. 1984;81:3684-3690.
67. Parrinello M, Rahman A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J Appl Phys*. 1981;52:7182-7190.
68. Hess B. P-LINCS: a parallel linear constraint solver for molecular simulation. *J Chem Theory Comput*. 2008;4:116-122.
69. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589.
70. Fleming PJ, Richards FM. Protein packing: dependence on protein size, secondary structure and amino acid composition. *J Mol Biol*. 2000;299:487-498.
71. Bruylants G, Redfield C. (15)N NMR relaxation data reveal significant chemical exchange broadening in the alpha-domain of human alpha-lactalbumin. *Biochemistry*. 2009;48:4031-4039.
72. Boyd J, Redfield C. Defining the orientation of the ^{15}N shielding tensor using 15N NMR relaxation data for a protein in solution. *J Am Chem Soc*. 1998;120:9692-9693.
73. Redfield C, Smith RA, Dobson CM. Structural characterization of a highly ordered "molten globule" at low pH. *Nat Struct Biol*. 1994;1:23-29.
74. Jeng MF, Englander SW. Stable submolecular folding units in a non-compact form of cytochrome c. *J Mol Biol*. 1991;221:1045-1061.
75. Ptitsyn OB. Molten globule and protein folding. *Adv Protein Chem*. 1995;47:83-229.
76. Redfield C. NMR studies of partially folded molten-globule states. *Methods Mol Biol*. 2004;278:233-254.
77. Redfield C. Using nuclear magnetic resonance spectroscopy to study molten globule states of proteins. *Methods*. 2004;34:121-132.
78. Bartelli NL, Sun S, Gucinski GC, et al. The cytoplasm-entry domain of antibacterial CdiA is a dynamic α-helical bundle with disulfide-dependent structural features. *J Mol Biol*. 2019;43:3203-3216.
79. Radford SE, Woolfson DN, Martin SR, Lowe G, Dobson CM. A three-disulphide derivative of hen lysozyme. Structure, dynamics, and stability. *Biochem J*. 1991;273:211-217.
80. Ewbank JJ, Creighton TE. The molten globule protein conformation probed by disulphide bonds. *Nature*. 1991;350:518-520.
81. Quezada CM, Schulman BA, Froggatt JJ, Dobson CM, Redfield C. Local and global cooperativity in the human alpha-lactalbumin molten globule. *J Mol Biol*. 2004;338:149-158.
82. Levin AM, Bates DL, Ring AM, et al. Exploiting a natural conformational switch to engineer an interleukin-2 "superkine". *Nature*. 2012;484(7395):529-533.
83. Ren J, Chu AE, Jude KM, et al. Interleukin-2 superkines by computational design. *Proc Natl Acad Sci USA*. 2022;119(12):e2117401119.
84. Yang H, Ulge UY, Quijano-Rubio A, et al. Design of cell-type-specific hyperstable IL-4 mimetics via modular de novo scaffolds. *Nat Chem Biol*. 2023;19(9):1127-1137.
85. Watson JL, Juergens D, Bennett NR, et al. De novo design of protein structure and function with RFdiffusion. *Nature*. 2023;620(7976):1089-1100.
86. Di Lella S, Herrmann A, Mair CM. Modulation of the pH stability of influenza virus hemagglutinin: a host cell adaptation strategy. *Biophys J*. 2016;110:2293-2301.
87. Horng JC, Demarest SJ, Raleigh DP. pH-dependent stability of the human alpha-lactalbumin molten globule state: contrasting roles of the 6–120 disulfide and the beta-subdomain at low and neutral pH. *Proteins*. 2003;52(2):193-202.
88. Watanabe S, Harayama M, Kanemura S, Sitia R, Inaba K. Structural basis of pH-dependent client binding by ERp44, a key regulator of protein secretion at the ER-Golgi interface. *Proc Natl Acad Sci USA*. 2017;114:E3224-E3232.
89. van der Kamp MW, Daggett V. Influence of pH on the human prion protein: insights into the early steps of misfolding. *Biophys J*. 2010;99:2289-2298.
90. Jesus CSH, Cruz PF, Arnaut LG, Brito RMM, Serpa C. One peptide reveals the two faces of α-helix unfolding-folding dynamics. *J Phys Chem B*. 2018;122(14):3790-3800.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Vaz DC, Rodrigues JR, Loureiro-Ferreira N, et al. Lessons on protein structure from interleukin-4: All disulfides are not created equal. *Proteins*. 2023;1-17. doi:10.1002/prot.26611