International Journal of Biological Macromolecules 263 (2024) 130348

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: www.elsevier.com/locate/ijbiomac

Functional and structural characterisation of RimL from *Bacillus cereus*, a new N^{α}-acetyltransferase of ribosomal proteins that was wrongly assigned as an aminoglycosyltransferase

H. Leonardo Silvestre [a], J.L. Asensio [b], T.L. Blundell [a], A. Bastida [b,*], V.M. Bolanos-Garcia [a,c,**]

[a] Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, United Kingdom
[b] Departamento de Química Bio-orgánica, IQOG, Spanish National Research Council, C/ Juan de la Cierva 3, E-28006 Madrid, Spain
[c] Department of Biological and Medical Sciences, Oxford Brookes University, Oxford OX3 0BP, United Kingdom

ARTICLE INFO

Keywords:
N^{α}-acetyltransferase
GNAT family
Acetyl-CoA
RimL
L7/L12 ribosomal protein
Acetylation
Protein substrate recognition

ABSTRACT

Enzymes of the GNAT (GCN5-relate *N*-acetyltransferases) superfamily are important regulators of cell growth and development. They are functionally diverse and share low amino acid sequence identity, making functional annotation difficult. In this study, we report the function and structure of a new ribosomal enzyme, N^{α}-acetyl transferase from *Bacillus cereus* ($RimL^{BC}$), a protein that was previously wrongly annotated as an aminoglycosyltransferase. Firstly, extensive comparative amino acid sequence analyses suggested $RimL^{BC}$ belongs to a cluster of proteins mediating acetylation of the ribosomal protein L7/L12. To assess if this was the case, several well established substrates of aminoglycosyltransferases were screened. The results of these studies did not support an aminoglycoside acetylating function for $RimL^{BC}$. To gain further insight into $RimL^{BC}$ biological role, a series of studies that included MALDI-TOF, isothermal titration calorimetry, NMR, X-ray protein crystallography, and site-directed mutagenesis confirmed $RimL^{BC}$ affinity for Acetyl-CoA and that the ribosomal protein L7/L12 is a substrate of $RimL^{BC}$. Last, we advance a mechanistic model of $RimL^{BC}$ mode of recognition of its protein substrates. Taken together, our studies confirmed $RimL^{BC}$ as a new ribosomal N^{α}-acetyltransferase and provide structural and functional insights into substrate recognition by N^{α}-acetyltransferases and protein acetylation in bacteria.

1. Introduction

Post-translational modification (PTM) of proteins is common to all kingdoms of life [1]. It is estimated that about 50 to 70 % of soluble yeast proteins are acetylated, while for human proteins the estimate raises up to 80–90 %, making it one of the most widespread protein modifications [2]. In *Bacillus subtilis* acetylated proteins account for nearly 38 % of the total protein content in this bacterium [3]. The proteins responsible for these modifications are acetyl transferases. They are part of the GNAT (GCN5-relate *N*-acetyltransferases) super family, which share a common acetyl donor, Acetyl-CoA. Protein acetylation is known to occur at the N^{α} of lysine groups [4], O^{β} acetylation at Ser or Thr residues and N^{α} of the N-terminal amino acid [2]. Because acetyl transferases exhibit a high diversity of functions [5], a wide variety of acceptors are recognised by these enzymes, making the prediction of *bona fide* substrates (*e.g.*, the acyl group receptors) challenging.

In bacteria, proteins known to be lysine acetylated mostly belong to central metabolic and protein synthesis pathways, although some other lysine-acetylated proteins have been reported to participate in protein folding, amino acid and nucleotide biosynthesis and detoxification responses. Lysine acetylation can occur in one, two or multiple amino acids of a protein [3,6]. In eukaryote organisms, lysine acetylation is a PTM that occurs in abundance, usually associated to histones [7]. In contrast, acetylation at the N^{α}-terminal in bacteria is far less documented. Indeed, only seven examples of prokaryotic proteins are known to be acetylated and of these, the catalytic activity of only three proteins has been described in detail [8]. The best characterised acetyl transferases known to acetylate N^{α} terminal are certain members of the Rim protein family, which are responsible for the acetylation of ribosomal proteins. Three proteins comprise this family: RimL, RimJ, and RimI,

* Corresponding author.
** Correspondence to: V.M. Bolanos-Garcia, Department of Biological and Medical Sciences, Oxford Brookes University, Oxford OX3 0BP, United Kingdom.
E-mail addresses: agatha.bastida@csic.es (A. Bastida), vbolanos-garcia@brookes.ac.uk (V.M. Bolanos-Garcia).

https://doi.org/10.1016/j.ijbiomac.2024.130348
Received 25 September 2023; Received in revised form 29 January 2024; Accepted 19 February 2024
Available online 22 February 2024

which are responsible for acetylation of the ribosomal proteins L7/L12, S5, and S18, respectively. While S5 and S18 are constitutively acetylated, the extent of L7/L12 acetylation depends on the cells growth rate [9]. Knock-out mutants of RimL, RimJ, and RimI do not show any distinguishable phenotype. Of all Rim proteins, the orthologues from *Salmonella typhimurium* (Rim^{ST}) are the most extensively studied, functionally and structurally [8,10]. $RimJ^{ST}$ and $RimI^{ST}$ are designated as N^{α}-alanine acetyltransferases and $RimL^{ST}$ as a N^{α}-serine acetyltransferase. However, the terminal amino acid residues of these Rim proteins are not conserved across bacteria, and therefore these specificities relate to *Salmonella typhimurium* alone. This feature and the profusion of acetyl transferases in the bacterial genome makes difficult the accurate prediction of their functions [5]. Moreover, the fact individual genes from whole genome sequencing initiatives are normally annotated based on sequence similarity [2,11], imparts important limitations to the accurate prediction of function of proteins exhibiting high amino acid sequence divergence.

In this work, we describe the function and structure of the protein $RimL^{BC}$ from *Bacillus cereus* (ATCC 14579; DSM31; UniProt entry Q81D84), an enzyme that was previously annotated as an aminoglycosyl transferase. Acetyl-CoA binding to $RimL^{BC}$ was confirmed using diverse, complementary biophysical techniques including Saturation-Transfer Difference-Nuclear Magnetic Resonance (STD-NMR), thermal shift assays, and Isothermal titration calorimetry (ITC). Further evidence that the $RimL^{BC}$ enzyme acts as a N^{α} Acetyl transferase was provided after solving the crystal structure of native $RimL^{BC}$ and a double mutant in which two amino acid residues key for $RimL^{BC}$ function, C132 and K135, where substituted by alanine. It is worth mentioning that initial predictions of the 3D structure of $RimL^{BC}$ by the Artificial Intelligence (AI)-based programme alpha fold failed to accurately predict $RimL^{BC}$ substrate specificity. This limitation to accurately predict 3D protein structure of enzymes of the GNAT (GCN5-relate *N*-acetyltransferases) super family using AI methods evidences the need to solve 3D structures of protein molecules using experimental approaches. We also identified the ribosomal protein L7/L12 from *Bacillus cereus* as a *bona fide* substrate of $RimL^{BC}$. Matrix-Assisted Laser Desorption/Ionisation coupled to Time-Of-Flight mass spectrometry (MALDI-TOFF) assays showed the K80 and K83 residues of L7/L12 are acetylated by native $RimL^{BC}$ using Acetyl-CoA as the acetyl group donor but not by a double mutant where two key catalytic amino acid residues were substituted with Alanine ($RimL^{BC[C132A-K135A]}$). Taken together, our studies show that $RimL^{BC}$ is a GNAT protein member that exhibits N^{α} Acetyl transferase activity and that the ribosomal protein L7/L12 is a substrate of $RimL^{BC}$. Last, we advance molecular models of the $RimL^{BC}$–AcetylCoA and the $RimL^{BC}$–L7/L12 interactions, which provide clues about acetyl group donor and protein substrate recognition by N^{α}-acetyltransferases in bacteria, respectively.

2. Materials and methods

2.1. Molecular cloning, overexpression and purification of $RimL^{BC}$, double mutant $RimL^{BC[C132A-K135A]}$ and ribosomal L7/L12

The gene *BC_2494* from *Bacillus cereus* strain ATCC 14579 (UniProt entry: Q81D84) was amplified by PCR using primers designed to complement specifically 15 bp at the 5'ends of codifying and complementary DNA strands. The recognition sequences for the restriction enzymes *Hind*III and *Xho*I were introduced in the primers for the subsequent cloning of the amplicon in these sites (5′ ATA CGC AAG CTT GTG ATT AAA CTA GAA AGT TTT 3′ and 5′ ATA CGC CTC GAG TCA TTA CTT CTT CCA TTC ATA TTC TAA 3′, respectively). PCR amplification was specific with only one band of the expected length observed after DNA gel electrophoresis. The amplicon band was purified using a gel extraction kit, purified and double digested with the above mentioned restriction enzymes. The digested amplicon was then ligated to the vector pET-28b (+) previously digested with the restriction enzymes *Xho*I and *Hind*III and dephosphorylated with alkaline phosphatase to yield the plasmid pET- $RimL^{BC}$. The sequence of the gene encoding for the $RimL^{BC}$ protein was confirmed by Sanger sequencing before transformation of *E. coli* BL21(DE3) cells for protein overproduction. Time-course expression of recombinant $RimL^{BC}$ containing an N-terminal His_6-tag was monitored by sodium dodecyl sulphate-polyacrilamide gel electrophoresis (SDS-PAGE). For this, *E. coli* BL21(DE3) cells transformed with the plasmid pET-$RimL^{BC}$ were grown at 37 °C in LB medium containing 26 μg/ml Kanamycin. When the cell density determined at OD_{600nm} was 0.3–0.5, the cells were induced with 0.5 mM of IPTG. The culture was incubated for 3 h post-induction at 30 °C. For the purification of the recombinant $RimL^{BC}$ protein thus overproduced, cells were resuspended in a buffer solution containing 50 mM Tris pH 8.0, 200 mM NaCl, 30 mM imidazole and disrupted through the use of a high pressure homogeniser. The lysate was then centrifuged for 40 min at 4 °C at 18,000 rpm (Sorvall-S34 rotor) and the resulting supernatant loaded onto a prepacked 5 ml Ni^{2+} column (GE-Healthcare). The soluble $RimL^{BC}$ protein was eluted from the column using an imidazole gradient (up to 500 mM imidazole), followed by gel filtration using a 26/60 Superdex 75 column (Suppl Fig. 1A). The fractions containing $RimL^{BC}$ were concentrated and the buffer changed to 50 mM Tris pH 8.0, 100 mM NaCl. Samples were flash-frozen in liquid nitrogen and stored in -80 °C until use. $RimL^{BC[C132A-K135A]}$ was overexpressed and purified using identical protocols and experimental conditions described above for native $RimL^{BC}$.

The gene encoding the ribosomal protein L7/L12 (Uniprot Q81J50) from *B. cereus* was cloned in pET-28b (+) using the primers forward 5' CTG GCG GAT CCG ATG ACT AAA GAA CAA ATC ATT 3′ (*BamHI*) and reverse 5' CG CTT AAG CTT TTA CTT AAC TTC TAC AGC AGC GCC 3′ (*HindIII*) and the same methodology described for generating the plasmid pET-$RimL^{BC}$. L7/L12 was overexpressed in *E coli BL21* (DE3) and purified by IMAC. Overexpression of ribosomal L7/L12 was confirmed by SDS-PAGE analysis, which showed overproduction of a protein of *ca.* 12.6 KDa, the molecular mass of L7/L12 (Suppl Fig. 1A).

A double mutant of the gene encoding for $RimL^{BC[C132A-K135A]}$, was generated by nested PCR using this set of primers: forward 5′ AAT ACA TCG GCC ATT TCA GCG TAC GAA GCC ATA GGG TTT GTA AAA 3′; reverse 5′ TTT TAC AAA CCC TAT GGC TTC GTA CGC TGA AAT GGC CGA TGT ATT 3′.

2.2. Aminoglycoside acetyltransferase assay

The reaction was measured spectrophotometrically by following the increase in absorbance at 412 nm due to the formation of 5-thio-2-nitrobenzoate resulting from the reaction between the sulfhydryl group of the product of the acetyltransfer reaction, CoA-SH, and 5,5′-dithiobis (2-nitrobenzoic acid) (DTNB). Assay mixtures contained 50 mM HEPES buffer pH 8.0, 100 to 300 μM of Acetyl-CoA, and 0.05–1 mM of the aminoglycosides Kanamycin A, Kanamicyn B, Neomycin, Streptomycin, Paromomycin, and Ribostamycin in a volume of 1 ml. The reaction was initiated by the addition of $RimL^{BC}$ at 20 mM concentration. The absorbance of each assay mixture was measured periodically.

2.3. Thermal shift assay

To monitor protein unfolding, the fluorescent dye Sypro orange was used. The thermal shift assay was carried out with an iQ5 Real Time Detection System (Bio-Rad). The system contained a heating/cooling device for accurate temperature control and a charge-couple device detector for simultaneous imaging of the fluorescence changes in the wells of the microplate. Solutions of 100 μl contain 90 μl of 10 mM Tris buffer pH 8.0, 200 mM NaCl, 2.8× Sypro Orange, 4.4 μM of $RimL^{BC}$ plus 10 μl of ligands at 10 mM were added to the wells of the 96-well iCycler (Bio-Rad) iQ PCR plate. The plate was heated from 25 to 90 °C with a heating rate of 0.5 °C/min. Fluorescence intensity was measured at an Excitation/Emission wavelength of 490/530 nm. The ligands used were including Kanamycin A, Kananamycin B, Neomicin, Gentamicin C (1a),

Gentamicin C (2), Sisomicin, and Tobramycin as acceptor and Acetyl-CoA as donor. Eight technical replicas were performed for each experiment.

2.4. Isothermal titration calorimetry

ITC experiments were performed at 25 °C in 10 mM Tris buffer at pH 8.0, 200 mM NaCl using a VP-ITC calorimeter (Microcal, Inc.). The titration curves were measured by stepwise injection of a ligand (2 mM of Acetyl-CoA) into the reaction cell loaded with the protein solution (1.5 ml, 60 μM), with a total of 20 injections of ligand. The thermal effect (*e.g.*, heat exchange) per mol of cofactor injected was plotted against the cofactor/protein ratio after correcting for the heat of dilution. The binding parameters were obtained by nonlinear regression of the data using the simplest model compatible with the experimental curves. Three technical replicas were performed for each experiment.

2.5. Analytical ultracentrifugation

Sedimentation velocity experiments with RimL^{BC} in the 1–20 μM concentration range were performed using an analytical ultracentrifuge (Beckman XL-A) equipped with absorbance optics, using an An50Ti rotor. Whole-cell buoyant molecular masses ($M^{c}_{w,a}$) were determined by fitting a sedimentation equilibrium model for a single sediment solute to individual datasets with the program EQASSOC.

2.6. STD NMR

NMR spectra were acquired with a Bruker Avance 600 MHz spectrometer at 298 K. NMR samples were prepared in 10 mM of phosphate buffer pH 7.8, employing 50 μM protein concentration and 1–2 mM Acetyl-CoA or the assayed aminoglycosides. STD data sets were acquired with 1024 scans, using a train of 50 ms Gaussian-shaped pulses for the selective saturation of the RimL^{BC} protein envelope at $\delta = 7.2$ or -0.5 ppm. The total saturation time was fixed to 2 s and the off-resonance frequency was set at $\delta = 100$ ppm. A short 10 ms spin-lock pulse was employed to remove the background protein signals. No water-suppression scheme was applied.

2.7. Small angle X-ray scattering

High and low angle scattering data were collected at ESRF BM29 BIOSAXS (Experiment MX-1946 and MX-2072). Reconstruction of the molecular shape of dimeric RimL^{BC} from the scattering profile alone was carried out using the *ab initio* program Gasbor [12]. Numerous independent shape restorations were performed using both no symmetry as well as 2-fold symmetry constraints. To check the results of *ab initio* modelling, the program DAMMIN [13] was also used. These models were similar to those obtained with GASBOR but at a lower resolution due to the algorithm employed which uses a smaller number of atoms at larger, regularly spaced distances. The first crystal structure of RimL^{BC} described in this report (PDB entry 6ERD) was superimposed onto the low resolution structure of RimL^{BC} in solution and visualised using the program Pymol [14].

2.8. RimL^{BC} and RimL$^{BC[C132A-K135A]}$ protein crystallisation and structure resolution

For the crystallisation of RimL^{BC}, several hundred crystallisation conditions were screened using the sitting drop method. A single hit was observed for one condition, which consisted of 30 % PEG 6000, 0.1 M HEPES pH 7.0 and 1 M LiCl. The protein crystals were subsequently grown in bigger drops and tested at a synchrotron facility, diffracting at 2.5 Å. Encouraged by these results, a molecular replacement strategy was initially tested, but due to the very low sequence homology between members of the GNAT protein family, it was not possible to find the phases and solve the structure. This prompted us to develop a MAD approach, where selenium methionine was incorporated into the recombinant protein, purified and crystallised. Initially 1 mM DTT was utilised as the reducing agent. However, it was observed that the presence of this reducing agent prevented crystal growth. Other reducing agents were tried, with 10 mM β-mercaptoethanol being the most successful. Protein crystals were obtained in the above described condition, with and resulting diffraction of up to 2.0 Å resolution. The resulting images were processed in iMosflm [15], scaled and analysed using the AutoSol package of Phenix [16]. In the first round of Se atoms search, Phenix found 21 of the possible 36 Se atoms resulting from the original construct. After further refinement of Se positions, Phenix found 20 Se atom positions. This would imply that the His-tag present in the construct was partly cleaved. After this round, model building began and in the end more than 98 % of the protein was modelled and built. Water molecules were added as the model building proceeded. Table 1 shows the statistics for data collection and refinement for the SeMet dataset. Atomic coordinates and structure factors have been deposited in the PDB under the ID 6ERD. The RimL$^{BC[C132A-K135A]}$ double mutant was crystallised and solved by molecular replacement. The max resolution achieved for the RimL$^{BC[C132A-K135A]}$ double mutant was 1.52 Å. Structure factors and atomic coordinates of RimLBC$^{[C132A-K135A]}$ have been deposited in the Protein Data Bank under the ID 8Q6U.

2.9. RimL^{BC}-Acetyl-CoA molecular docking

The 3D structure of RimL^{BC} (PDB ID: 6ERD) was optimised using the Protein Preparation Wizard tool integrated in Maestro (Schrödinger, New York, NY, USA) [17]. Acetyl-CoA was prepared using the LigPrep module in the Schrödinger 2022 package and minimised using the OPLS4 force field. The docking grid of the enzyme was generated using the receptor grid generation tools in Glide. The size of the docked molecules was set to within 15 Å. Molecular docking was performed using the Glide-dock XP module [18,19]. Docking score, glide e-model, ionisation penalty, glide gscore, and TPSA were used to select the docking poses.

2.10. L7/L12 protein acetylation assay and analysis by High Resolution Mass Spectrometry (Q-Exactive HF)

The reaction consisted of 1 mM Acetyl-CoA, 50 μM RimL^{BC}, 150 μM L7/L12, and 50 mM HEPES, pH 7.5, in a final volume of 200 μl. The reaction mixture was incubated for 12 h at 25 °C. After dialysis to remove the excess of Acetyl-CoA, the proteins were separated by SDS-PAGE (4–12 % gradient precast gels). As a control, the same reaction but without RimL^{BC} was used. Bands of the L7/L12 protein were cut before analysis by Mass Spectrometry MALDI TOF/TOF Bruker Ultraflex (mass range 300–150,000 uma).

3. Results and discussion

RimL^{BC} and L7/L12 were produced in bacterial cells as a soluble histidine-tagged protein and purified by immobilised metal affinity chromatography (IMAC) followed by size exclusion chromatography (SEC) (Suppl Fig. 1A). It is worth noting that initial IMAC runs using a lower ion strength and imidazole concentration (150 mM NaCl and 30 mM imidazole, respectively, resulted in the co-purification of RimL^{BC} with a native *E. coli* protein of approx. 67 kDa (Suppl Fig. 1B), most likely Glucosamine-6-phosphate synthase (GlmS) [20], an enzyme that catalyses the formation of D-glucosamine 6-phosphate from D-fructose 6-

Table 1
SEC-MALS data of full length RimL^{BC} in aqueous solution.

Molecular mass (kDa)	Calculated mass (μg)	Mass fraction (%)
43.4	68.47	100

phosphate using L-glutamine as the ammonia source. Increasing ion strength and imidazole concentration solved this problem as GlmS was efficiently removed from the IMAC column before RimL^{BC} elution (Suppl Fig. 1B). Densitometry analysis of the SDS-PAGE following SEC purification (Suppl Fig. 1C) indicated a protein purity higher than 90 %. A similar level of purity was obtained for L7/L12 after SEC (Suppl Fig. 1A). In order to determine the molecular mass of RimL^{BC} in aqueous solutions, SEC coupled to multi-angle light scattering (SEC-MALS) was used (Table 1 and Fig. 1A). The results from this analysis revealed a protein of a molecular mass of 43 kDa, which is double the calculated for a RimL^{BC} monomer, indicating that RimL^{BC} is a stable homodimeric protein in aqueous solutions. The oligomerisation state of RimL^{BC} was confirmed by analytical ultracentrifugation (AUC) experiments carried out at different protein concentrations (1–20 μM range) and by Small Angle X-ray Scattering (SAXS) studies. The AUC and SAXS results further confirmed that RimL^{BC} is dimeric under all the experimental conditions tested (Suppl Fig. 1D–G). As the protein was originally annotated as possessing aminoglycoside N6' acetyltransferase activity (UniProt Q81D84), a functional test (*e.g.*, the DTNB assay) was developed to test the potential binding of several aminoglycosides including Kanamycin A, Kananamycin B, Neomicin, Gentamicin C (1a), Gentamicin C (2), Sisomicin, and Tobramycin as potential RimL^{BC} substrates, using Acetyl-CoA as the acetyl donor. No acetyltransferase activity was observed for all these aminoglycosides in DNTB assays. In order to further confirm that none of these molecules was a *bona fide* RimL^{BC} substrate, thermal shift assays, in which a fluorescent dye is employed to assess the physical interaction between a protein and a specific ligand, were developed. In a thermal shift assay, a physical interaction between a protein and a ligand is evidenced by an increase in the melting temperature (Tm) compared to that of the protein in the absence of any ligand. The thermal

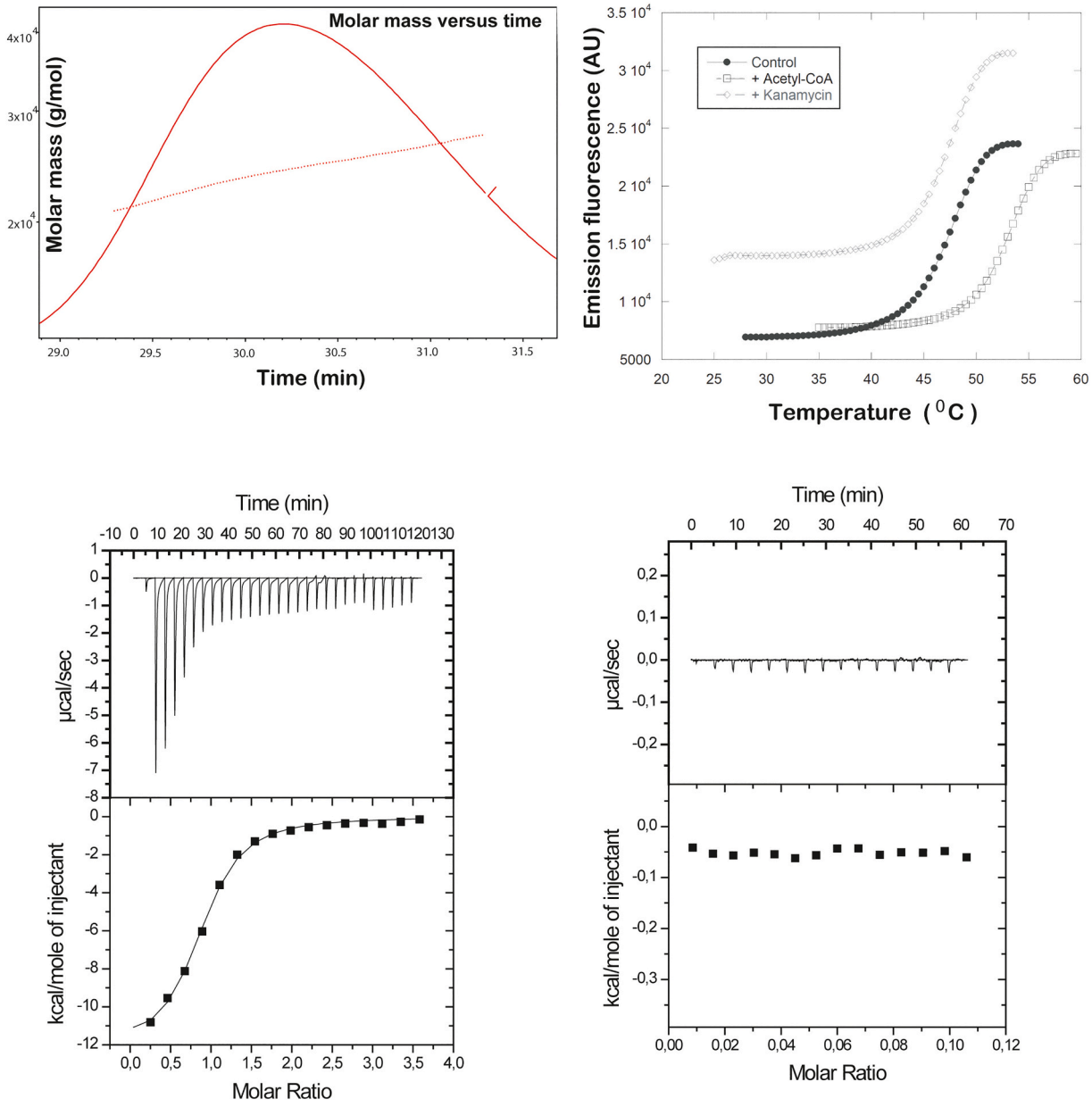

Fig. 1. Biophysical studies of RimL^{BC}. (A) SEC-MALS profile showing RimL^{BC} self-associates to form a dimer in aqueous solutions. (B) Thermal Shift of RimL^{BC} (4.4 μM) in 10 mM of Tris buffer pH 8, 200 mM of NaCl. Tm of the blank = 48 °C, Tm of RimL^{BC} plus Kanamycin (1 mM) =48 °C and of RimL^{BC} in the presence of Acetyl-CoA (1 mM) Tm = 53 °C. The lack of thermal shift in the presence of the aminoglycoside Kanamycin and the 5 °C shift in the presence of Acetyl-CoA indicated RimL^{BC} binding to the latter molecule (C) ITC of RimL^{BC} (40 μM) in 10 mM of Tris buffer pH 8.0, 200 mM NaCl upon titration with Acetyl-CoA and (D) upon titration with Kanamycin A confirmed RimL^{BC} binding to Acetyl-CoA but not to Kanamycin A.

denaturation of RimL^{BC} alone was found to be 48 °C, and no variation in the Tm was observed for RimL^{BC} in the presence of any of the aforementioned aminoglycosides. Interestingly, for both Acetyl-CoA and CoA, a 5 °C shift was observed (Tm = 53 °C), thus indicating that both molecules were able to bind to RimL^{BC} (Fig. 1B). In order to provide further evidence for this conclusion, STD-NMR experiments of protein/Acetyl-CoA and protein/aminoglycoside mixtures were performed. These experiments showed that acetyl-CoA binds to RimL^{BC} with a moderate affinity, as revealed by an efficient transfer of magnetisation to several acetyl-CoA protons (Fig. 2). On the other hand, no STD-NMR signals were observed for the aminoglycosides Kanamycin A, Kanamycin B, Neomicin, Gentamicin C (1a), Gentamicin C (2), Sisomicin, and Tobramycin under any of the experimental conditions tested (data no shown). Saturation transfer to the terminal acetylsulfanyl moiety of acetyl-CoA was the most intense, suggesting that this ligand region must be deeply buried in the protein binding pocket. In contrast, magnetisation transfer to the adenine nucleotide seemed limited probably due to a more exposed location of this fragment.

Furthermore, to quantitatively measure protein-ligand association, isothermal titration calorimetry (ITC) experiments were conducted, in which the RimL^{BC} protein was titrated with Acetyl-CoA or the aminoglycoside Kanamycin A. The ITC data showed that one molecule of Acetyl-CoA binds to each RimL^{BC} monomer with an affinity constant of 40 μM at pH 8.0 (Fig. 1C). The process is enthalpically driven while the entropic term opposes binding. In contrast, no specific association with RimL^{BC} was observed for the aminoglycoside Kanamycin A (Fig. 1D),

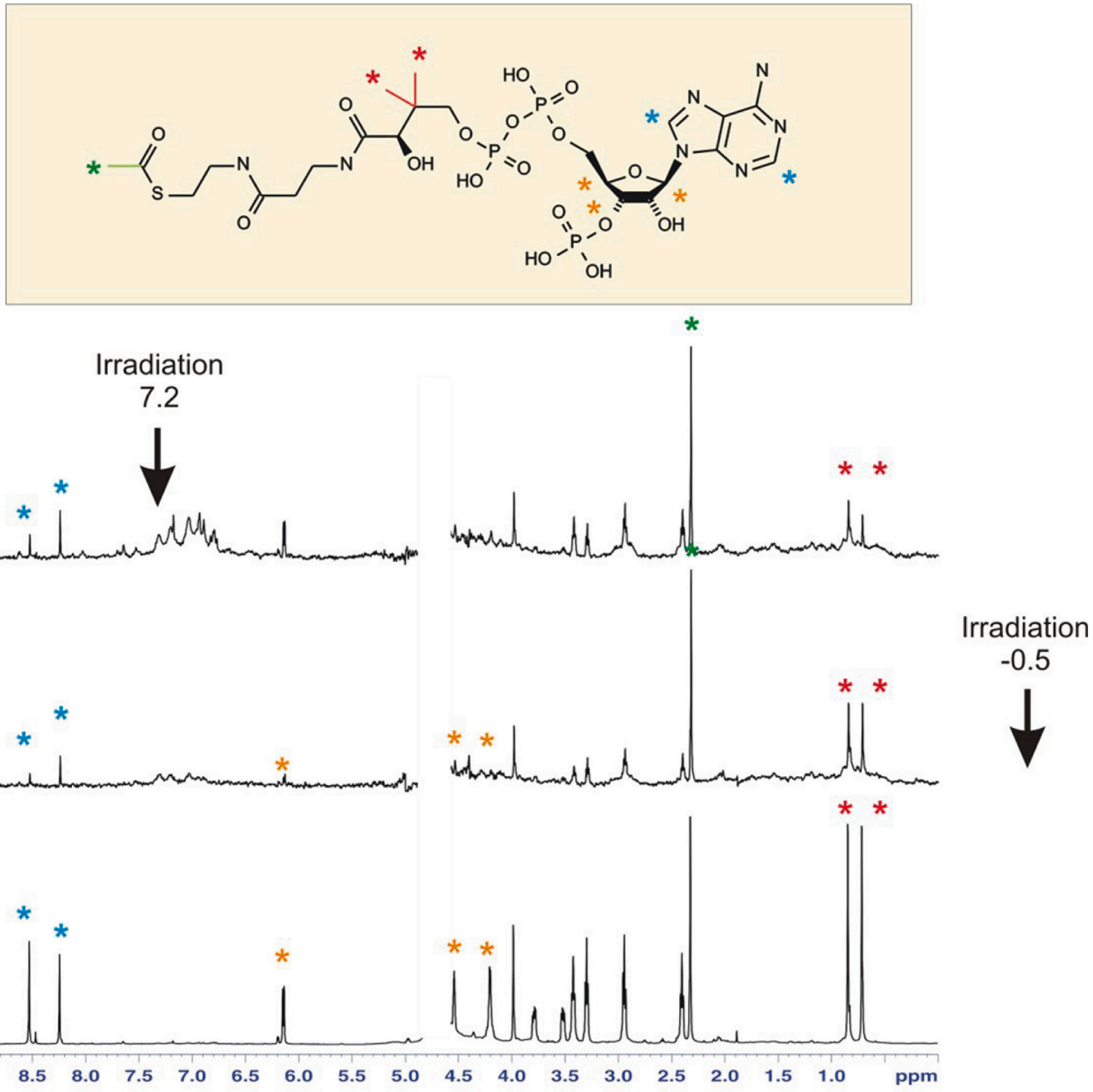

Fig. 2. 600 MHz STD NMR spectra obtained in 10 mM phosphate buffer at pH 7.8 and 25 °C, upon saturating (2 s) the RimL^{BC} protein (50 μM) at 7.2 (upper panel) or − 0.5 ppm (medium panel), in the presence of Acetyl-CoA (1 mM). A reference 1D-spectrum of the mixture is shown at the bottom panels. Assignments for selected Acetyl-CoA signals are shown with coloured asterisk.

further confirming the affinity of RimL^{BC} for Acetyl-CoA and that Kanamycin A is not a ligand of this enzyme.

Taken together, the thermal shift, ITC, and STD-NMR experiments consistently showed that RimL^{BC} has no affinity for N$^{\alpha}$ aminoglycosides. Therefore, we concluded this protein was previously wrongly annotated as an aminoglycoside acetyltransferase. In an attempt to validate further this conclusion, we set out to crystallise and solve the structure of RimL^{BC} using an anomalous X-ray scattering method. This approached enabled us to solve the crystal structure of RimL^{BC} at a maximum resolution of 2.0 Å. Details of the X-rays protein crystallography data and statistics are shown in Table 2. The crystal structure of RimL^{BC} revealed this enzyme adopts a typical GNAT fold, which consists of the following topology: β1, α1, α2 (motif C), β2, β3 (motif D), β4, α3 (motif A), β5, α4 (motif B), β6, and β7 (Fig. 3A). Most of the β-sheets run on anti-parallel fashion, except for sheets β4 and β5 that run parallel, making the central core of the protein to adopt a "V" like conformation. Interestingly, the early use of the programme alpha-fold (*e.g.*, before we solved the crystal structure of RimL^{BC}) predicted a protein of aminoglycosyltransferase activity. Since the structural information was released on the PDB (ID 6ERD), alpha-fold has changed the structure prediction to that of an aminoglycoside N6'-acetyltransferase. This example highlights the essential role that remains for experiment determination of individual enzymes (and of course, of macromolecular assemblies) using experimental approaches. Moreover, current AI-based approaches such as alpha-fold models perform less well than experimental structures as targets for the computational docking algorithms used in drug design [21]. This fact plus the possibility of unveiling unanticipated 3D structural features such as specific metal ions, obligate cofactors and stoichiometry, provides further support to the need for continuing developing experimental structural biology [22].

SEC-MALS, AUC and SAXS indicated that RimL^{BC} is a homodimer in aqueous solutions. This seems in agreement with RimL^{BC} self-association in the crystalline state, where RimL^{BC} showed a contact interface between two protomers to form a dimer of dimers (Fig. 3B–C). There is a twofold symmetry between one protomer and the other and several charged residues located at the interface between the protomers further stabilise the homodimer. Such stabilising interactions occur mostly at the edges of the β6-β6' sheets and involve the establishment of salt bridges between residues 183Glu-158His, 187Arg-208Glu, 187Arg-183Glu, 203Ser-187Arg, 158His-188Glu and 207Tyr-188Glu of one protomer and the corresponding symmetry-related amino acids of the other protomer. There are two further stabilising interactions worth noting: a salt bridge formed between residues 159Arg-183Glu of the same polypeptide and an interaction between 185Leu and 183Glu that is mediated by a water molecule. A similar overall topology to that of RimL^{BC} is found in the enzyme RimL from *Salmonella enterica* subsp. *enterica serovar Typhimurium str. LT2* (RimLST), described by Blanchard and collaborators [8]. The amino acid sequence alignment shown in Fig. 3D revealed the extent of amino acid residues conservation between RimL^{BC} and RimLST.

Table 2
X-ray data collection and refinement statistics.

Structure	SeMet RimL^{BC}	RimL$^{BC[C132A-K135A]}$ double mutant
Space group	$P2_1$	P 1 21 1
Unit cell	a = 68.48, b = 84.13 c = 72.07	
Resolution range Å	72–2.0 (2.05–2.00)	19.32–1.52
Unique reflections	109,729	31,604
Completeness	99 %	57.4
Redundancy	3.2 (4.5)	
I/σ (I)	12.5 (2.2)	
Refinement		
Rwork (%)	0.17	0.21
Rfree (%)	0.23	0.24
Average B-factor		
Protein	23.48	21.66
Water	27.67	

In an attempt to gain further structural insight into RimL^{BC} catalytic activity, we attempted to cocrystallise RimL^{BC} in complex with Acetyl-CoA. However, this could not be achieved, possibly due to hydrolysis of the Acetyl-CoA thioester bond during the course of the crystallisation experiments. The unsuccessful attempts to crystallise RimL^{BC} in complex with Acetyl-CoA, led us to develop a molecular docking approach to model the interaction between RimL^{BC} and Acetyl-CoA. Because the crystal structure of RimLST (PDB ID 1S7N) was solved in the presence of CoA as a ligand, this structure was used to study the possible mode of binding of CoA to RimL^{BC}. The model structure of RimL^{BC} with CoA (Fig. 4A) suggests that the adenine moiety, which electron density was clearly visible in the crystal structure of RimL^{BC}, is located on the back surface of the central pocket. Interestingly, in the model structure of RimL^{BC} with CoA, the sulfhydryl group of CoA does not appear to establish contacts with any amino acid residue nearby, a feature that was previously observed in the RimLST-CoA complex [8]. Instead, the crystal structure of RimLST in complex with CoA shows that the loop connecting α4 to β4 is involved in the interaction with the CoA ribose moiety and the 3'phosphate group.

Because a closer inspection of the crystal structure of RimL^{BC} indicated that residues C132 and K135 may be important for the interaction with CoA (Fig. 4B), we set out to crystallise and solve the structure of a RimL^{BC} double mutant in which C132 and K135 were substituted by Alanine (RimL$^{BC[C132A-K135A]}$). The crystal structure of the RimL$^{BC[C132A-K135A]}$ double mutant was solved by molecular replacement at high resolution (1.52 Å) and together with the crystal structure of apo RimL^{BC}, used for computational analyses of the RimL^{BC}-Acetyl-CoA interaction. The molecular docking studies support the view that the substitution of C132 and K135 result in the loss of protein contacts with Acetyl-CoA (Fig. 4C).

For bacterial ribosomes, it is known that certain protein subunits of the ribosome are acetylated, including L7/12, S8 and S5 [25–27]. The initial bioinformatics and ulterior biochemical and structural analyses described above suggested that ribosomal proteins could be a substrate of RimL^{BC}, resulting in their acetylation. In an attempt to clarify if this was the case, spectra fragmentation analysis by MS/MS of the ribosomal protein L7/L12 from *Bacillus cereus*, which has been identified as a substrate of N$^{\alpha}$-acetyltransferases in other organisms [8], was carried out in the presence of Acetyl-CoA with and without RimL^{BC}. The spectra fragmentation studies showed that the 50S ribosomal protein L7/L12 from *B. cereus* was acetylated in the residues K80 and K83 only in the presence of RimL^{BC} (Fig. 5), confirming L7/L12 as a substrate of this N$^{\alpha}$-acetyltransferase.

Interestingly, the surface area of contact between two RimL^{BC} protomers is 987.3 A^2 (N HB-21, N SB-12), which is much smaller compared to the interface of RimLST. In the latter enzyme, at the dimer interface a large cavity of 1863 A^3 is formed, which has been suggested to constitute the binding region of RimLST ribosomal protein substrates (Fig. 6A) and where N$^{\alpha}$ acetylation occurs [8]. In the model proposed by Blanchard and collaborators, it was postulated that the predominantly helical N-terminal region of L7/L12 binds to the trough formed upon RimLST dimerisation (Fig. 6A), after which a large conformational change should occur to bring to close proximity the L7/L12 C-terminal region and the RimLST CoA binding site.

However, the proposed L7/L12 substrate binding region of RimLST is at odds with the series of docking analyses carried out in this study on both RimLST and RimL^{BC} using well established computational approaches [28–31]. These computational analyses suggest a very different mode of interaction, where the C-terminal region of L7/L12 was consistently predicted to dock in a RimLST and RimL^{BC} region that is in close proximity to the CoA binding site. Interestingly, the C-terminal region of L7/L12 was predicted to dock to the same RimLST and RimL^{BC}

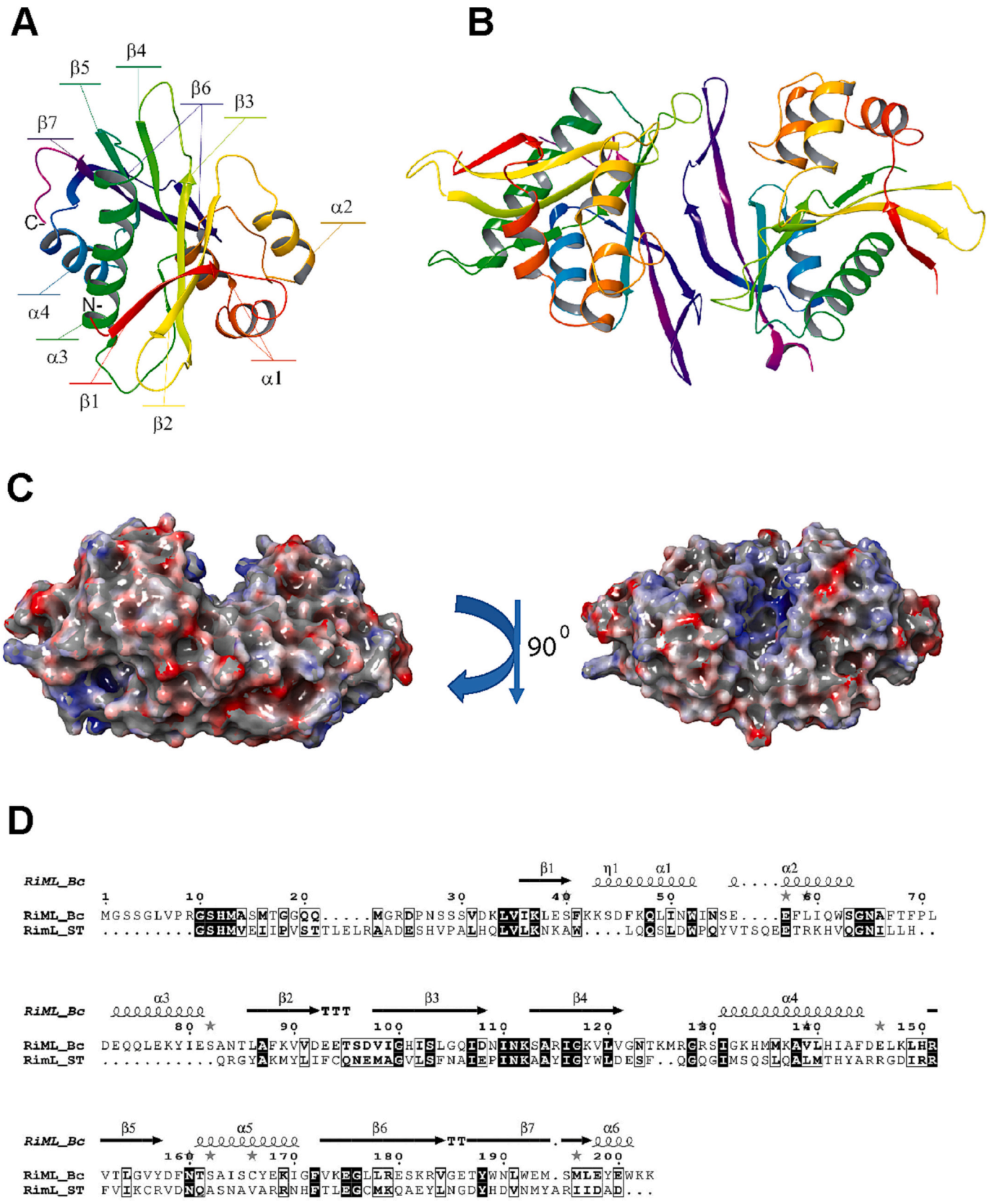

Fig. 3. 3D structure of RimL^{BC} (A) Details of its secondary structure organisation. (B) RimL^{BC} self-associates to form a homodimer in solution, an oligomerisation state that was recapitulated in the crystalline state. (C) Surface representation of the RimL^{BC} dimer (face centered, left). Same view after a 90^{0} rotation showing the buried pocket formed between the interface of both proteins (top view, right). Protein structure cartoons generated with Pymol [14]. (D) Amino acid sequence alignment of RimL^{BC} and RimLST showing the extent of amino acid residues conservation between these N$^{\alpha}$-acetyltransferases. The secondary structure elements above the aligned sequences are from the crystal structure of RimL^{BC}• Sequence alignment generated with Clustal Omega [23]. Figure generated with EstPript [24].

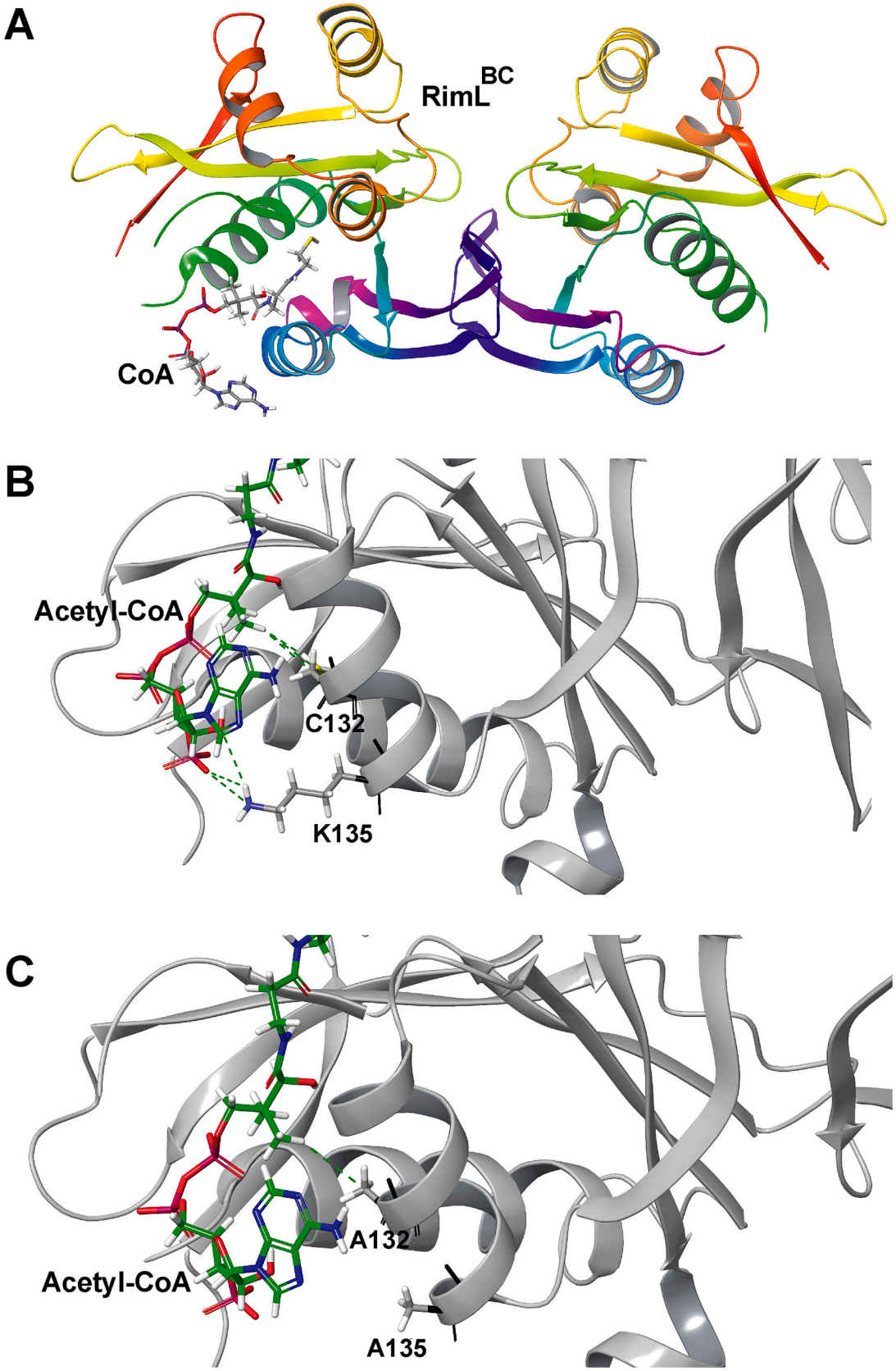

Fig. 4. (A) 3-D model structure of $RimL^{BC}$ (PDB ID 6ERD) in complex with CoA. (B) Cartoon representation of $RimL^{BC}$ highlighting the contacts of residues C132 and K135 with Acetyl-CoA. (C) Cartoon representation of $RimL^{BC[C132A-K135A]}$ double mutant showing the impact of these amino acid substitutions on the interaction with Acetyl-CoA.

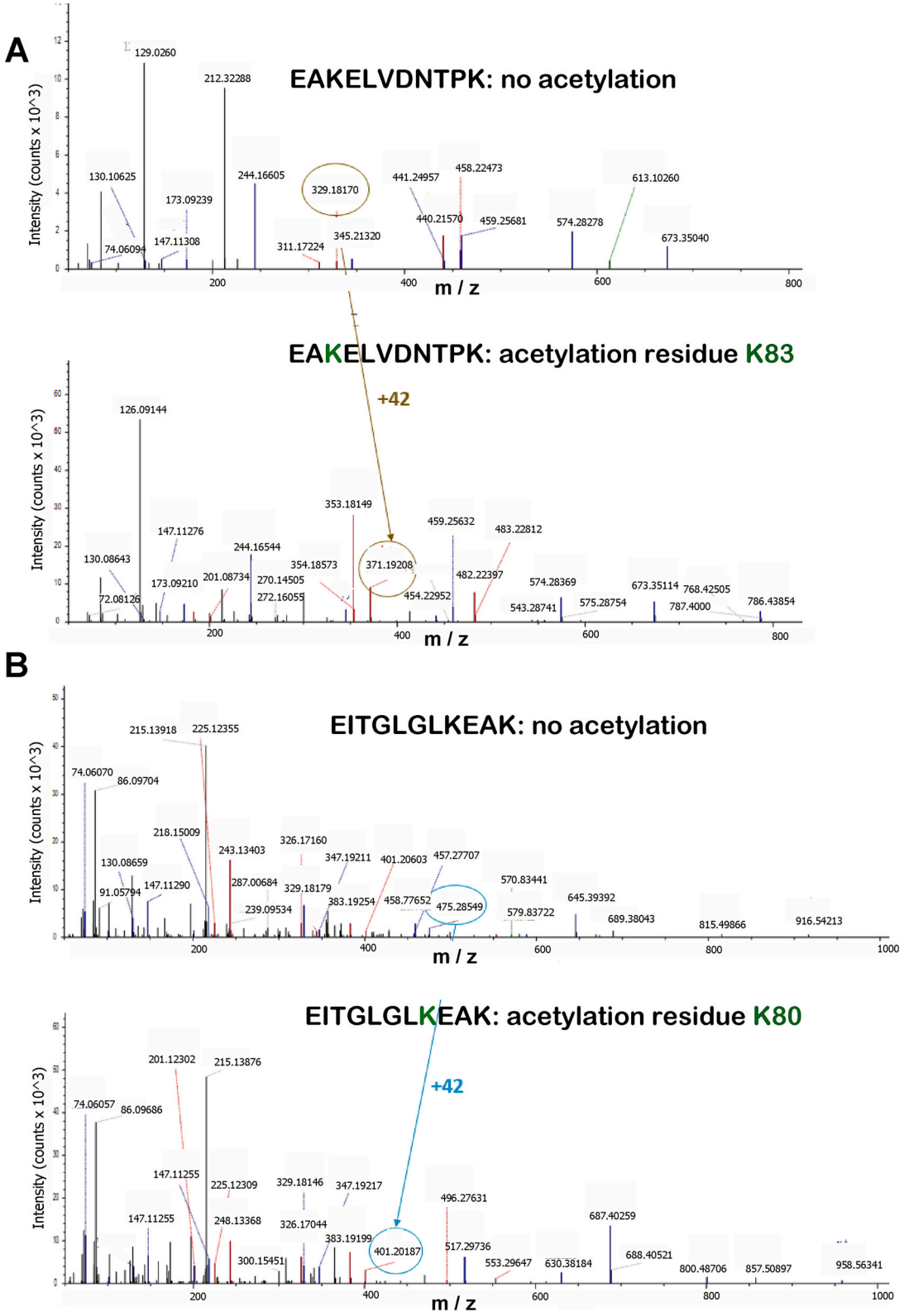

Fig. 5. Spectra of fragmentation by MS/MS of the ribosomal protein L7/L12. (A) L7/L12 protein in the presence of Acetyl-CoA (EA**K**ELVDNTPK) (upper panel), and in the presence of RimL^{BC} with Acetyl-CoA (EAκELVDNTPK acetylated K83) (lower panel). The amino acid sequence of 50S Ribosomal L7/L12 from *Bacillus cereus* was MTKEQIIEAVKSMTVLELNDLVKAIEEEFGVTAAAPVAVAGGAGEAAAEKTEFDVELTSAGAQKIKVIKVVREITGLGLKEA-κELVDNTPKVIKEAAAKEEAEEIKAKLEEVGAAVEVK. (B) Spectrum of fragmentation by MS/MS of L7/L12 in the presence of Acetyl-CoA (EA**K**ELVDNTPK) (upper panel), and RimL^{BC} in the presence of Acetyl-CoA and L7/L12 (EITGLGLκEAK acetylated K80) (lower panel). The amino acid sequence of 50S Ribosomal L7/L12 from *Bacillus cereus* DSM31 was MTKEQIIEAVKSMTVLELNDLVKAIEEEFGVTAAAPVAVAGGAGEAAAEKTEFDVELT SAGAQKIKVIKVVREITGLGLκEAKELVDNTPKV IKEAAAKEEAEEIKAKLEEVGAAVEVK.

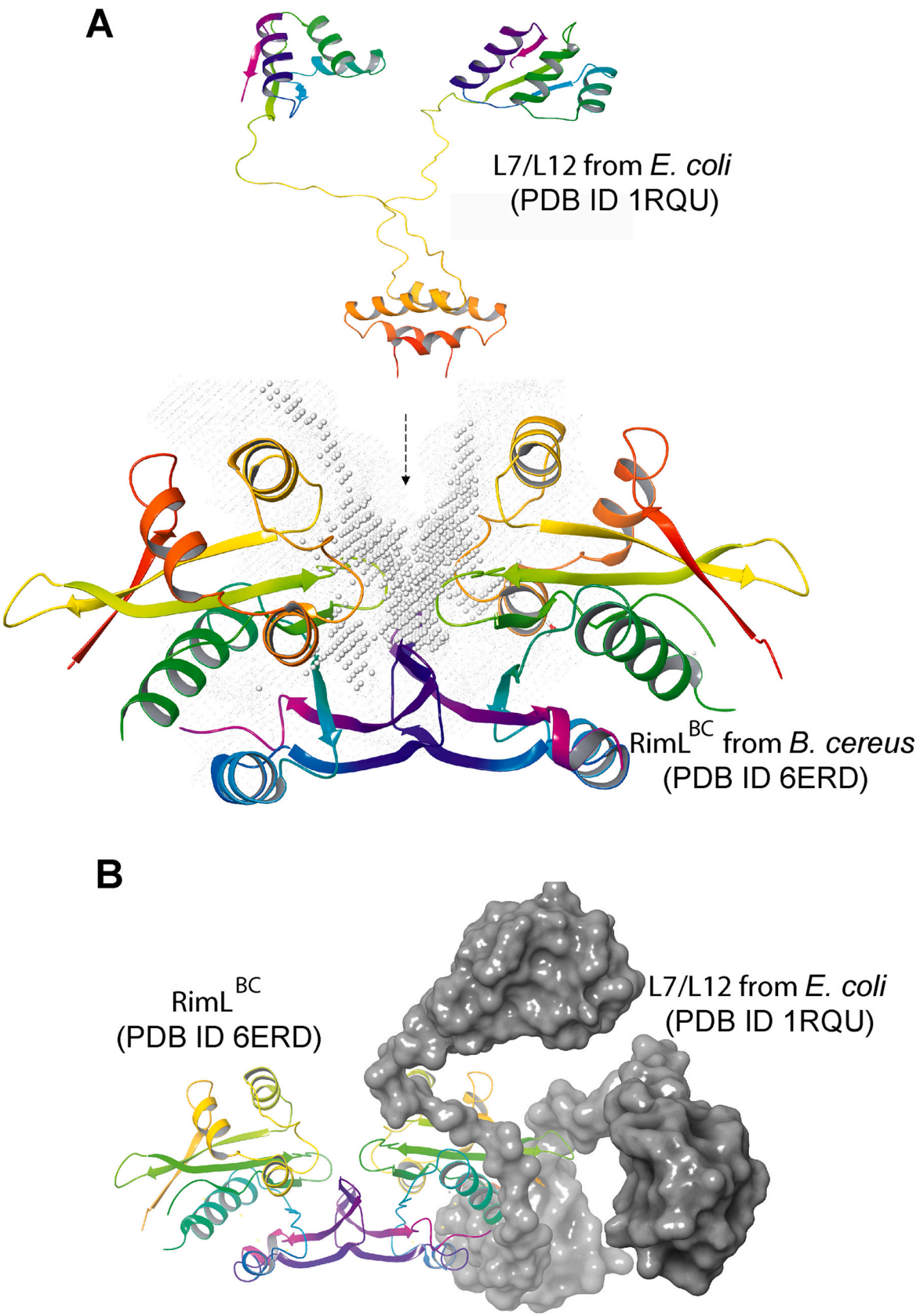

Fig. 6. (A) 3-D model structure showing the plausible mode of $RimL^{BC}$ (PDB ID 6ERD) interaction with the protein substrate L7/L12 (PDB ID 1RQU) proposed by Blanchard and collaborators [8]. (B) Alternative mode of $RimL^{BC}$-L7/L12 interaction postulated in this study.

region when using the L7/L12 full length amino acid sequence and a truncated segment encompassing the C-terminal region (Fig. 6B). These predictions encourage us to attempt the co-purification of full length L7/L12 and RimL^{BC} region by IMAC followed by SEC. However, initial protein co-purification trials showed no evidence of the formation of a stable binary complex, an observation that was further supported by SEC-MALS runs conducted at 20 °C and pH 8.0. These findings suggest that additional RimL^{BC} residues are implicated in the establishment of a productive L7/L12 and RimL^{BC} interaction. Future studies on full length proteins using modern techniques such as native mass spectrometry may be helpful to confirm the L7/L12-RimL^{BC} physical interaction and establish the stoichiometry of L7/L12-RimL^{BC} complex formation.

4. Conclusions

The present study reports the functional and structural characterisation of RimL from *Bacillus cereus* (RimL^{BC}; UniProt entry Q81D84), an N$^{\alpha}$-acetyltransferase enzyme that was previously wrongly assigned as an aminoglycosyltransferase. Such incorrect annotation is not entirely surprising given the profusion and low sequence homology of acetyl transferases in bacteria species. Our study also reports the identification of the *B. cereus* ribosomal protein L7/L12 as a true substrate of RimL^{BC}, in a process involving RimL^{BC}'s use of Acetyl-CoA as the acetyl group donor. Because of our inability to crystallise RimL^{BC} in complex with Acetyl-CoA, we advance a molecular model of the plausible RimL^{BC}–AcetylCoA interaction as well as that of RimL^{BC} with L7/L12. These models provide some molecular clues about the recognition of the acetyl group donor and one protein substrate by RimL^{BC}. We postulate that a similar mode of molecular recognition of acetyl group donor and protein substrates occurs in other bacteria N$^{\alpha}$-acetyltransferases.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2024.130348.

Abbreviations

ATCC	American Type Culture Collection
AUC	Analytical Ultracentrifugation
DTNB	5,5′-dithiobis (2-nitrobenzoic acid)
GNAT	GCN5-relate *N*-acetyltransferases
IMAC	Immobilised Metal Affinity Chromatography
ITC	Isothermal titration calorimetry
MALDI-TOF	Matrix-Assisted Laser Desorption/Ionisation coupled to Time-Of-Flight mass spectrometry
MALS	Multiple Angle Light Scattering
PCR	Polymerase Chain Reaction
PTM	Post-Translational Modification
PDB	Protein Data Bank
STD-NMR	Saturation-Transfer Difference-Nuclear Magnetic Resonance
SEC	Size Exclusion Chromatography
SDS-PAGE	Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis

Funding

AB. Ministerio de Educación y Ciencia. Estancias de profesores-investigadores en centros extranjeros de enseñanza superior e investigación. Programa Salvador de Madariaga. PRX17/00085 and PR2009-0087.

CRediT authorship contribution statement

H. Leonardo Silvestre: Writing – review & editing, Writing – original draft, Investigation. **J.L. Asensio:** Writing – original draft, Investigation. **T.L. Blundell:** Investigation, Writing – review & editing, Funding acquisition. **A. Bastida:** Writing – review & editing, Writing – original draft, Project administration, Investigation, Funding acquisition, Conceptualization. **V.M. Bolanos-Garcia:** Writing – review & editing, Writing – original draft, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Structure factors and atomic coordinates of native (RimL^{BC}) and double mutant (RimL$^{BC[C132A-K135A]}$) proteins have been deposited in the Protein Data Bank (ID 6ERD and 8Q6U, respectively).

Acknowledgments

Dr. Petra Pernot, beamline BM29 European Synchrotron Radiation Facility (ESRF) and Dr. Ed Lowe, Dept. of Biochemistry, University of Oxford for their help in SAXS data collection. Dr. David Staunton, Dept. of Biochemistry, University of Oxford for his help in SEC-MALS data collection.

References

[1] C.T. Walsh, S. Garneau-Tsodikova, G.J. Gatto, Protein post-translational modifications: the chemistry of proteome diversifications, Ang. Chem Int. Ed. Engl. 44 (2005) 7342–7372.
[2] B. Polevoda, F. Sherman, N-terminal acetyltransferases and sequence requirements for N-terminal acetylation of eukaryotic proteins, J. Mol. Biol. 325 (2003) 595–622.
[3] M. Zhang, T. Liu, L. Wang, Y. Huang, R. Fan, K. Ma, Y. Kan, M. Tan, J.Y. Xu, Global landscape of lysine acylomes in *Bacillus subtilis*, J. Proteomics 271 (2023) 104767.
[4] J.D. Jones, C.D. O'Connor, Protein acetylation in prokaryotes, Proteomics 11 (2011) 3012–3022.
[5] T. Arnesen, P. Van Damme, B. Polevoda, K. Helsens, R. Evjenth, et al., Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 8157–8162.
[6] J. Zhang, R. Sprung, J. Pei, X. Tan, et al., Lysine acetylation is a highly abundant and evolutionarily conserved modification in Escherichia coli, Mol. Cell. Proteomics 8 (2009) 215–225.
[7] M.D. Shahbazian, M. Grunstein, Functions of site-specific histone acetylation deacetylation, Annu. Rev. Biochem. 76 (2007) 75–100.
[8] M.W. Vetting, L.P. de Carvalho, S.L. Roderick, J.S. Blanchard, A novel dimeric structure of the RimL N-acetyltransferase from *Salmonella typhimurium*, J. Biol. Chem. 280 (23) (2005) 22108–22114.
[9] S. Ramagopal, A.R. Subramanian, Growth-dependent regulation iu production and utilization of acetylated ribosomal protein L7, Mol. Biol. 94 (1975) 633–641.
[10] M.C. Vetting, D.C. Bareich, M. Yu, J.S. Blanchard, Crystal structure of RimI from *Salmonella typhimurium* LT2, the GNAT responsible for N$^{\alpha}$-acetylation of ribosomal protein S18, Protein Sci. 17 (10) (2008) 1781–1790.
[11] L.I. Hu, B.P. Lima, A.J. Wolfe, Bacterial protein acetylation: the dawning of a new age, Mol. Microbiol. 77 (2010) 15–21.
[12] D.I. Svergun, Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing, Biophys. J. 76 (6) (1999) 2879–2886.
[13] D.I. Svergun, M.V. Petoukhov, M.H.J. Koch, Determination of domain structure of proteins from X-ray solution scattering, Biophys. J. 80 (6) (2001) 2946–2953.
[14] The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.
[15] A.G.W. Leslie, Joint CCP4 +, ESF-EAMCB Newsletter on Protein Crystallography. 26 (1992).
[16] P. D. Adams, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Cryst. D66 (2010) 213–221.
[17] G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments, J. Comput. Aided Mol. Des. 27 (2013) 221–234.
[18] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, D.T. Mainz, Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, J. Med. Chem. 49 (2006) 6177–6196.
[19] G.L. Warren, C.W. Andrews, A.M. Capelli, B. Clarke, J. LaLonde, M.H. Lambert, M. Lindvall, N. Nevins, S.F. Semus, S. Senger, et al., A critical assessment of docking programs and scoring functions, J. Med. Chem. 49 (2006) 5912–5931.
[20] V.M. Bolanos-Garcia, O.R. Davies, Structural analysis and classification of native proteins from *E. coli* commonly co-purified by immobilised metal affinity chromatography, Biochim. Biophys. Acta 1760 (2006) 1304–1313.

[21] M Karelina, J.J. Noh, R. O. Dror. How accurately can one predict drug binding modes using AlphaFold models? eLife 12 (2023) RP89386.
[22] R.J. Read, E.N. Baker, C.S. Bond, E.F. Garman, M.J. van Raai, AlphaFold and the future of structural biology, Acta Crystallogr. D Struct. Biol. 79 (7) (2023) 556–558.
[23] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal omega, Mol. Syst. Biol. 7 (2011) 1–6.
[24] X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server, Nucleic Acids Res. 42 (2014) W321–W324.
[25] Y. Gordiyenko, S. Deroo, M. Zhou, H. Videler, C.V. Robinson, Acetylation of L12 increases interactions in the Escherichia coli ribosomal stalk complex, J. Mol. Biol. 380 (2) (2008) 404–414.
[26] X. Liu, J.P. Reilly, Correlating the chemical modification of Escherichia coli ribosomal proteins with crystal structure data, J. Proteome Res. 8 (10) (2009) 4466–4478.
[27] L. Xie, X. Wang, J. Zeng, M. Zhou, X. Duan, Q. Li, Z. Zhang, H. Luo, L. Pang, W. Li, G. Liao, X. Yu, Y. Li, H. Huang, J. Xie, Proteome-wide lysine acetylation profiling of the human pathogen Mycobacterium tuberculosis, Int. J. Biochem. Cell Biol. 59 (2015) 193–202.
[28] I.T. Desta, K.A. Porter, B. Xia, D. Kozakov, S. Vajda, Performance and its limits in rigid body protein-protein docking, Structure 28 (9) (2020) 1071–1081.
[29] S. Vajda, C. Yueh, D. Beglov, T. Bohnuud, S.E. Mottarella, B. Xia, D.R. Hall, D. Kozakov, New additions to the ClusPro server motivated by CAPRI, Proteins: Struct., Funct., Bioinf. 85 (3) (2017) 435–444.
[30] D. Kozakov, D.R. Hall, B. Xia, K.A. Porter, D. Padhorny, C. Yueh, D. Beglov, S. Vajda, The ClusPro web server for protein-protein docking, Nat. Protoc. 12 (2) (2017) 255–278.
[31] D. Kozakov, D. Beglov, T. Bohnuud, S. Mottarella, B. Xia, D.R. Hall, S. Vajda, How good is automated protein docking? Proteins: structure, Function, and Bioinformatics. 81 (12) (2013) 2159–2166.