**SUPPORTING INFORMATION**

# Discovery of cryptic largimycins in *Streptomyces* reveals novel biosynthetic avenues enriching the structural diversity of the leinamycin family

**Adriana Becerril,[†,‡,§] Ignacio Pérez-Victoria,[#,§] Suhui Ye, [†,‡] Alfredo F. Braña, [†,‡] Jesús Martín, [#] Fernando Reyes, [#] José A. Salas[†,‡] and Carmen Méndez[*,†,‡]**

[†]Departamento de Biología Funcional e Instituto Universitario de Oncología del Principado de Asturias (I.U.O.P.A), Universidad de Oviedo, Oviedo, Spain

[‡]Instituto de Investigación Sanitaria de Asturias (ISPA), Oviedo, Spain

[#]Fundación MEDINA, Centro de Excelencia en Investigación de Medicamentos Innovadores en Andalucía, Armilla, Granada, Spain.

[§]These authors contributed equally

*** Correspondence:** Carmen Méndez (cmendezf@uniovi.es)

**Strains, culture conditions, plasmids and DNA manipulations**

*S. argillaceus* ATCC 12956 and *Streptomyces canus* ATCC 12646 were used as source of DNA and/or as LRGs producers. SM19 and SM30 (Becerril et al., 2018) and SM30a (tomato paste, 40 g $l^{-1}$; oat flour, 15 g $l^{-1}$; rice molasses, 2 g $l^{-1}$; tap water; pH 4.5) media were used for LRGs production. *Escherichia coli* DH10B (Invitrogen) and *E. coli* ET12567/pUB307 (Kieser et al., 2000) were used as cloning host and as donor strain for conjugation experiments, respectively. When required, media were supplemented with antibiotics at the following final concentrations: apramycin (25 µg $ml^{-1}$), thiostrepton (50 µg $mL^{-1}$), ampicillin (100 µg $mL^{-1}$), kanamycin (50 µg $mL^{-1}$), nalidixic acid (25 µg $mL^{-1}$), and hygromycin (200 µg $mL^{-1}$). Standard procedures were used to transform/conjugate strains (Kieser et al., 2000; Sambrook and Russell, 2001). PCR amplifications were carried out using oligonucleotides in Table S1, Herculase (Stratagene) and 2.5% dimethyl-lsulfoxide (DMSO). Amplicons were purified and sequenced to confirm their identity. Plasmid pUO9090 (M. C. Martín, unpublished results) was used for subcloning. Plasmids pHZ1358 (Sun et al., 2009) and pBSKTTE (this work) were used to generate mutants. Plasmids pEM4T (Menendez et al., 2006), pEM4ATc (Malmierca et al., 2018) and pSETEH (R. Salcedo, unpublished results) were used for expressing genes in *Streptomyces*. pBSKTTE was constructed by cloning a BamHI-XbaI fragment containing *oriT* and *ermEp** from pEM4T into the same sites of pBSKT (Lombó et al., 1999). Bioinformatic analyses were carried out using BlastP (Altschul et al., 1997) and antiSMASH (Weber et al., 2015) programs.

**Generation of mutants**

Several mutants were generated either by inserting a plasmid into the target gene, or by replacing most of the gene by an apramycin resistance cassette that was inserted in the same direction of transcription. To this aim, several plasmids were constructed, introduced by conjugation into *S. argillaceus*, and the corresponding mutants selected by being thiostrepton-resistant (pBSKTTE-based plamids) or apramycin-resistant and thiostrepton-sensitive (pHZ1358-based plasmids). Mutants were genetically confirmed by Southern hybridization or by PCR amplification using specific primers (Figure S1 and Table S1), followed by sequencing the PCR products. Complementation of mutants was carried out by expressing the wild-type copy of the mutated gene *in trans* under the control of the erythromycin resistance promoter (see below). Primers for PCR amplification and DNA sequencing are shown at Table S1. Construction of these plasmids was carried out as follows:

pBSKTTE-AT (to generate *S. argillaceus* ΔlrgG): a 1.26 kb DNA fragment internal to *lrgG* was PCR amplified using primers MutAT831_A y MutAT831_B, digested with BamHI and EcoRI and subcloned into the same sites of pBSKTTE.

pHZ-lrgR3 (to generate *S. argillaceus* ΔlrgR3): a 1.81 kb DNA fragment containing *lrgT1* and the 5'-end of *lrgR3* was amplified using primers TetR21I up and TetR21I rp, digested with BglII and EcoRI and subcloned into the same sites of pUO9090 upstream of the apramycin resistance cassette generating pUO-TetR21I. Also, a 2.05 kb DNA fragment containing the 3'-end of *lrgR3*, *lrgP1* and the 5'-end of *lrgC1* was amplified using oligonucleotides TetR21D up and TetR21D rp and subcloned into the BamHI and EcoRV sites of pUO-

TetR21I, downstream of the apramycin resistance gene generating pUO-TetR21. Then, the whole fragment was rescued with SpeI and subcloned into the XbaI site of pHZ1358.

pHZ-lrgR4 (to generate *S. argillaceus* ΔlrgR4): a 2.04 kb DNA fragment containing *lrgW1* and the 5'-end of *lrgR4* was amplified using primers TetR48I up and TetR48I rp, digested with BglII and EcoRI and subcloned into the same sites of pUO9090 upstream of the apramycin resistance gene, generating pUO-TetR48I. Then a 2.0 kb fragment containing the 3'-end of *lrgR4*, *lrgO* and the 5'-end of *lrgB* was also amplified using oligonucleotides TetR48D up and TetR48D rp, subcloned into pCR-Blunt and rescued as an EcoRV fragment (using one site from the vector) to be subcloned into the same site of pUO-TetR48I in the right orientation, downstream of the apramycin resistance cassette. Finally, the insert was subcloned as a SpeI fragment into the XbaI site of pHZ1358.

pHZ-PH19 (to generate *S. argillaceus* Δorf18): a 1.98 kb DNA fragment containing the 5'-end of *orf18* and *orf16* and *orf17* was amplified using oligonucleotides PH19I up/PH19I rp, digested with EcoRI and PstI, and subcloned into the same sites of pUO9090 upstream of the apramycin resistance cassette, generating pUO-PH19I. Then, a 2.02 kb DNA fragment containing the 3'-end of *orf18*, *lrgT1* and the 5'-end of *lrgR3* was PCR amplified using primers PH19D up/PH19D rp, digested with BamHI and XbaI, and subcloned into the same sites of pUO-PH19I, downstream of the apramycin resistance gene. Finally, the whole insert was rescued as a SpeI fragment and subcloned into the XbaI site of pHZ1358.

pHZ-Trans20 (to generate *S. argillaceus* ΔlrgT1): a 2.02 kb DNA fragment containing the 5'-end of *lrgT1*, *lrgR3* and *lrgP1* was amplified using oligonucleotides orf20TranspI up and orf20TranspI rp, digested with BglII and KpnI and subcloned into the same sites of pUO9090, upstream of the apramycin resistance gene, generating pUO-Trans20I. Then, a 2.13 kb fragment containing the 3'-end of *lrgT1*, *orf18* and *orf17* and the 5'-end of *orf16* was amplified using primers orf20TranspD up and orf20TranspD rp, digested with BamHI and EcoRV, and subcloned into the same sites of pUO-Trans20I, downstream of the apramycin resistance cassette, generating pUO-Trans20. Finally, the whole insert was rescued as a SpeI fragment and subcloned into the XbaI site of pHZ1358.

pHZ-PH22 (to generate *S. argillaceus* ΔlrgP1): a 2.02 kb DNA fragment containing the 5'-end of *lrgP1* and *lrgC1* was amplified using oligonucleotides orf22PHI up and orf22PHI rp, digested with EcoRI and PstI, and subcloned into the same sites of pUO9090 upstream of the apramycin resistance cassette, generating pUO-PH22I. Then, a 1.99 kb DNA fragment containing the 3'-end of *lrgP1*, *lrgR3* and the 5'-end of *lrgT1* was amplified using primers orf22PHD up and orf22PHD rp, digested with BamHI and EcoRV, and subcloned into the same sites of pUO-PH22I downstream of the apramycin resistance gene. Finally, from the resultant construct pUO-PH22 the insert was released with SpeI and subcloned into the XbaI site of pHZ1358.

pHZ-cit23 (to generate *S. argillaceus* ΔlrgC1): a 2.06 kb DNA fragment containing the 5'-end of *lrgC1*, *lrgP1* and *lrgR3* was amplified using oligonucleotides cit23I up and cit23I rp, digested with BglII and HindIII, and subcloned into the same sites of pUO9090 upstream of the apramycin resistance gene, generating pUO-cit23I. Afterwards, a 2.00 kb DNA fragment containing the 3'-end of *lrgC1*, *lrgR1* and the 3'-end of *lrgQ* was amplified

using primers cit23D up and cit23D rp, digested with EcoRV and XbaI, and subcloned into the same sites of pUO-cit23I downstream of the apramycin resistance gene, generating pUO-cit23. Finally, this construct was digested with SpeI and the released fragment was subcloned into the XbaI site of pHZ1358.

pHZ-PH53 (to generate *S. argillaceus* ΔorfF52): a 2.05 kb DNA fragment containing the 5'-end of *orf52*, *orf53* and the 3'-end of *orf54* was amplified using primers orf53PHI up and orf53PHI rp, digested with EcoRI and HindIII, and subcloned into the sites of pUO9090 upstream of the apramycin resistance gene, generating pUO-PH53I. Then, a 2.04 kb DNA fragment containing the 3'-end of *orf52*, *lrgC3* and the 3'-end of *lrgW2* was amplified using primers orf53PHD up and orf53PHD rp, digested with EcoRV and XbaI, and subcloned into the same sites of pUO-PH53I, generating pUO-PH53. Finally, the insert was subcloned as a SpeI fragment into the XbaI site of pHZ1358.

pHZ-cit52 (to generate *S. argillaceus* ΔlrgC3): a 1.99 kb DNA fragment containing the 5'-end of *lrgC3*, *lrgW2* and the 3'-end of *lrgB* was amplified using oligonucleotides Cit52 I up and Cit52 I rp, digested with EcoRI and HindIII, and subcloned into the same sites of pUO9090 upstream of the apramycin resistance cassette, generating pUO-cit52I. Then, a 2.01 kb DNA fragment containing the 3'-end of *lrgC3*, *orf52*, *orf53* and the 3'-end of *orf54* was amplified using primers Cit52 D up and Cit52 D rp, digested with EcoRV and XbaI, and subcloned into the same sites of pUO-cit52I downstream of the apramycin resistance gene, generating pUO-cit52. Finally, the insert was rescued as a SpeI fragment and subcloned into the XbaI site of pHZ1358.

pHZ-Ox49 (to generate *S. argillaceus* ΔlrgO): a 1.99 kb DNA fragment containing the 3'-end of *lrgW1*, *lrgR4* and the 5'-end of *lrgO* was amplified using oligonucleotides Ox49I up and Ox49I rp, digested with EcoRI and PstI and subcloned into the same sites of pUO9090 upstream of the apramycin resistance cassette, generating pUO-Ox49I. Then, a 2.0 kb DNA fragment containing the 3'-end of *lrgO*, *lrgB* and the 5'-end of *lrgW2* was amplified using oligonucleotides Ox49D up and Ox49D rp, digested with EcoRV and XbaI and subcloned into the same sites of pUO-Ox49I, downstream of the apramycin resistance gene, generating pUO-Ox49. Finally, the insert was rescued as a SpeI fragment and subcloned into the XbaI site of pHZ1358.

**Plasmid constructs for gene expression**

Several plasmid constructs were generated to overexpress specific *lrg* genes under the control of the erythromycin resistance promoter *ermEp**, using plasmids pEM4T, pEM4ATc or pSETEH. These plasmids were introduced by conjugation into *S. argillaceus* wild type or mutant strains and the corresponding recombinant strains selected by being thiostrepton-resistant (pEM4T-based plamids) or apramycin-resistant (pSETEH- and pEM4Tc-based plasmids). Strains were genetically confirmed by PCR amplification using specific primers (Table S1), followed by sequencing the PCR products. Construction of these plasmids was carried out as follows:

pEM4T-R1 (to overexpress *lrgR1*): a 700 bp DNA fragment containing *lrgR1* was amplified using oligonucleotides Reg24 up and Reg24 rp, digested with BamHI and subcloned in the right orientation into the same site of pEM4T.

pEM4T-R2 and pEM4ATc-R2 (to overexpress *lrgR2*): a 1.11 kb DNA fragment containing *lrgR2* was PCR amplified using oligonucleotides Reg831_A and Reg831_B, digested with BamHI and EcoRI and subcloned into the same sites of pEM4T and pEM4ATc, respectively.

pSETEH-cit23 (to complement *S. argillaceus* ΔlrgC1): a 1.39 kb DNA fragment containing *lrgC1* was amplified using oligonucleotides ermECit23 up and ermECit23 rp, digested with NheI and SpeI and subcloned in the right orientation, into the XbaI site of pSETEH.

pSETEH-cit52 (to complement *S. argillaceus* ΔlrgC3): a 1.30 kb DNA fragment containing *lrgC3* was amplified using oligonucleotides ermECit52 up and ermECit52 rp, digested with NheI and SpeI and subcloned in the right orientation, into the XbaI site of pSETEH.

pSETEH-Ox49 (to complement *S. argillaceus* ΔlrgO) : a 1.31 kb DNA fragment containing *lrgO* was amplified using oligonucleotides ermEOx49 up and ermEOx49 rp, digested with NheI and SpeI and sucloned in the right orientation into the XbaI site of pSETEH

**UPLC Analysis and Purification of largimycins**

Strains were grown in a two-step culture method (Fernández et al. 1998). A seed culture was prepared in 50-ml Erlenmeyer flasks containing 10 mL of Trypticase Soy Broth (TSB) medium, incubated for 48 hours at 30ºC and 250 rpm. This culture was used to inoculate 250 mL Erlenmeyer flasks each containing 50 mL of SM30a. Production of LRGs was monitored daily for 8 days. Culture samples of 1 mL were extracted with an equal volume of ethyl acetate containing 1% formic acid, with shaking for 60 min. Organic extracts were dried under vacuum, and residues were dissolved in methanol to evaluate LRGs production by UPLC. Analyses were performed by reversed-phase chromatography on an Acquity UPLC equipment with a BEH C18 column (1.7 mm, 2.1 x 100 mm; Waters, Milford, MA, USA) with acetonitrile and 0.1% trifluoroacetic acid (TFA) in water as eluent. Samples were eluted with 10% (v/v) acetonitrile for 1 min, followed by a linear gradient from 10 to 100% acetonitrile over 7 min at a flow rate of 0.5 mL $min^{-1}$ and a column temperature of 35ºC. Detection and spectral characterization of peaks were carried out with a photodiode array detector and Empower software (Waters). Chromatograms were extracted at 300 or 330 nm.

For purification purposes, *S. argillaceus* WT-R2, *S. argillaceus* ΔlrgO or *S. canus* WT-R2 strains were grown by the two-step culture method mentioned above, but using five 2-liter Erlenmeyer flasks, each containing SM30a medium (400 mL) in the production step. These cultures were incubated for 3 days (5 days for LRG O1). To purify LRG A1, LRG A2, LRG A3 and LRG A4, cultures were centrifuged and filtered, and applied to a solid-phase extraction cartridge (Sep-Pak Vac C18, 10 g, Waters). The retained material was eluted using a linear gradient from 0 to 100% methanol in 0.05% (v/v) TFA in water for 55 min, at 5 mL $min^{-1}$. Fractions were taken every 5 min and analyzed by UPLC. Fractions containing the desired compounds were evaporated *in vacuo* and dissolved in a small volume of a mixture of DMSO and methanol. To purify LRG O1, cultures were extracted with ethyl acetate plus formic acid, the organic extract was dried down under vacuum, and residues were dissolved in a small volume of DMSO:methanol (1:1). Afterwards, in all cases the desirable products were finally purified by preparative HPLC using a SunFireC18 column (10 mm, 10 x 150 mm, Waters) at a flow

of 5 mL/min with mixtures of acetonitrile or methanol and 0.05% TFA in water, in isocratic conditions optimized for each compound. The purification procedure afforded LRG A1 (0.8 mg), LRG A2 (0.9 mg), LRG A3 (0.3 mg), LRG A4 (0.9 mg) and LRG O1 (0.8 mg) all as amorphous yellowish solids.

**Spectroscopic analysis of largimycins and molecular modelling**

Structural elucidation of each compound was carried out by ESI-TOF mass spectrometry and NMR spectroscopy. HRMS spectra were collected by LC-MS analyses using an Agilent 1200RR HPLC equipped with a SB-C8 column (2.1 × 30 mm, Zorbax) and coupled to a Bruker maXis Spectrometer. Chromatographic and ionization conditions were identical to those previously described (Pérez-Victoria et al., 2016; Martín et al., 2014). UV/vis (DAD) spectra were also collected in the same chromatographic analyses. NMR spectra were recorded in DMSO-$d_6$ or $CD_3OD$ at 24ºC on a Bruker AVANCE III-500 MHz (500 and 125 MHz for 1H and 13C NMR, respectively) equipped with a 1.7 mm TCI MicroCryoProbe™, using the residual solvent signal as internal reference ($\delta_H$ 2.50 and $\delta_C$ 39.5 for DMSO-$d_6$, $\delta_H$ 3.31 and $\delta_C$ 49.0 for $CD_3OD$). Molecular modelling was used in combination with NMR data to determine the relative configurations using 3D structural models generated with Chem3D Pro 12.0 starting from the reported X-ray structure of LNM E2 (Huang et al. 2015). The structures were first constructed to roughly satisfy the observed $^3J_{HH}$ and key NOESY correlations and then submitted to energy-minimization by molecular mechanics with the MM2 force field using as gradient convergence criteria an RMS value of 0.001. Molecular modelling images (Figure S9) were generated with PyMOL (DeLano, 2002).

**Structure elucidation of largimycins**

Largimycin A1 (**1**) was assigned the molecular formula $C_{33}H_{40}N_4O_{14}S_3$ based on the observed ion $[M+H]^+$ at *m/z* 813.1781 (calcd. for $C_{33}H_{41}N_4O_{14}S_3^+$ = 813.1776, $\Delta$m = 0.6 ppm) alongside their corresponding isotopic pattern, indicating 16 degrees of unsaturation. The connectivity of **1** was determined by detailed 1D ($^1$H) and 2D NMR (COSY, HSQC and HMBC) spectroscopic analyses. Interestingly, the NMR spectroscopic data of **1** (Table S3 and Figure S4) resembled those reported for LNM E2 and LNM E3 (Huang et al., 2015), confirming that LRGs are related to LNMs. Interpretation of the HSQC and HMBC spectra revealed the presence of 13 quaternary carbons (including one $\alpha$, $\beta$ unsaturated ketone at $\delta_C$ 201.3, six ester/amide carbonyls in the range $\delta_C$ 169-175, three $sp^2$ carbons at $\delta_C$ 139.1, 152.6 and 153.5 and three $sp^3$ carbons at $\delta_C$ 49.0, 51.3 and 61.7), 9 methines (including five olefinic/aromatic carbons, one oxygenated methine, two methines likely corresponding to the $CH_\alpha$ of two amino acid moieties and a final aliphatic methine at $\delta_C$ 45.5 and $\delta_H$ 3.94 likely bonded to a sulfur atom), 8 aliphatic methylenes (including a methylene within an oxirane ring resonating at $\delta_C$ 49.0 and $\delta_H$ 3.31, 2.86) and finally 3 methyl groups (two of them corresponding to two *N*-acetyl groups). Analysis of COSY correlations identified different spin systems (Figure 4B). Two of them are comprised by the $\alpha$ and $\beta$ protons of two amino acids which, based on the proton and carbon chemical shifts of the corresponding positions, were assigned to two cysteine moieties which turned out to be *N*-acetylated (CysNAc), based on the key HMBC correlations observed between these $\alpha$ protons and the carbonyl of the acetyl groups (Figure 4B ). Another spin system, comprising H-10 to H-13, contains four olefinic protons

corresponding to *Z* and *E* double bonds, as indicated by the measured coupling constants. The characteristic proton and carbon chemical shifts of those methines (positions 10 to 13) are indeed very similar to the chemical shifts reported for the olefinic methines of LNMs E2, E3 and E4 (Huang et al., 2015) suggesting they corresponded to identical structural motifs. The *E* double bond of this spin system is conjugated with a ketone, as indicated by the key long-range correlation between H-11 and the C-9 carbonyl, as found in known LNMs. On the *Z* double bond end, this spin system is conjugated with an aromatic heterocycle, as revealed by the key HMBC correlations between H-15 and C-13, C-14 and C-16. The chemical shifts at the heterocycle discard a possible thiazole ring in favor of an oxazole. Thus, the extended $\pi$-system between positions 9 and 16 of **1** is equivalent to the one contained in known LNMs, just differing in the divalent heteroatom present in the azole heterocycle, being sulfur in LNMs and oxygen in **1**. Another spin system, comprising H-7 and H-8, is connected to the extended $\pi$-system via C-8, as demonstrated by the long-range correlations from H-10 to C-8 and from H-8 to C-9. Interestingly, the proton and carbon chemical shifts at positions 7 and 8 were also remarkably similar to those reported for LNMs E2 and E3, indicating the sulfur substitution at C-7, thus accounting for the third sulfur atom in the molecular formula, the other two being localized in the S-conjugated CysNAc units. It was thus evident that **1** likely contained a tetrahydrothiopyran ring analogous to that found in the previously mentioned LNMs. The HMBC correlations of the methyl protons H-20 with C-5, C-6 and C-7 confirmed the expected substitution position of this methyl group within the saturated heterocycle. The COSY correlations between H-4 and H-5 combined with the HMBC correlations from H-5 to C-3, C-6 and C-7, together with those from H-4 to C-3 unambiguously closed the tetrahydrothiopyran ring. The second substituent at C-6 was found to correspond to the sulfur atom on one of the CysNAc units based on the long-range correlation observed between the $\beta$ methylene protons of the amino acid and C-6. The chemical shifts of the isolated methylene at position 2 also resembled those reported for LNMs E2 and E3. The HMBC correlations of these methylene protons, H-2, with the carbonyl C-1 and the quaternary carbon C-3 provided the connectivity path for the extension of the macrocycle, while the long-range correlation from H-2 to C-21 provided the linkage between the C-3 side chain substituent and the macrocycle. Such side chain contains only two observable protons belonging to a methylene, H-22, within an oxirane ring. In the edited HSQC spectrum, this methylene cross peaks show a characteristic and diagnostic phase, opposite than expected and equal to methyl and methine groups, due to the partial $sp^2$ character of this carbon within the epoxide ring and its associated $^1J_{CH}$, much larger than the 125 Hz the spectral acquisition is optimized for, thus rendering the unexpected phase for this methylene. Additional evidence for the presence of the epoxide functional group is provided by the value of the corresponding geminal coupling between the H-22 methylene protons, 5.5 Hz, characteristic for oxiranes. The HMBC correlations from H-22 to C-3, C-21 and C-23 rendered the full connectivity of the C-3 side chain. On the other hand, to establish how the macrocycle extends after the oxazole ring it was essential to find a proton displaying a long-range correlation with the quaternary carbon C-16 of the aromatic heterocycle. H-18, corresponding to the only observed oxygenated methine, correlates with such carbon in the HMBC spectrum and also with C-17, a $sp^2$ carbon at $\delta_C$ 153.5, a chemical shift compatible with an oxime functional group. H-18 belongs to another small spin system involving also methylene H-19,

which also displays a long-range correlation with the putative oxime carbon C-17. This methylene was found to be also bound to the sulfur atom of the second CysNAc unit based on the long-range correlations involving the thioether bridge, from H-19 to the β carbon of the amino acid and from the β methylene protons of the amino acid to C-19. Having all the NMR signals accounted for, it was confirmed the mentioned oxime functional group in order to meet with the number of heteroatoms in the molecular formula of **1**. To also satisfy all the degrees of unsaturation, it was necessary to close the macrocycle via an ester bond between the oxime oxygen and carbonyl C-1, finally providing the full connectivity of **1** (Figure 4B), which resembles that of LNMs E2 and E3. The oxime double bond was assigned a *Z* stereochemistry on the basis of the observed $\delta_C$ 73.2 for C-18 which would be 2-4 ppm smaller in case of the *E* geometric isomer (Hawkes et al., 1974), as indicated by comparison with model compounds (Kajiro et al., 1999) and empirically-based prediction of $^{13}C$ NMR chemical shifts (Elyashberg et al., 2009) (a detailed explanation is provided later). To determine the relative configuration of the chiral centers in LRG A1, a 3D structural model of **1** was generated starting from the reported X-ray structure of LNM E2 (Huang et al., 2015). The model was first constructed to roughly satisfy the observed $^3J_{HH}$ (Table S3) and key NOESY correlations (Figure 4C) before its energy-minimization. The resulting minimized structure (Figure S9A) perfectly accounted for the observed NMR couplings and NOEs. Not surprisingly the pattern of NOESY correlations involving the protons in the macrocycle backbone is identical to that reported for LNM E2 and E3. The relative configuration at the epoxide chiral center, C-21, in the C-3 side chain was established on the basis of the strong NOESY correlation between $H_a$-22 and $H_b$-4 and the weaker NOESY correlation between $H_a$-22 and the methyl protons H-20. Modelling proved that such NOEs would not be observed in the corresponding epimer of **1** at C-21 (data not shown). Interestingly, the determined relative configuration at C-21 matches that for this chiral center in LNM itself (Noriaki and Shimizu, 1993). The absolute configurations at C-3, C-18 and the $C_\alpha$ of both S-conjugated CysNAc units were assigned to be the same as those found in LNM, L-Thr and mycothiol respectively, on the basis of biosynthetic and phylogenetic (Pan et al., 2017) arguments as explained in the main text, thus providing the full absolute stereochemistry of **1** (Figure 4A).

Largimycin A2 (**2**) was assigned the molecular formula $C_{28}H_{31}N_3O_{11}S_2$ based on the observed ions $[M+NH_4]^+$ at *m/z* 667.1744 (calcd. for $C_{28}H_{35}N_4O_{11}S_2^+$ 667.1738, $\Delta$m = 0.9 ppm) and $[M+H]^+$ at *m/z* 650.1476 (calcd. for $C_{28}H_{32}N_3O_{11}S_2^+$ = 650.1743, $\Delta$m = 0.5 ppm) alongside their corresponding isotopic patterns, indicating 15 degrees of unsaturation. The connectivity of **2** was determined by detailed 1D ($^1H$) and 2D NMR (COSY, HSQC and HMBC) spectroscopic analyses assisted by comparisons with the NMR data of **1**. As expected, the NMR spectroscopic data of **2** (Table S4 and Figure S5) partially resembled those of **1** and interestingly also those reported for LNM E4 (Huang et al., 2015). Analysis of COSY correlations identified the same number and type of spin systems found in **1** but one of the CysNAc related spin systems which was missing (Figure 4B), in agreement with the number of sulfur atoms in the molecular formula of **2**, one less than in **1**. The methylene at position 19, which was S-conjugated with a CysNAc unit in **1**, now displays in the edited-HSQC spectrum of **2** the characteristic diagnostic phase already mentioned for oxirane rings indicating an epoxide functional group involving C-18/C-19. The pattern of HMBC correlations of **2** also matches that found for **1** but

the important exception of the β methylene protons of the CysNAc unit which now displayed a long-range correlation with C-7 rather than C-6 (as found in **1**). Such distinctive feature unambiguously determined the presence of a tetrahydrothiophene ring, involving C-3 to C-6, in the structure of **2** in contrast with the tetrahydrothiopyran ring, involving C-3 to C-7, present in **1**. The determined connectivity of **2** thus resembles that of LNM E4 (Huang et al., 2015). The oxime double bond was assigned a *Z* stereochemistry based on the observed $\delta_C$ of 49.9 for C-18 and the same arguments employed for **1** (see detailed explanation later). The relative stereochemistry of **2** could be determined by analysis of the observed $^3J_{HH}$ (Table S4) and key NOESY correlations (Figure 4C) combined with molecular modelling (Figure S9A) as previously described for **1**. Likewise, the absolute configurations of the key chiral centers at C-3, C-18 and the $C_\alpha$ of the CysNAc unit were assigned to be the same as **1** based on their shared biosynthetic origin, thus providing the full absolute stereochemistry of **2**.

Largimycin A3 (**3**), was assigned the molecular formula $C_{23}H_{23}ClN_2O_8S$ on the basis of the observed ion $[M+H]^+$ at *m/z* 539.0937 (calcd. for $C_{23}H_{24}N_2O_8S^+$ = 539.0936, Δm = 0.2 ppm) alongside its corresponding isotopic pattern (Figure S6). Based on this molecular formula (indicating 13 degrees of unsaturation), its shared biosynthetic origin with LRGs A1 and A2, and the reported structure for LNM E4 (Huang et al., 2015) we have proposed a tentative chemical structure for LARG A3 (**3**) (Figure 4A).

Largimycin A4 (**4**) was assigned the molecular formula $C_{23}H_{24}N_2O_9S$ based on the observed ion $[M+NH_4]^+$ at *m/z* 522.1544 (calcd. for $C_{23}H_{28}N_3O_9S^+$ = 522.1541, Δm = 0.6 ppm) alongside its corresponding isotopic pattern, indicating 13 degrees of unsaturation. The single sulfur atom in the molecular formula suggested the absence of any CysNAc moiety. The connectivity of **4** was established after detailed 1D ($^1$H) and 2D NMR (COSY, HSQC and HMBC) spectroscopic analyses further assisted by comparisons with the NMR data of **1** and **2**. The NMR spectroscopic data of **4** (Table S4 and Figure S7) are remarkably similar to those of **1** but two important differences. On the one hand, the resonances corresponding to positions 18 and 19 essentially match those found for **2**, indicating the presence of the same oxirane ring. On the other hand, the observed $\delta_C$ of 72.8 for C-6 clearly indicates a hydroxy substitution of the quaternary C-6 carbon instead of the CysNAc-S-conjugate substituent present in **1**. In this manner, the connectivity of **2** could be easily established. The pattern of COSY and HMBC correlations of **4** (Figure 4B) corroborated the assignments**.** The oxime double bond was assigned a *Z* stereochemistry based the observed $\delta_C$ 50.7 for C-18 and the same arguments already employed for **1** and **2** (see detailed explanation later). Not surprisingly, the NOESY correlations observed for **4** (Figure 4C) also match those found in **1**, indicating the expected identical relative configuration of both compounds. As already indicated, the cluster *scan* and *lrg* belong to the same clade VII of the previously mentioned phylogenetic classification (Pan et al., 2017), thus the absolute configurations of the chiral centers at C-3 and C-18 were assigned to be the same as **1**, providing the full absolute stereochemistry of **4**.

Largimycin O1 (**5**) was assigned the molecular formula $C_{29}H_{39}N_3O_8S_2$ based on the observed ion $[M+H]^+$ at *m/z* 622.2255 (calcd. for $C_{29}H_{40}N_3O_8S_2^+$ = 622.2251, Δm = 0.6 ppm) alongside its corresponding isotopic pattern, indicating 12 degrees of unsaturation. The connectivity of **5** was established after detailed 1D ($^1$H) and

2D NMR (COSY, HSQC and HMBC) spectroscopic analyses further assisted by comparisons with the NMR data of **1**, **2**, **4**, LNM E1 (Huang et al., 2015) and GNM B (Pan et al., 2017). The NMR spectroscopic data of **5** (Table S5 and Figure S8) confirmed a largimycin related structure but showed some important differences compared to the previously elucidated LRGs A1, A2 and A4. The edited HSQC spectrum clearly revealed the presence of three olefinic protons additionally to those found in **1**, **2** and **4**. Two of these protons correspond to a $sp^2$ methylene which turned out to be conjugated with the characteristic extended $\pi$-system between positions 9 and 16 of LRGs, as indicated by the long-range correlations from H-10 to both C-9 and C-21 (Figure 4B). Such olefinic exomethylene unit is identical to the one displayed by GNMs and WSMs (Pan et al., 2017) and differs from the ketone at C-9 found in **1**-**4** and known LNMs. The third proton, H-7, according to the key COSY and HMBC correlations (Figure 4B) was found to be contained in the C-1 to C-8 fragment, a substructural motif identical to that found in GNM B (Pan et al., 2017) (including the methyl substituent at C-6), as reflected by the similar NMR carbon chemical shifts of both GNM B and **5** in these positions. Interestingly, C-1 to C-7, the C-6 methyl and the side chain substituent at C-3 also showed very similar chemical shifts to those of LNM E1 (Huang et al., 2015), indicating that the C-3 alkyl branch in **5** and LNM E1 are identical, as reflected by the key COSY and HMBC correlations involving H-22 and the H-23 methyl group (Figure 4B). After having localized the C-3 thiol substituent in this manner, the second sulfur atom contained in the molecular formula of **5** was determined to belong to a CysNAc-S-conjugate substituent at C-19 as already described for **1**. Interestingly, the oxygenated methine proton H-18 is coupled not only to the methylene H-19, as found for **1**-**4**, but also to another methine proton, H-17, which is coupled to the exchangeable proton 17-NH. Thus, the methine at position 17, corresponded to the $\alpha$ position of an amino acid (also reflected in its chemical shifts), another remarkable difference compared to **1**-**4**, where C-17 is the quaternary carbon of the oxime functional group. Long-range correlations of H-17 and H-18 with C-16 confirmed the expected bond between C-17 and the heterocycle (Figure 4B). Finally, the key HMBC correlation between H-17 and 17-NH with the carbonyl C-1 closed the macrocycle with an amide bond, providing the full connectivity of **5**. LRG O1 is a macrolactam as are the known LNMs, GNMs and WSMs (Figure 1). The observed NOESY correlations observed for **5** (Figure 4C) are analogous to those found in GNM B, confirming the expected double bonds stereochemistry for **1**. Based on its shared biosynthetic origin, the absolute configuration at C-3 and C-18 were assigned to be the same as **1**, while the membership of cluster *lrg* to the mentioned clade VII (Pan et al., 2017) also allows assigning to C-22 the same absolute configuration determined for LNM E1 (Huang et al., 2015) and to C-17 the same stereochemistry as the $\alpha$ position of L-Thr, thus providing the full absolute stereochemistry of **5**.

Determination of the oxime double bond stereochemistry in LRGs A1, A2 and A4 has relied on the $^{13}C$ NMR chemical shift of C-18 in each largimycin. The $^{13}C$ chemical shift at the $\alpha$ position in oximes is an unambiguous and convenient probe for determining the configuration of the oxime double bond (Hawkes et al., 1974). This is based on the fact that a *syn* $\alpha$ carbon resonates at upper field than the equivalent *anti* $\alpha$ carbon. Thus, for discriminating *E*/*Z* oxime stereoisomers, a simple comparison of the $\delta_C$ at the relevant $\alpha$ carbon is enough to identify the correct double bond configuration. Such approach is straightforward when the

$^{13}C$ NMR data is available for both stereoisomers. However, in the case of largimycins we only have the experimental data for one stereoisomer. To identify whether it corresponds to the *E* or the *Z* isomer we have compared the experimental $\delta_C$ at C-18 in each largimycin with the empirically-based predicted chemical shift. The application of empirical methods of $^{13}C$ NMR chemical shift prediction has already been successfully employed for determining relative stereochemistry (Elyashberg et al., 2009). To validate this approach, *R*-2-acetoxyiminoindan was employed as model compound. This molecule contains an oxime ester functional group conjugated with an aromatic ring, it also displays an oxygenated methine in the aliphatic $\alpha$ carbon (with respect to the oxime) and experimental $^{13}C$ NMR data is available for both *E* and *Z* oxime stereoisomers (Kajiro et al., 1999). Chemical shift prediction with ACD/Labs predictor shows an excellent agreement between the experimental and the predicted $\delta_C$ at the aliphatic $\alpha$ carbon, thus validating the use of comparisons between experimental and empirically based predicted chemical shifts for establishing the oxime double bond stereochemistry in largimycins (Figure S9B). Next, the $\delta_C$ at C-18 for the possible *E* and *Z* oxime isomers in each largimycin was likewise calculated with ACD/Labs predictor. Comparison of the predicted value with the experimental one unambiguously determined the *Z* configuration for the oxime double bond in all largimycins (Figure S9B).

**Alignments of Lrg proteins with Lnm proteins**

LrgT1 *vs* LnmY

```
Lnm  25   LILVTLLMAELLIQVDQTIVNVALPFIQRDLDFTESGLPWVVNAYGLLYGGLLPLGGRIG  84
          L+LV + MA+L++ +D TIVNVALP +Q  L F+   L WVVNAY L +GGLL LGGR+G
Lrg  23   LMLVVIAMAQLMVILDATIVNVALPSVQTSLGFSTQNLSWVVNAYILAFGGLLLLGGRVG  82

Lnm  85   DLFGRRRVLIIGVTVFTLASVVCGFSTDPTVMVIARAVQGMGGALTAPVVLSLIITSFEE  144
          DL GRR   I GV +FTL S++ G + +   ++  R +QG G A+  P VLSLI T F
Lrg  83   DLLGRRPAFIGGVLLFTLGSLLGGLAQNSGWLLGMRVIQGAGAAIVVPTVLSLIATGFPT  142

Lnm  145  GPSRQRAFALWGSAQAAGALFGLIVGGLLTSGPGWEFSFFVAVPIGALVVGLAATTI--K  202
             R RAFA++     AG   GL+ GGLLT    W +   V VPIG + + LAA     K
Lrg  143  ERERNRAFAVFAGVSGAGGAIGLVAGGLLTEWASWRWVLLVNVPIG-VALALAAPLFIGK  201

Lnm  203  ESRADERPSIDVAGAVTFTAALLMTVFAFLDTQRS-WTSPLRGGLLVVAAVLFAVFLAVE  261
            R++ R   DV GA+T T  + + V+ F+    S W      G  + A VL   F+ +E
Lrg  202  TPRSEGR--FDVGGAITSTGGVALLVYGFIHASDSGWREATTIGSFIGAVVLLVAFVFIE  259

Lnm  262  RRQGHPFVPLQVFRLRNTSGALAIMAIFGATQMSYFFFVTLYLQEILGFSALQTGLAYLP  321
           R   P +PL++F  RN SG  A+      + +  FFF+TL+ Q +LG+S L+TGLA+LP
Lrg  260  SRTPQPIIPLRLFASRNRSGIYALGVAMMGSLIGMFFFLTLFFQNVLGYSPLKTGLAFLP  319

Lnm  322  LIATLLVFAQVCMKTVARVGLTRMLMTGLLCLGLGMGWLGLAASSGSFVGTVLGPTIISG  381
          L  +++V + V M+ + ++G    L+ G L +   + WL   ++  +++G +LGP ++ G
Lrg  320  LSVSIIVVSGVMMQLIPKIGQRLPLVAGTLLITGSLIWLSAISADSAYLGDLLGPMLLYG  379

Lnm  382  IGLGLTMMPAGTLATTGVDPADAGAVSGLFNTAILVGGSLGLAVLTTVT-QAVGGLGGTH  440
           G G+  MP   +  + V+  D GA SGL N    +GGSL LAV+ TV   A  G  G
Lrg  380  FGAGMVFMPMTMVGVSDVEMQDTGAASGLLNATQQMGGSLSLAVIITVYGAATNGATGDP  439

Lnm  441  GYTVAFLCSAGLAAA  455
           + +A   SAG   A
Lrg  440  AHVLAKGASAGFLVA  454
```

## LrgR1 *vs* LnmO

```
Lnm  16   LRDFLRDQSPHTTTTMVRRNQSAYSCGGQDRNVYFLESGRLKTVMFSRSGKECLLRIHTP  75
          LRD +        T  + R ++ YS G  D ++Y +E G++K V  S  GK CLL +
Lrg  13   LRDRMVGSGRSAPTVRLVRGENVYSSGQSDNSLYLVEEGQVKIVSGSLDGKRCLLSVCVE  72

Lnm  76   GSIFGELCSLGGIREESAIAMRDSVVHRMSYEHFLSSITEAGIVDEFINHLTRRLAEQQQ  135
          G  FGEL  L G+R E+A AM+ SV+ ++S  H  + + +    +EF  HL  +L+ QQ+
Lrg  73   GEFFGELGILQGMRAETATAMKRSVLRKLSAAHVAAVLQDLENAEEFALHLMEQLSWQQR  132

Lnm  136  SITHLVTVDSEQRLGEILLDLARKLGRGDGRRPHIAERITHEELAGMVGTTRSRVGHFLK  195
           I ++VT++ EQRL   LLDLARK+GR  G    +  RITHEEL+ MVGTTRSRVG FLK
Lrg  133  LIANMVTMNCEQRLAVTLLDLARKVGRRQGAELRMERRITHEELSEMVGTTRSRVGFFLK  192

Lnm  196  GFRDRGLVEVTRESHLIIDQSRLAAYL  222
          F D GL+  +  + L + ++RLA Y+
Lrg  193  RFSDDGLLVHSGSAQLTVHETRLARYV  219
```

## LrgC2 *vs* LnmA and LrgC2 *vs* LnmZ

### *LrgC2 vs LnmA*

```
Lnm  25   FAELRETDPLARVRLPYGGEGWMVTRYDDVRAANSDPRFSR-AQIGEDTPRTTPLAR---  80
          ++++R    +  +R P G    +V RY + R   +D RF++   +  D  R     +
Lrg  32   YSDMRAEGVVHTIREPNGLHRRLVLRYAEARDVMNDARFAKDPGLAWDQLRDAGYVKGER  91

Lnm  81   --RSDTILSL---DPPEHTRLRRLLSKAFTARRMGAMQSWLEELFAGLLDGV--ERTGHP  133
            R+D +  L   DPP+HTRLR+L+SKAFT RRM AM+  + E+   LLDG+  +RT
Lrg  92   DNRADYLYHLVNTDPPDHTRLRKLISKAFTNRRMEAMRPRVREIAEQLLDGLAGQRT---  148

Lnm  134  ADIVRDLAQPFTIAVICRLLGVPYEDRGRFQHWSEVIMSTTAYSKEEAVSAD---ASIRA  190
           D++ D A P    VIC +LGVP  DR  F+ W+  +++      E A++     A++++
Lrg  149  VDLIEDYAHPLATTVICEILGVPNADRENFRIWATAMLTAPDAVVEGALTPQEGYAAMQS  208

Lnm  191  YLADLVSARRAAPHD-------------DLLGVLVSARDDDDRLTEDELITFGVTLLVAG  237
          +  +L++ RR    D             D++  L+ ARD+ +RLTE E+++  + LL AG
Lrg  209  FFTELLARRREELKDLAETDEDTSDQQPDVITGLIVARDEGNRLTEIEMVSTAMLLLSAG  268

Lnm  238  HETSAHQLGNMVYALLTHEDQLSLLREQPELLPRAVEELLRFVPLGNGVGNARIALEDVE  297
           E + + + N   AL  + +Q +LLR +P+L+  AVEE LR+ P    +   R++ EDVE
Lrg  269  QEPTVNLIANGTLALFDNPEQFALLRAKPDLVASAVEEFLRYDPPVE-LSTMRVSTEDVE  327

Lnm  298  LSGGTVRAGEGVVAAAVNANRDPRAFDDPDRLDITREKNPHLAFGHGAHYCLGAQLARME  357
          ++G  + AG  V  +  + +RD R F++PD+LDITR  NPHLAFG+G H+CLGA LAR+E
Lrg  328  VAGTVIPAGSVVTVSIASTSRDERQFENPDQLDITRTDNPHLAFGYGLHHCLGAPLARIE  387

Lnm  358  LRVAIGGLLERFPGLRLAVPADQVEWKTGGLFRGPQRLPI  397
           + AI  L+ R+PG+ L+   D + W+   + RG   LP+
Lrg  388  GQEAIAALVGRYPGISLSGTRDDLRWRPTRIMRGLVELPV  427
```

### *LrgC2 vs LnmZ*

```
Lnm  29   YAKLFEDGDPIRVQLPFGEPAWLVTRYDDARFVLTDRRFSRHLATQRDEPRMTPRAVPE-  87
          Y+ +  +G    ++ P G    LV RY +AR V+ D RF++      D+ R       E
Lrg  32   YSDMRAEGVVHTIREPNGLHRRLVLRYAEARDVMNDARFAKDPGLAWDQLRDAGYVKGER  91

Lnm  88   --------SILTMDPPDHTRLRTLVSKAFTPRRIESKRAWIGELAAGLVADMKAGGAPAE  139
                  ++  DPPDHTRLR L+SKAFT RR+E+ R  + E+A  L+ D  AG    +
Lrg  92   DNRADYLYHLVNTDPPDHTRLRKLISKAFTNRRMEAMRPRVREIAEQLL-DGLAGQRTVD  150

Lnm  140  LVGSYALAIPVTVICELLGVPEDDRTRLRGWCDAALST------GELTDEECVQSFMDLQ  193
          L+  YA  +  TVICE+LGVP  DR   R W  A L+       G LT +E    +  +Q
Lrg  151  LIEDYAHPLATTVICEILGVPNADRENFRIWATAMLTAPDAVVEGALTPQE---GYAAMQ  207

Lnm  194  KYFEDLVKERRAEPRD-------------DLTSALIEARDAHDRLAEPELIGLCISILIG  240
           +F +L+  RR E +D             D+ + LI ARD  +RL E E++   + +L
```

```
Lrg  208  SFFTELLARRREELKDLAETDEDTSDQQPDVITGLIVARDEGNRLTEIEMVSTAMLLLSA  267

Lnm  241  GFETTASEISSFVHVLQQRRELWTRLCADPEAIPAAVEELLRFVPFAANGISPRYALEDM  300
          G E T + I++    L    E +  L A P+ + +AVEE LR+ P      + R + ED+
Lrg  268  GQEPTVNLIANGTLALFDNPEQFALLRAKPDLVASAVEEFLRYDP-PVELSTMRVSTEDV  326

Lnm  301  TVGGVLVREGEPVIVDTSAVNRDGLVFDNADEVVIDRADNRHMVFGHGAHHCLGAHLARV  360
           V G ++  G  V V  ++ +RD   F+N D++ I R DN H+ FG+G HHCLGA LAR+
Lrg  327  EVAGTVIPAGSVVTVSIASTSRDERQFENPDQLDITRTDNPHLAFGYGLHHCLGAPLARI  386

Lnm  361  ELQEALKALVEGMPGLRLSG---DVEWKADMIIRA  392
          E QEA+ ALV   PG+ LSG   D+ W+   I+R
Lrg  387  EGQEAIAALVGRYPGISLSGTRDDLRWRPTRIMRG  421
```

## LrgR2 *vs* LnmO

```
Lnm  7    TEPTGFDARLRDFLRDQSPHTTTTMVRRNQSAYSCGGQDRNVYFLESGRLKTVMFSRSGK  66
          T+   F  RLR     +   T T  V R    Y  GG+DR++Y +E+G++KT+M + SGK
Lrg  28   TDSVRFRTRLRALCATRLHDTPTLTVPRRGHVYVDGGRDRHIYVIETGQVKTLMTTGSGK  87

Lnm  67   ECLLRIHTPGSIFGELCSLGGIREESAIAMRDSVVHRMSYEHFLSSITEAGIVDEFINHL  126
           CLL I   G + GEL  L   R ++A+A+  + + R+  + FLS + E G++ E++ HL
Lrg  88   RCLLSISGAGDVLGELGLLASERSDTAVALEPATLRRIPGDRFLSVLAEEGLLPEYVRHL  147

Lnm  127  TRRLAEQQQSITHLVTVDSEQRLGEILLDLARKLGRGDGRRPHIAERITHEELAGMVGTT  186
            R+ EQQ+ I  +VT++ E+RL   LL L++++G   G    I  RIT EELA MVGTT
Lrg  148  GERVIEQQKIIVDMVTMECERRLAARLLLLSQQIGTRHGPLVLIRARITQEELAEMVGTT  207

Lnm  187  RSRVGHFLKGFRDRGLVEVTRESHLIIDQSRLAAYLESR  225
          RSRVG FLK FR+ GLV++ + S L++D  RL  Y+E R
Lrg  208  RSRVGLFLKRFREAGLVDMAQGS-LVVDDRRLGDYIEGR  245
```

## LrgE *vs* LnmE and LrgE *vs* LnmH

### *LrgE vs LnmE*

```
Lnm  26   LLHSGNAGFIIHRAGQLNHAFRAQGRQFATDLVGHLNKVVEGVATIVVHEEILGTADRLH  85
          LL+S +AGF+IHR GQL + FR +G  F+ DLV  +N    G  +I   EE+ GT  RLH
Lrg  27   LLNSADAGFLIHRVGQLKNEFREEGMSFSADLVDLINNAQVGYVSIFAFEELFGTQSRLH  86

Lnm  86   WLIHMKQPNDYSRFLEIADHDRSFKEITEADRIAAAEGAGNWERMFVEGSFQERVYVPQH  145
          WL+H+K+P+DY R L++ DH + + E++  DR+ A +G GNWERMFVEGS  E +  PQH
Lrg  87   WLLHLKEPSDYRRMLDMVDHSQKWAEVSAGDRLPA-KGGGNWERMFVEGSMTETIICPQH  145

Lnm  146  GLDEHDDDHDHDHHDEPSDTFVPPARHQTGLPDSRMRSSVDSGLTIIRTAQTAFRFRTEA  205
          GL  HD+D DH     P DTF PPAR+Q+ L   ++  SV++ LT+ R  Q  +  R EA
Lrg  146  GLSHHDEDDDH-----PLDTFQPPARYQSRLAPEQLLHSVNTPLTVHRRLQVRYALREEA  200

Lnm  206  REFAFAWASEVNRALGGELTVYLYEETFGQQDRIHWMIHLDSLDTYRKLTELSRHDADYQ  265
          R+F F WA  V+ A+ G  T +LYEE +GQQDR+H +IHL S + Y +L +L + D + +
Lrg  201  RKFWFEWAGHVSEAMVGRATAFLYEEMWGQQDRLHLLIHLASAEAYHELMDLQQTDPELR  260

Lnm  266  ALFGRQFVPDFKGGGGWEQTFVSPTIQDTVLTPLH  300
           L   Q VPD KGGGGW +T +  T+ DT+  PLH
Lrg  261  RLMASQVVPDGKGGGGWHRTVLDGTMTDTLWAPLH  295
```

### *LrgE vs LnmH*

```
Lnm  6    FHSTNSGVIVERTAQLKQAHRAEGRRIALEQAAYLNDKFAGQTTVTVHEETFGVRDRLHW  65
           +S ++G ++ R  QLK   R EG   + +    +N+   G  ++   EE FG + RLHW
Lrg  28   LNSADAGFLIHRVGQLKNEFREEGMSFSADLVDLINNAQVGYVSIFAFEELFGTQSRLHW  87

Lnm  66   LVHLPKLSGHREISRLLD----------GDEGPA-ADADWHGMFVDGSFSETALIPQHWG  114
          L+HL + S +R +  ++D          GD  PA    +W  MFV+GS +ET + PQH G
Lrg  88   LLHLKEPSDYRRMLDMVDHSQKWAEVSAGDRLPAKGGGNWERMFVEGSMTETIICPQH-G  146
```

```
Lnm  115  MYGTDEALPEGTVIDAAAPDLRVPPAQRQTSMSPERTLNSSGAGLMIHRVAQPKYAFRAE  174
          +   DE        D    D   PPA+ Q+ ++PE+ L+S    L +HR  Q +YA R E
Lrg  147  LSHHDED-------DDHPLDTFQPPARYQSRLAPEQLLHSVNTPLTVHRRLQVRYALREE  199

Lnm  175  ARLFARRITESINTRLPGIASSFLYEEAFGPADRVHWLIHMKSEDTYYDLIDMHMRMDDA  234
          AR F       ++  + G A++FLYEE +G  DR+H LIH+ S + Y++L+D+  + D
Lrg  200  ARKFWFEWAGHVSEAMVGRATAFLYEEMWGQQDRLHLLIHLASAEAYHELMDLQ-QTDPE  258

Lnm  235  TRAIYLDEIIAPEKGGGTWNRLFVEESMGDIAFSPV  270
           R +   +++   KGGG W+R  ++ +M D  ++P+
Lrg  259  LRRLMASQVVPDGKGGGGWHRTVLDGTMTDTLWAPL  294
```

## LrgM2 *vs* LnmM

```
Lnm  8    QAYGIEALNVWCGLARLTAADLFAGRGLDPERLDNLMMSERSIGLPIEDPVTNAVNAARP  67
           A G+EA N + G A L   +LF  RGLD +R DNLMM ++S+ LP EDPVTNAVNAARP
Lrg  1    MAVGVEAANAYVGRAALDVRELFHARGLDLDRFDNLMMVQKSVNLPCEDPVTNAVNAARP  60

Lnm  68   LIEALSPEERARIELVVTSTESGVDYSKSLTSYVHKYLGLNRHCRLIEVKQACFGATAAV  127
          +++ L+ E+R  IE V+  TESG+D+ K +++YVH YLGL R CR  EVK AC+G TAA+
Lrg  61   ILDRLTGEQRDSIEAVIVGTESGLDFGKPISTYVHHYLGLGRRCRSFEVKHACYGGTAAL  120

Lnm  128  QTAVGYLASGISPGAKALVIATDVAVVDEKAEYSEPAAGHGAAAMLLSDRPRVLAMDLGA  187
          +TA+G LA    P A+ALVIA D         Y EP+ G GAAA+LL   P VL +D GA
Lrg  121  RTALGLLAGSPRPRARALVIAADAGGSVVGMPYWEPSHGAGAAALLLGPEPEVLDLDPGA  180

Lnm  188  FGNYSYETLDSARPSPRFDIADVDRSLFAYLDCLKNAYADYAARVTDVDFTRDFDHLVMH  247
           G YSYE +D+ RP P     D D SL AYL+CL+ ++ADY   V   D    FDHL  H
Lrg  181  AGLYSYEVMDTCRPRPDLAAGDSDLSLLAYLECLERSFADYRDTVAGADIVETFDHLAFH  240

Lnm  248  TPFAGLVKAGHRKMMREQGVTGPR-IDEDFARRVAPSLIYPGSVGNLCSGSVYLALASLL  306
          TPF G+VK  HR +MR      P  I+ DF  R+A SL+ P  +GN  +  +++AL SL+
Lrg  241  TPFGGMVKGAHRHLMRRVKRMPPDAIEADFRARMAASLVNPARIGNTYAAGLFVALLSLI  300

Lnm  307  DSGVVTAPSRVGLFSYGSGCSSEFFSGIVDEQSAATVAEQGIGKRLEARARITFDEYLAV  366
          + G  + P RVGLFSYGSGC+SEF+SG+V   + A +A+    + L+AR R+   +Y  +
Lrg  301  EHGDFSEPRRVGLFSYGSGCASEFYSGVVGAGAPALLAQYRTAEHLDARHRLAVADYDRL  360

Lnm  367  LEHNLECLVPVENRTVDPAEWEPLLDRVGDRPEILTFTGVKDYHRQYAW  415
           +H  E    V++  +D + +  + +      E L    ++++HR+Y W
Lrg  361  AKHTAEGAPGVQDAVLDVSPYADVYEPALHGREQLVLDRIENFHRRYRW  409
```

## LrgF *vs* LnmF

```
Lnm  4    IGPTHRGVRLTAEPHVLRATLTSPDGLNSLSGAALDALGAALDRAEADPECRVLLLEGSG  63
          +G  HR +RL+  P VL  TL  PD  NS+  A +  L AALD AEADP+CRV++L G
Lrg  3    LGTDHRTIRLSGTPGVLTVTLDRPDAQNSIDTAMIRELHAALDTAEADPDCRVVVLTGGD  62

Lnm  64   GTFCTGLDFEEAAGDPAGGASQAGRGGAEFLALMRRFGETPLAVVACVDGRAAGGGVGLA  123
          G FCTG+D   AA +      +A +GGAEFL L++R   TP  VVA VDGR AGGGVGL
Lrg  63   GVFCTGMDLLGAAAEGRPDRERAAQGGAEFLGLLKRLTLTPRVVVATVDGRVAGGGVGLV  122

Lnm  124  AAADLVIATERSEFSLPEALWGLVPCCVLPVLVRRTGFQPAYAMALSTQPVSARRAADFR  183
          AA+D V+AT RS FSLPEALWGL+PC V P L+RRTGFQ AYAM+LST PV+A +A
Lrg  123  AASDFVLATPRSTFSLPEALWGLLPCSVAPFLIRRTGFQKAYAMSLSTLPVTADQALLSG  182

Lnm  184  LVDEVVPDPDAAVRRLLVRLTRLDPATIGELKQYFRAMWFTTEDTDAFALREFTRLIDSP  243
          LVD++      A+RRL  R+T+++ AT+G+LK+YF  MW  T + +  A+ EF RL+ S
Lrg  183  LVDQLADQLQPALRRLAFRVTKVEGATVGDLKRYFAKMWIITPEVERAAIDEFARLMSSD  242

Lnm  244  VARRRITDYTTTRRLPWE  261
             RRI D+   +R PWE
Lrg  243  GVGRRIADFAERQRFPWE  260
```

## LrgG *vs* LnmG

```
Lnm  2    VALVFPGQGSQRKGMGADLFARFPDLTRQADTVLGHSVEELCRSSGDGRLDRTEYAQPAL  61
           A +FPGQG+QR+GMG DLF R+P+    AD +LG SV ELC      RL  T Y QPAL
Lrg  5    TAWLFPGQGAQRRGMGRDLFDRYPEAMAAADRILGFSVRELCLGDAGERLTDTRYLQPAL  64

Lnm  62   FVVSALS---YLARDPGLPQPTLLAGHSLGEYGALFAAGCFDFATGVRLVRERGALMGRA  118
          FVV+ L+   Y AR+P    P  LAGHSLGEY AL AA  FDF TG+ LV  RG LMGRA
Lrg  65   FVVNELTRRAYAAREPA---PDFLAGHSLGEYNALLAADAFDFETGLALVARRGELMGRA  121

Lnm  119  QGGGMLAVLGVDGDEVQALLAGTGARQVDVANYNTPTQTVLSGPLDELRMVSAALGQRPG  178
           GG M AV+G     +  LLA  G   VD+AN N+  Q VLSGP + LR  + A+
Lrg  122  TGGAMTAVVGPGAARIADLLAEAGLGDVDLANLNSAEQVVLSGPAESLRRAAGAVTAAGA  181

Lnm  179  VRCVPVRVSAAFHSRHMRPAAQEFATFLTGFSFADPHRTVISSVTARPYGAGQVAELLSR  238
           RCVP+RVSAAFHSR+M  AA+EFA FL GF   DP   VI++VTARPY  G+V  LL+
Lrg  182  GRCVPLRVSAAFHSRYMADAAREFAAFLDGFELRDPRLPVIANVTARPYRPGEVRRLLAA  241

Lnm  239  QIESPVRWSETMAYLRERGTTELEEMGPGKVLTGLWKQGRADGAKARAVAPAPVAVVAGV  298
          Q+ S VRWSE+M +L  RG   + E GPG+VLTGLW     D A   A AP   A
Lrg  242  QVHSAVRWSESMRHLLARGVDRIAEQGPGRVLTGLWDAAVKDAANGPAPAPRRAAGGDPA  301

Lnm  299  PAAAAARAPASPAPVAARAATAAPARPSDPAPPAPRTAPSPASPVPPASVSRGQRAEELG  358
          P AA A  PA  A       T   + PS  A  A   A  PA  VP      G  AE LG
Lrg  302  PPAAGAAPPAVGAAADPAPGTGPASAPSPAAGSATPLADRPAEGVP------GPTAERLG  355

Lnm  359  SAEFRQDYGIRYAYLAGAMFRGIASAELVIRMGRAGLMGFFGAGGLGLDKVESALVRIKD  418
          SA FR DYG+RYAYLAG+M++GIAS ELV RMGRAGL+G+FG GGL ++++E+A+  ++
Lrg  356  SAAFRADYGVRYAYLAGSMYKGIASTELVARMGRAGLLGYFGTGGLRMERIEAAIAALRA  415

Lnm  419  ALGPDGRYGMNLLHSIDDPAYEHAVVDLCLKHGVHDVEAAGFTQLTPAVVQFRFSGAHRD  478
           LGP G +GMNLLH++ DPA E A V L L+HGV  VEA+GFTQ T A+V +R SGAHRD
Lrg  416  ELGPGGGFGMNLLHALHDPALEEATVRLFLRHGVRFVEASGFTQPTRALVLYRLSGAHRD  475

Lnm  479  AAGRAVAVRRVLAKVSRPEVAAAFMAPAPAAILRRLTADGRLTPQEAEIAAELPVGQDIC  538
           AGRAVA  R+LAKVSRPEVA AFM PAP A++R L ADG LT  EA     LP+  ++C
Lrg  476  GAGRAVAPNRLLAKVSRPEVATAFMEPAPEALVRGLLADGLLTAAEAGAGRALPLAGEVC  535

Lnm  539  VEADSGGHTDGGAALTLLPSMIRHRDAAMARHGYGRRIRIGAAGGIGAPEAVAAAFVLGA  598
           EADSGGHTD   A TL+P+M   RD  +A  GY   IR+GAAGG+GAPE++AAAFVLGA
Lrg  536  AEADSGGHTDAAVAYTLMPAMTDLRDRIVAERGYPDGIRVGAAGGLGAPESLAAAFVLGA  595

Lnm  599  DFVLTGSVNQCSPEAGTSDAVKDILAGLDVQDTAYAPAGDMFEIGARVQVVRKGTLFAAR  658
          DFV+TGSVNQC+P+AGTSDAVKD+LA  DVQDT YAPAGDMFEIGARVQV+R+GTLFAAR
Lrg  596  DFVVTGSVNQCTPQAGTSDAVKDLLAQADVQDTGYAPAGDMFEIGARVQVLRRGTLFAAR  655

Lnm  659  GNKLYQLYRSHDSWESIDAGTRRSVEETYFKRPFAEVWEETRAYHLGRGRDAEIEKADRL  718
           NKLYQ YR +D  E+IDA  RR++EE +F+R   +VW+ETR YHL  GR  E+E+A+R
Lrg  656  ANKLYQAYRRYDGLEAIDAKLRRTIEEGWFRRDLEQVWKETREYHLRAGRPQEVERAERD  715

Lnm  719  PKHRMALAFRWYFARSVRWGLEGEPTQKVNYQIQCGPAIGAFNHVVRGTGLEDWRHRHVD  778
          PKHRMAL FRWYF  S R  + G+P ++VNYQI  GPAIGAFN    GT L DWR+RHVD
Lrg  716  PKHRMALVFRWYFVHSTRTAMRGDPAERVNYQIHTGPAIGAFNRFAAGTALADWRNRHVD  775

Lnm  779  LIAEHLMTGAADVLARR  795
           +AE LM GAA+VL  R
Lrg  776  AVAEALMAGAAEVLRDR  792
```

## LrgP2 *vs* LnmH and LrgP2 *vs* LnmE

### *LrgP2 vs LnmH*

```
Lnm  6    FHSTNSGVIVERTAQLKQAHRAEGRRIALEQAAYLNDKFAGQTTVTVHEETFGVRDRLHW  65
           HS N+GV+VER  QL+   R+EGR+ A E + YLN ++ G  T  V+EETFG +D LHW
Lrg  26   LHSANAGVVVERVGQLRAEFRSEGRQFARELSEYLNTRYVGVATTFVYEETFGTKDTLHW  85

Lnm  66   LVHLPKLSGHREISRLLDGDEG------------PAADADWHGMFVDGSFSETALIPQHW  113
```

```
          L+H+  L  +  + R+   DEG                  W  MF+DG   ET LIPQ +
Lrg  86   LLHMRSLEAYETLVRMGSQDEGWREVMFRNRIPEERGGGSWDRMFLDGGLKETVLIPQSF  145

Lnm  114  GMYGTDEALPEGTVIDAAAPDLRVPPAQRQTSMSPERTLNSSGAGLMIHRVAQPKYAFRA  173
          GMYGT +      V D+      VP A+ QTS  P+  L+S+ +G+++HR  + KY FRA
Lrg  146  GMYGTADTELSSVVEDSGVDRFVVPTAEHQTSQKPDELLHSANSGIIMHRTGELKYEFRA  205

Lnm  174  EARLFARRITESINTRLPGIASSFLYEEAFGPADRVHWLIHMKSEDTYYDLIDMHMRMDD  233
          E R FAR +TES N  L G A+ FLYEEAFG +DR+HW IH++   +YY+L+ +  R
Lrg  206  EGREFARALTESWNASLGGEATIFLYEEAFGLSDRIHWFIHLRKLSSYYNLMGLRARTSP  265

Lnm  234  ATRAIYLDEIIAPEKGGGTWNRLFVEESMGDIAFSP  269
          A R ++  + I  EKGGG W R+FV+ S+ D+A +P
Lrg  266  AAREVFTKQWIPEEKGGGGWERMFVQGSLQDLALTP  301
```

*LrgP2 vs LnmE*

```
Lnm  24   SDLLHSGNAGFIIHRAGQLNHAFRAQGRQFATDLVGHLNKVVEGVATIVVHEEILGTADR  83
          +D+LHS NAG ++ R GQL   FR++GRQFA +L  +LN    GVAT  V+EE  GT D
Lrg  23   ADVLHSANAGVVVERVGQLRAEFRSEGRQFARELSEYLNTRYVGVATTFVYEETFGTKDT  82

Lnm  84   LHWLIHMKQPNDYSRFLEIADHDRSFKEITEADRIAAAEGAGNWERMFVEGSFQERVYVP  143
          LHWL+HM+    Y   + +   D  ++E+   +RI    G G+W+RMF++G  +E V +P
Lrg  83   LHWLLHMRSLEAYETLVRMGSQDEGWREVMFRNRIPEERGGGSWDRMFLDGGLKETVLIP  142

Lnm  144  QH-GL-DEHDDDHDHDHHDEPSDTF-VPPARHQTGLPDSRMRSSVDSGLTIIRTAQTAFR  200
          Q  G+    D +      D   D F VP A HQT      +  S +SG+ + RT +  +
Lrg  143  QSFGMYGTADTELSSVVEDSGVDRFVVPTAEHQTSQKPDELLHSANSGIIMHRTGELKYE  202

Lnm  201  FRTEAREFAFAWASEVNRALGGELTVYLYEETFGQQDRIHWMIHLDSLDTYRKLTEL-SR  259
          FR E REFA A     N +LGGE T++LYEE FG  DRIHW IHL  L +Y  L  L +R
Lrg  203  FRAEGREFARALTESWNASLGGEATIFLYEEAFGLSDRIHWFIHLRKLSSYYNLMGLRAR  262

Lnm  260  HDADYQALFGRQFVPDFKGGGGWEQTFVSPTIQDTVLTPLHPGAPA  305
               + +F +Q++P+ KGGGGWE+ FV  ++QD  LTP H G  A
Lrg  263  TSPAAREVFTKQWIPEEKGGGGWERMFVQGSLQDLALTPQHWGMYA  308
```

LrgI *vs* LnmI

```
Lnm  23   VMVRVLSAARPDAQPVTAACELRGLGLDSMTAARLWLAVQGECAADVPLGWLTEATTVGE  82
          ++ R+L+  RPD         L   GL S+   RLW  V+ E   D+P   L  A T+ E
Lrg  14   ILHRLLAEVRPDGAEPLPPGGLADWGLTSLALTRLWAGVRREFGIDLPPARLA-AATLDE  72

Lnm  83   YAQRVADHASQAVPVQGAGAVGAQVAADPDALHEPFPLTPLQEAYLIGKEPELQADAVGC  142
              VA   +   P + A AVGA          EPFP+T LQ+AYL+ K  +L  +AVGC
Lrg  73   LTATVATGTAGTAP-RPATAVGADG-------REPFPVTDLQQAYLVAKGSDLGGEAVGC  124

Lnm  143  HLYREFDVPALDTERLRAAWQRLVEHHDILRATVTEDGRQQITAQAPRWDLAVHGSATRA  202
          HLYREF VP LD +RLR AWQ L++ H +LRA V +DG Q I  +A  W L VHG
Lrg  125  HLYREFAVPDLDPDRLRRAWQDLIDQHGMLRAVVDDDGTQHIRPRADDWQLPVHGPGG--  182

Lnm  203  EFTETATAVRARMSHHLFPAGHCPPFAIEVTLGPDGTGRVHFGIDAIVTDGQGLDLLTAQ  262
             E    VRAR+SH  +  G  P FAIEVT    G   VH  +D ++TDG G  +L  Q
Lrg  183  --AEDRDTVRARLSHRCYRPGAWPLFAIEVTRTGTGPDIVHLSLDTLITDGHGYAVLLQQ  240

Lnm  263  WEACYADPSHLLPAPTAPLSVRDCVVALDAARRTEAHRRDLDHWVRRLRELPGAPGLFTA  322
          W   Y  P    P P A  SV + V AL A R ++AHR DL +W   L +LP  PGL
Lrg  241  WHQRYHHPEQ--PLPAAGPSVAELVPALLAERGSDAHRADLAYWKDELADLPAGPGLVAP  298

Lnm  323  DAPERGTGLSCVRRSSRTARLTAAEWRSLRARAEELAVSPTSLVLTVFTEALARHGAHEP  382
                G  C  R    A L A +WR+L  RA  L VSPT+LVL +F EAL R
Lrg  299  VPAAAPAG-GCRERRPLDAELPADQWRALTVRAARLGVSPTALVLALFAEALDRRAPQRR  357

Lnm  383  FSLVVTTSRRPQLPPEADHLVGPFTGTTFVEAVPPQQHTFEEAARLTHEGLWQALEHSAV  442
           +LVVTTS RP L  E  HLVGPFT T  +        T +E A   H  L + L H  V
Lrg  358  TALVVTTSSRPYLAAETPHLVGPFTSTAVLVLEREADETLDEQAAAVHARLGEHLRHGLV  417
```

```
Lnm  443   CGVSAQRALRGGGPGPLPVVFTSMLDAAGRP-RARGFAAAPVYAVSQTSGVWLDHQMWEQ  501
            G+ A R  R G P P  VVFTS+LD    P  A GF AA  Y VSQT+ V LDHQMWEQ
Lrg  418   SGIEALREQRTGRPAPA-VVFTSLLDVGPPPGVAGGFGAAIEYGVSQTTDVALDHQMWEQ  476

Lnm  502   DGALHLRWDTADGCFAPGVVEAAFASLCNGLRALAVAGPVVTRPLNDLQQAYFVARAAGE  561
           DGAL  RWD     FAPG VE AFA+  N L           RP   LQ+AY VAR A
Lrg  477   DGALRYRWDVDPTRFAPGAVETAFAAFGNALAGACAEPAGDPRPPRALQEAYMVARTAAG  536

Lnm  562   PGPWRGCQVVVPYDTDERVDPVRLESAVVRLVEAYDVLRSAVTQDGVVEVRAGAPRRWTV  621
            GP  GCQ    ++ D  +D   L  A+ RLV+ + VLR+A   DGV + R G P +W +
Lrg  537   DGPAEGCQCYQSFEVDA-LDLEALADALRRLVDGHAVLRAAFAADGVTDRRRG-PGQWRI  594

Lnm  622   PVVAGGCPDE-------VRDAMAAANFPLGRYPQFEVRVVRGDDGDTVL-MSMDLTLTDA  673
           PV+  G P         +RD M    FPLG +P  ++RV R D G +V+  + DL + D
Lrg  595   PVIEAGDPAAHPALRALIRDEMTNRPFPLGSWPLVDLRVTRDDSGFSVVHCAFDLLVADG  654

Lnm  674   RGIHLTGRELMRLYADPAAEPRPAEAARDSARDADEQARSRAHWQDRLRALPPGVPLPGP  733
             IH   R+L RLYADPAA P PA      A  A EQAR   +W++RL  LP G  L  P
Lrg  655   LSIHRLYRDLWRLYADPAARPVPAGPDPAPASPAAEQAR---YWRERLSELPAGPELGPP  711

Lnm  734   RDADGPDR-RVRLAGAPLALRPLTDRCAEHGLSLDAVLLTAFTDVLARTYGTDFAVPVVR  792
                  R R RLAG     R L  R AE+GL  D +LL A T  ++      +A+PVV
Lrg  712   GGPARAGRGRTRLAGRIDGYRRLARRAAEYGLHPDDLLLAALTRAVSSHTAQPWALPVVL  771

Lnm  793   WDHGLDPQRPGEFTALSWLPCAPRELSFTARARTYQEGLERDADVSGSG-LPELRRAVAR  851
           W    +  RPGE + ++W+  AP E+  T  AR Y+  L+ D    G+G L E+RR V R
Lrg  772   WPKTAEAARPGEHSFMTWVTAAPAEVPLTTAARDYRRVLDADLAEGGTGGLGEMRRQVLR  831

Lnm  852   SGGAGYPVVYTCALDLTDRPLPGSVRAGQWLSCTPDVFLDCITTVDAGQLQLAWDAVDGR  911
              A +PVVYT  +DLTDRPLP  VR G+WL+ TP   LD +   +  +L+  WD V
Lrg  832   GSDAAHPVVYTALVDLTDRPLPAGVRQGEWLTSTPGCALDSVAVAEGDELRYCWDIVHAD  891

Lnm  912   APQGGWSELHAEYRRSVARLADDAAAWQEPAGGDTSGADDGEVRGAELHKILHEWNDTTR  971
            P G    + A +   +  LADDA    +PAGG T+          E H +L+ WN+TT
Lrg  892   FPDGLAERMFAVFENGLRLLADDATWAGDPAGGLTA---------EERHTVLYAWNETTV  942

Lnm  972   AFPDDRLMHQLFEEQAAQQPRAEALRWRGGGTMTYQELNRRANRIAARLAAEDVGPETVV  1031
             PDD   H LF+ QAA  P A ALRWRGG TMTY ELNRRANRIA RL A  VGP  +V
Lrg  943   PVPDDGPAHLLFQRQAASTPEAVALRWRGG-TMTYGELNRRANRIAHRLTAAGVGPRHLV  1001

Lnm  1032  AVSVPRGPMMVAVVLGILKAGGVYLPMEPHLPAERAAVILEEAHAEVVVTTADREGWPVP  1091
            V + RGP MVA + GILKAG  YLP++P LP  R A +L  A A  V+TT+D     +P
Lrg  1002  GVRIRRGPEMVAALHGILKAGAGYLPIDPALPGTRVAAMLGLARATTVLTTSDTPAAALP  1061

Lnm  1092  DGYARVCADAAVEGPHPADADNCPRPVTQPHNTAYIIFTSGSTGRPKGVAVAHRPVLNLI  1151
            G   +  D       PA  ++ P   +   +TAY+IFTSGSTG PKGV VAHR V NL+
Lrg  1062  AGVTGIETDTDPAINDPAAREDDPETRSTRDDTAYVIFTSGSTGTPKGVQVAHRSVRNLL  1121

Lnm  1152  NWCRRTFGFGPGDMGLCVTSLGFDLSVFDVFGLLGTGAALYIADAEQQRDPALLLDVLIE  1211
           N+C RTF   P D+GL VTSL FDLSVFD+ GLLG GA +Y+AD  QQRDP LLLD+L+
Lrg  1122  NFCHRTFELRPSDLGLAVTSLSFDLSVFDMLGLLGCGAGVYLADETQQRDPELLLDILLT  1181

Lnm  1212  EPVTFWNSAPTTLAQVGPLLD-TVGTAGTGDLRLVFLSGDFTPLPLPDEVRAVFPRADMI  1270
           EP+TFWNSAPTTL Q+ PLL    G A    LR+V L+GDF PL LP  +R  FP A+ I
Lrg  1182  EPITFWNSAPTTLHQLTPLLTPDTGDAAAAHLRIVALAGDFIPLSLPGAIREAFPNAETI  1241

Lnm  1271  SLGGATEATVWSNWFRIGAIDPAWRSIPYGRPIDNSRYHVLDEALRPCPVGVEGDLYIGG  1330
           +LGG TE TVWSN +R+  +DP WRSIPYGRPIDN+R++VLDE L PCP+GVEGDLY GG
Lrg  1242  ALGGPTETTVWSNVYRVTTVDPDWRSIPYGRPIDNTRHYVLDEHLEPCPIGVEGDLYTGG  1301

Lnm  1331  ECLALGYVNQPELTADRFIPDPFHEDPQERLYKTGDRALYYPDGNLSFQGRADGQVKVRG  1390
           ECLA+G+ NQPELTA  F+PDPF + P ER+Y+TGDRAL+ PDGNL   GR D QVK+RG
Lrg  1302  ECLAVGFCNQPELTARLFLPDPFVDTPGERMYRTGDRALWLPDGNLRVTGRGDRQVKIRG  1361

Lnm  1391  FRVELAEIEHRLRAHDGVKDAVVLAREDG--CGDRTLVAYLVALPGS-APSGRELRGFAG  1447
            RVEL E+EHRLRAH  V++ V + R+D    GD  LVAY+V  P + A +  ELR
Lrg  1362  HRVELGEVEHRLRAHPAVQEVVAVLRDDEPPSGDARLVAYVVTDPAAPAVTVAELRAHTA  1421
```

```
Lnm  1448  QTLPEYMVPNFIGFLAGFPATANGKLDRAALPWPL  1482
           + LP YMVPNF+  L  FPATANGKLDR ALPWPL
Lrg  1422  EALPGYMVPNFVALLPSFPATANGKLDREALPWPL  1456


Lnm  1527  VSVPSRDELCAEIADLFAQALGVESVDADTDLWDQGATSFTMVRVSGSLQRSYKQRFPVS  1586
           + V S   L  E+A +FA+ LGV++VD   DLWDQGATSFTMV++S  L++ Y+QR PVS
Lrg  1534  IPVASPAALAEEVAAMFARHLGVDTVDPALDLWDQGATSFTMVQISAGLRKRYQQRVPVS  1593

Lnm  1587  ALLDNPSVSAIAGWVHAQLGGGADAESTAAAEAETATSVDAETTATTVTQTTAASDERPD  1646
           AL+  P+ + IA  +  +LG  A  +  AA E                         P+
Lrg  1594  ALISEPTAAGIARILADRLGLRAQPDPAAAPEPGAG----------------------PE  1631

Lnm  1647  SGPGPVDFFATEERERFKRQHWNRRPDEPGLPEVPLGDARFEDELHAWRASRRDFLDQPV  1706
           +GPG VD F+ +ER+ FK   WN R   PG   + L +       + WR S RD+ D P+
Lrg  1632  AGPGTVDMFSPQERDAFKAAAWNLRRPAPGARRIALPETTVHPAHYDWRGSHRDYRDTPL  1691

Lnm  1707  PHRSFSRLLGLLRETTGADGTGALYPSAGDTYSVQVYLHLTPDAVEGLDAGLYYYDPSRH  1766
           P  + +RLLGLLRE     GT  LYPSAGDTY+VQ Y+H+ PDAV+GL  G+YYYDP  H
Lrg  1692  PAEALTRLLGLLREAPVEGGTRRLYPSAGDTYAVQAYVHVKPDAVDGLAGGIYYYDPRGH  1751

Lnm  1767  SLRLLRSGVLPDRGAHFYYNRPVFDRSRFGIYLFGQRHGIEPLYAEESLRYLTLESGYMS  1826
           +L L+ +    DR  HFYYNRPVFD S FGI+L GQ  GI PLY E +  +LTLE+GY+
Lrg  1752  ALELVNAEPRIDRTVHFYYNRPVFDGSAFGIFLIGQTRGIAPLYQEVAEHFLTLEAGYVG  1811

Lnm  1827  QLLMLGQAAHGVGLCPIGALNTEQLSQWLGLDEGHVFLQAFLGGAAEHPQRTAGGTVPFF  1886
           QLLM GQAA G+GLCP+G L  + +     LD+GHVFL +F+GG  +    TA    PF
Lrg  1812  QLLMTGQAACGIGLCPVGTLTFDDIRDQFALDDGHVFLHSFMGGGVDR-TGTADLRPPFA  1870

Lnm  1887  TEPTDSDGNSGSGDSSTVITDAVAPVSAAAEDADAEPPAHTAEPAAVIGMAGRLPGAGDL  1946
            +P                         A   AD EP    A   AV+GMAGR PGA DL
Lrg  1871  EQP-------------------------ARAAADEEP---KAAEVAVVGMAGRFPGAEDL  1902

Lnm  1947  DAFWDNLVSGRTAIGPAPASR----PETAPSGARATGGFLPHIDRFDSLLFHVSPQEAPA  2002
             +W  L +GR A+GP PA R        P+G    GGFL  ID FDSL F ++P EA A
Lrg  1903  GEYWRQLSTGRCAVGPLPAGRGFAEDGRLPTGLH--GGFLTGIDNFDSLHFRIAPVEAAA  1960

Lnm  2003  LDPQARLMLESVWQCLDDAGHTADSLRRSAGRVGVFIGSMWHDYRQQGADRWNGGDSAEV  2062
           LDPQ RL+L++VW CL+DAGHTA+SLR +A RVGVF  +MWHD++  G + W+   +A V
Lrg  1961  LDPQLRLLLQTVWTCLEDAGHTAESLRAAAPRVGVFTAAMWHDHQHTGKETWDADAAARV  2020

Lnm  2063  AATASDIANRVSHFFDFRGPSLAVDTSCSSSFAALHLAVESLRRGECGAAVVGAVNLLAH  2122
           AA A D+  R+SH F F GPS+AVDT+CSSS  ALHLA E+LRRGEC AAVVGAVNL+AH
Lrg  2021  AALAGDMPGRISHCFGFDGPSVAVDTACSSSLTALHLAAEALRRGECDAAVVGAVNLVAH  2080

Lnm  2123  PYHWGLLDGLELLAADAPPAAYAAEGSGWHPGEGVGVLLLRPADAARRAKDTVHGLIEGT  2182
           PYH  LL    LLA   P  A+AA+ SGW PGEGV  +LLRPA AA R  DTV G++E T
Lrg  2081  PYHLALLAEAGLLAEGGPVRAFAADSSGWCPGEGVAAVLLRPAAAADRDGDTVRGVLEAT  2140

Lnm  2183  RIGHAGRAPRYGAPHTAALADSLARALADASVIPDEVDYVECAAAGAGIADAAELEALGS  2242
            IGHAG   R+G P   ALA S+  AL  A + P +VDYVE A AGAG+ADAAELEA
Lrg  2141  WIGHAGSGGRFGVPDPRALAGSVGAALDRAGITPAQVDYVELAVAGAGVADAAELEAFAE  2200

Lnm  2243  VLARCAGASPVPVGTLKPNIGHLEAASGLSQLIKVLLQIRHGRIAPTLVSGELSPLVDWD  2302
           V A     +PV  GT+KPNIGHLE+ASGLSQL+KVLLQ  H RIAPTL +G  S LVDWD
Lrg  2201  VFA----GNPVLAGTVKPNIGHLESASGLSQLLKVLLQFEHDRIAPTLTAGRRSDLVDWD  2256

Lnm  2303  GLPV----ELVDTPRALTPRAADGRATVLVNAVGATGSYGHVVVRAPHAHGTGPAAQDGL  2358
            LP+     LVD P A T R A      +VNAVGA+G+Y H V+RA  +   GPA
Lrg  2257  DLPLRVPEHLVDWPSAATARRA------VVNAVGASGAYAHAVLRAATSADDGPAP----  2306

Lnm  2359  AGAGAAPSASGPRTVVLSAASPEGLTAAAGRLRDHLAGAGRALCLDDVAWTLQTGRASLG  2418
            GA A   A G + VVLSA S +GL  AAGRL +HLA  G +  L D+A+TLQ GR  +
Lrg  2307  -GATADGRAPGRQAVVLSAGSADGLRRAAGRLAEHLAN-GASPALADLAFTLQDGRVPMP  2364

Lnm  2419  HRLTLSADGLDGVRAGLTAFLDGRACPGLATAAADPALAGVPAGAQDLARAAGDWLRGHA  2478
           HRL +    +  VR  L  F  GR  P LA A   P          D   AA  WLRG
```

```
Lrg  2365  HRLAVVTSDIAEVRTALAEFAAGRTAPALADAVVGPGRPAAGLVPADADAAAAGWLRGAP  2424

Lnm  2479  VDFARLWSAPARRVPLPVQDFTVLAQERHWLAAPAARRPDGAAGSAPAAPESGQSAPPAS  2538
           V +  LWS   RRVPLP    TV A E H L AP             A    +  P A+
Lrg  2425  VAWHALWSPGRRRVPLP---GTVFAGEEHRLTAPPR-----------TASAPARPVPVAA  2470

Lnm  2539  PQVQDDRADRAQEHVAACFAEVSGIPAEQLHPRVPLEHYGLSSRLVARFNERLRQD-VQG  2597
           P    +  D  +++V A +AE SGIP E+L PRVPLEHYGL+S +V + N RL +D  +
Lrg  2471  PAPVAEVPDTVRQYVLAVYAEASGIPVERLDPRVPLEHYGLTSYVVGQLNARLAEDFTEP  2530

Lnm  2598  VSSTVLFEYPDLAGVAAHLAAHHEGPW----SAAPDTQPSPPVPSPDPLPVPRTPAAALG  2653
           VS T+ FE+ DLAGVAA LAA  +GPW    + AP  +P+P
Lrg  2531  VSRTLFFEHQDLAGVAAELAARVDGPWQPVRTEAPGDRPAP-------------------  2571

Lnm  2654  ESAAADGPE--PIAVIGIAGRYPGAGDLETFWSNLAEGVDSVGPLPAERARDGWPTEQMW  2711
              A + PE   IAV+G+AGRYP A DL+ FW  L +G D++GPLPAER R GWP + MW
Lrg  2572  ---AGNRPEDTAIAVVGLAGRYPQAADLDRFWQLLEQGYDAIGPLPAERHRPGWPVDLMW  2628

Lnm  2712  GGFLDGVDRFDALFFGIAPRDAQLMDPQERQFLQVVWETLEDAGCTRARIREQLGSDVGV  2771
           G FLD VDRFD LFF I+PRDA LMDPQER FL+V WE LEDAG TRAR+REQ G  VGV
Lrg  2629  GAFLDDVDRFDPLFFAISPRDAVLMDPQERLFLEVAWEALEDAGYTRARLREQHGGRVGV  2688

Lnm  2772  FVGTMYNEYPFFGVERSLAGESADTGSAVAGIANRVSYFLDLHGPSLAVDTMCSSSLTAL  2831
           F G MYNEYPF GVE+SL G +ADTGSA+AGIANRVSYFLDL+GPS+ VDTMCS+SLTA+
Lrg  2689  FAGAMYNEYPFLGVEQSLLGPAADTGSALAGIANRVSYFLDLNGPSMTVDTMCSASLTAV  2748

Lnm  2832  HLAVESLRRGECAAAVAGGVNLSLHPHKFRQQTRLKMSSSDHRCRSFGAGGDGFVPAEGV  2891
           HLAV +LR+GEC AA+ GGVNLS HPHKFRQQ RL+M+S+DHRCRSFGAGGDGF P EGV
Lrg  2749  HLAVRALRQGECEAALVGGVNLSAHPHKFRQQHRLRMASTDHRCRSFGAGGDGFTPGEGV  2808

Lnm  2892  GAVLLKPLSAAEADGDRIHAVIRGTAVNHGGKTNGYMVPNPVAQGDLVRAALRRAGADPA  2951
           G VLLKPL+ A ADGDRIH VIRGTAVNHGG+T+GYMVPNPVAQG+LV AALR AG  P
Lrg  2809  GVVLLKPLARAIADGDRIHGVIRGTAVNHGGRTSGYMVPNPVAQGELVAAALRDAGVGPG  2868

Lnm  2952  TIGYVEAHGTGTQLGDPVEINGLNRAFAGASVAPASRAIGSVKANIGHAEAAAGIAGLTK  3011
           +IGY+EAHGTGT LGDPVEINGL RAFAG  V   + AIGSVK+NIGH EAAAG+AGLTK
Lrg  2869  SIGYLEAHGTGTALGDPVEINGLARAFAG--VEAGTCAIGSVKSNIGHLEAAAGLAGLTK  2926

Lnm  3012  VVLQLRHRHLVPSLHTEELNDAVDWASSPFEVVREGRPWAPLTGADGAPLPRRAGLSAFG  3071
           V+LQLRHR LVPSLH E+LN  +DWA SPF V RE  PW     ADG P+PRRAGLSAFG
Lrg  2927  VLLQLRHRRLVPSLHAEQLNPDIDWARSPFAVQREAAPWPARRAADGTPVPRRAGLSAFG  2986

Lnm  3072  AGGANAHVVVEEYVPGTAPEPTEPGVPGVLEPQLIVLSAHDLGRLRALAGRLRDRLGRDD  3131
           AGGANAH+V+EEY+P  A EP  P       PQL VLSA D  RL  LA R  D L R +
Lrg  2987  AGGANAHLVIEEYLPQPAAEPVRPA-----GPQLFVLSARDEQRLTELAHRWADFLARPE  3041

Lnm  3132  RPAPALADVAHTLQSGREPLRERVALVAYDVAGLCRALDLFASGDTGAWVHGRTPGGALP  3191
              P  AD+AHT QSGREPLRER+A+VA   A L   L  F  GD+G  V GRTPG   P
Lrg  3042  --LPPFADLAHTAQSGREPLRERLAVVAAGPAELRAKLLRFLDGDSGDVVRGRTPGADAP  3099

Lnm  3192  DGPKAVLDAAADRDAELLRLGRHWTGGGTVDWPGLHPVRR-RLVSLPSYPFAEDRHWLPE  3250
            GP     +  D    LL L RHWT GG VDW  LH   R R  + PSYPFA  R+W PE
Lrg  3100  AGPHPADGSYGD----LLLLARHWTAGGRVDWSRLHTGDRPRRAAAPSYPFARGRYW-PE  3154

Lnm  3251  PRTAAPAAAAATLTEPSGTT--LYGRTWRALPPLAAAAAPAP-SGRVLCVFSAPGEPVAR  3307
           P   A    AA + EP G    L+ + WR+    A    P P +GR++C+++      AR
Lrg  3155  P---AKTVEAAPVQEPDGRQPLLFTKRWRS----AGQPRPEPVTGRIVCLYTDRSREAAR  3207

Lnm  3308  ALAALLGPDRVTLVRAGADAGNGVPGITGIGDEAEAAAFAQGLRADGPDAVGGLIDLTDL  3367
            +A   G DRV  VR G  A +G  G TG   E +A A   GL    PD + G +DL++L
Lrg  3208  QVAEAFGRDRVIPVREGGPA-DGDAGFTG---EQDAIALLDGLADRHPD-LTGWLDLSEL  3262

Lnm  3368  GGPAHGDAGSWTARLVLLRRLVRTLRGHGGRVLHVTEGLYGPAGPAPSLAGARMAGFVRM  3427
             PA  D G WTARL  L+RL+    G   R++H      G  G        R+AG VR+
Lrg  3263  DRPA-ADPGPWTARLAALQRLLARRPGTALRIVHAVRDRGGADG-------RRLAGVVRL  3314

Lnm  3428  LGAEYGRVTGTVLDLDVSAVGPDAAARQILAEYTGPYGPGDVSVRGGVRHRPELVALPDA  3487
           LGAE+  V  TV++ D    GP   AR++LAE+ G     +V  RGG R  P L  +P
```

```
Lrg  3315  LGAEHRTVRATVVESDA---GPAELARRLLAEWAGGEATAEVRHRGGERLTPVLEPVPVP  3371

Lnm  3488  GHRSLTPAVDRAYLVTGGTRGIGARVARLLVRRGARRIALTGARPQPPRADWPLLSPGTP  3547
                    + YLVTGGTRGIGA VAR LV RGARR+ALT A P PPR  W
Lrg  3372  APAEFPADPAKVYLVTGGTRGIGAEVARRLVERGARRLALTCAHPLPPRHRWSAPELSDR  3431

Lnm  3548  EAETASLVAELEAQGARVLVHSGPLSERERTDRFLREVREVLGPIGGVVHCAGRGPVGRP  3607
           EA     V  LE  GA+V++H    +       FL  VR  LGPIGGVVHCAG    GRP
Lrg  3432  EAVAVRNVQALERAGAQVMLHGEAPATETGLGAFLTGVRRSLGPIGGVVHCAGLPSRGRP  3491

Lnm  3608  SFIGKELADFDPVLEPKTTGLEVLDELCAGDRPEFFVLFSSLSAVAPGLAAGVLDYAAAN  3667
           SF+ K  AD   V EPKT G ++L+ LCA D+PEFFVL SS S + P L AGV DYAAAN
Lrg  3492  SFVHKTAADIAEVFEPKTAGADLLERLCAADQPEFFVLMSSASTLLPRLGAGVTDYAAAN  3551

Lnm  3668  AFLDCYADHQVRSGRPWFRSVAWPTWSESGMGADRPDSCAPVGVGPLGDEEGLRVLERIL  3727
           A L+  AD + R+GR  F +V WPTW  +GMGAD+PD CAP G+  +  EEGLRVLE +L
Lrg  3552  AHLEYLAD-RTRAGRTRFHAVHWPTWLGTGMGADQPDGCAPAGLDAITVEEGLRVLEAVL  3610

Lnm  3728  A--LPAEQARIVPCPPIDGIAADPAALLGSPRDTDATASVGSTTSAGSTPMAGSTPMAGS  3785
              LPA+   +VP P       DP  LL + R   A     + T+A
Lrg  3611  TGTLPAQLLPVVPLPG----RFDPETLLHANRAPAAPVDAPAATAAA-------------  3653

Lnm  3786  TPAAGSAPVPTTTGATPPRPREEEHTVPNTSVTGPPWLAPLFSELLAIPEDALDPTALLG  3845
                     T +GA            P      P WL  LFSE L IP   LD TA
Lrg  3654  ----------TFSGAPAATAAPTRSDAP------PDWLLALFSEALEIPLPDLDATAPFS  3697

Lnm  3846  DLGVESVLLGEILLRLEELTGLSLDPATLLDHPTLELLGRHLADLGV--PSAPPAPAATT  3903
           DLGVESV+LGE++  +E+  G SL+PA LL+H TLE L  +L   G+   +A   PA
Lrg  3698  DLGVESVMLGELVELIEQQVGSSLEPAVLLEHQTLERLAGYLRKAGLDRAAADAPPAPEP  3757

Lnm  3904  APAVAPVTPVTPTAPVAVAPVTPVAPSG-----KIAVIGLSCRFPGAEDAAAFARNLLGG  3958
            PA   VT   P +    APVTP A S      +IA+IGL CRFPGA D  AF   L  G
Lrg  3758  LPAAGSVTKAQPQS----APVTPPAASREPADRRIAIIGLDCRFPGAPDPDAFWAALAAG  3813

Lnm  3959  TCSVTEVPPSRWDVGELYRPELEPGRSTSKWGGFLDGIEDFDPEWFGMSEDEARCLDPAV  4018
             SVTEVPPSRWD   LYRPE   G S S+WGGF+DGIEDFDPEWFGM+E+E RCLDPAV
Lrg  3814  RNSVTEVPPSRWDHRALYRPEHRIGSSISRWGGFVDGIEDFDPEWFGMTEEEGRCLDPAV  3873

Lnm  4019  RLFLEGSATCLTDAGYGARELAGRDVGVFAGARMSHYGRRVGERRGLVGMGSDQNFIAAR  4078
           RL LEG+A C  DAGY   EL GR+VGVF GAR+  YGRR+G R G   +G DQNF+AAR
Lrg  3874  RLVLEGTANCFADAGYRTEELQGREVGVFVGARLGDYGRRIGLRSGPAALGGDQNFVAAR  3933

Lnm  4079  IAHHFDLHGPNLVVDSACSSSLVALQLACRSLLDGESELALAGGVDVLLDEEPYLDFSAA  4138
           +AHHFDLHGPNL VDSACSS+L A+QLACRSLL+GESELA+AGGVD+LLDE PYL+FSA
Lrg  3934  VAHHFDLHGPNLTVDSACSSALAAVQLACRSLLEGESELAVAGGVDILLDERPYLEFSAV  3993

Lnm  4139  KALSRHGRCATFDEDADGFVPGEGCGVVLLKPLEKALRDGDRIHAVIDAVAVNNDGRTMG  4198
           +ALS  GRCATFD DADGFVPGEGCG+VLLKPL +AL DGDR+ AV+DAVA+NNDGRTMG
Lrg  3994  RALSPTGRCATFDRDADGFVPGEGCGLVLLKPLARALADGDRVLAVVDAVALNNDGRTMG  4053

Lnm  4199  LTTPNPAAQAKVVRRALAAAGRRADEVGLIEAHGTGTMIGDPIELRALTEVFREETGRTG  4258
           LTTPNP AQAK +RRALA AG  A+ VG++EAHGTGTMIGDPIELRALT+V+RE T   G
Lrg  4054  LTTPNPVAQAKAIRRALATAGMSAERVGMVEAHGTGTMIGDPIELRALTDVYRETTDARG  4113

Lnm  4259  FCAIGSVKTNVGHLLSAAGMAGLIKAVLAVRDGRIAPTLFCERPNPRFDFAASPFYPSRT  4318
           FCAIGSVK+N+GHLLSAAG AGL+KAVLA+R+  + PTL CE PNPRFDFA SPFYP+
Lrg  4114  FCAIGSVKSNIGHLLSAAGAAGLVKAVLALRNRLLPPTLHCEHPNPRFDFADSPFYPNTA  4173

Lnm  4319  AHDWVPEPGRVRVAGVSAFGLGGTNAHAVVSQLDPVLAAAHRPRPALPAPN-FARRRLWL  4377
             DW P+PGR RVA VSAFGLGGTNAH +VS+ D    AAH P         F RRRLWL
Lrg  4174  LRDW-PDPGRPRVAAVSAFGLGGTNAHLIVSEPDENAVAAHPPTRRPLPRPVFDRRRLWL  4232

Lnm  4378  EA  4379
           EA
Lrg  4233  EA  4234
```

LrgJ *vs* LnmJ

```
Lnm  15   CHLVLEHSDFIMQNHRVHGVSVMPGVTFLDIVFRILRDRGFDTARAELRNVLFHEAIATS  74
          C LVL H DFIMQNHRVHGVSVMPGVTFLDIV R+L  +G D    ELR ++F EAIAT+
Lrg  8    CRLVLRHDDFIMQNHRVHGVSVMPGVTFLDIVLRVLAAQGLDPTTVELRGIVFAEAIATA  67

Lnm  75   EGCDRDIRITVSTSTDGSRWITAESRRREGGESAADYQENFRGELVLHDVPEPGPLDVDR  134
          EG DR+IR+ +    DG R +T  SR   G E   +++EN R EL          LD
Lrg  68   EGDDREIRVVIGEPADGVRPVTGTSRWLRGDEPYGEWRENIRAELHPAGPSAVPDLDAAA  127

Lnm  135  LRTTARRVADLDEMYARARAEEIRHGSAMRCFGRLYYGDGELLAELGLDGEAAALDEHFH  194
          L   A R   + +MYA  R+ EI HG+AM+C G ++ G   LLAEL L       D  F
Lrg  128  LVAGAVRTRSMADMYAHTRSREIVHGAAMQCAGPMHLGADHLLAELSLVLPETGEDRAFL  187

Lnm  195  LHPAKMDCATIAAFAQVPPPDQDPFIPVFIESFRAPRPLTGVAYAHMPRPETYAQSGDIM  254
          LHPAKMD ATI A+ Q      +PFIP+FI+ FRA  PL G    H+P+PE    SG++
Lrg  188  LHPAKMDAATIVAYGQREITAAEPFIPMFIDRFRAHGPLHGAFLVHVPQPEELTPSGELF  247

Lnm  255  HNDCALYDADGRFLAGFTKLTCKRIRNPELITRLLDAPDVTRTAAPAPAAVS-PSPVVAP  313
           +  +L+D  GR +A F +LTCKRIR PE I  LL    +  T A  PAAV+ P    AP
Lrg  248  RSTFSLHDRQGRTVAEFDRLTCKRIRRPESIRDLL----LETTGARQPAAVAGPQAPTAP  303

Lnm  314  ASSDGGAGPDAVRAHLRELVGTLLGRAPHAIRTDAGFYDLGLDSGHMLDISRRLEEYVCA  373
          A+          R  LR  +  +L R+   +    GFYD+GL S  ML IS  LEE V +
Lrg  304  AAD--------YRHWLRGRIARMLDRSADTVDDGLGFYDMGLTSVDMLRISNELEEVVGS  355

Lnm  374  PLYPTLLFEFSDIDSLAAHLYAEFGAQVRSAPANPPATPATPGEDAGAPPASAARSTARA  433
           LYPTLLFE + +D LA HL   +GA    AP                    A    AR
Lrg  356  ALYPTLLFEHTTVDGLARHLEETYGAPAAPAPQPSAPAAPA---------EPARPHRARL  406

Lnm  434  VAPALGCHRPVWTPLPADPGAFAADGARTVVLVGADAATAAALRDAAAPARVVRAERASA  493
          +       RPVW  +         D    + ++GAD A  A L            ERA+A
Lrg  407  L-------RPVWQRVAHSAPTVPTD----LAVLGADPALLAVL-----------TERAAA  444

Lnm  494  FQRLAADHYRLDPADPDQL-ASLTAALATDGISATAYVRCARTRDTDGAGSALPD---AY  549
             L      +DP D   L A L+A     G S T         D  G G +  D
Lrg  445  TGALVV---AVDPGDRALLGARLSALPERRGRSWTLL-------DATGLGGSSADPAEVA  494

Lnm  550  LESWALAVAVTGTRPTGPVPVLFLHPRDPAAPRPHEDALGALARTVAAEAPQLRCRAVGH  609
          + +WA   A +  RPT    VLF H R     +P   A+ AL RTV AE P L  RAV
Lrg  495  VAAWAACAAASERRPTRVRAVLFAHRR-----QPEYAAVAALGRTVTAELPALAVRAVEV  549

Lnm  610  DATATAGDLAAVIAAESTDLSAESEVRHTGGTRLTVRHETLAVPSGNGAGVLREDGVYLV  669
           A   A +LA ++ AE+ D SAESEVRH GG R   R + +  P+  G   LRE GVYL+
Lrg  550  AAAGPA-ELAELLLAEAADPSAESEVRHAGGERSARRFQVVESPARAGEAGLRERGVYLI  608

Lnm  670  TGGGGSLAALLVDRLVTRGPVRLVLTGRSAPGPELTQRIEGWRRRGAEVTHVRGDVAHTD  729
          TGGGG L  +L + L      R++L+GRSAPG  L  R+  W+R GA+  +   DV
Lrg  609  TGGGGGLGPMLAEHLARTRRARILLSGRSAPGEALLGRMREWQRYGADAQYRTADVTDPA  668

Lnm  730  DVLAAVTCARETYGRIDGVFHCAGSVDDGMFFRKDPERSAAVLAAKVAGTRNLDEATADD  789
           V   V  ARE YGRIDGV HCAG V DG+F  K PE+   VLA K++G R+LDEAT  D
Lrg  669  QVRELVAAARELYGRIDGVVHCAGVVRDGIFLAKRPEQIREVLAPKISGIRHLDEATRAD  728

Lnm  790  GLAFFALFSSVSASVANPGQADYAYGNAFMEHFAEQRAARADRPGVSVAVGWPLWADGGM  849
          GL F A +SS+SA + NPGQ+DYAY NA+++H+       A RPG S++V WPLWA+GGM
Lrg  729  GLDFLAAYSSLSAVIGNPGQSDYAYANAYIDHYL------AARPGRSLSVDWPLWAEGGM  782

Lnm  850  RVSEDVLRRSADTSGLHALPADAGLDALFGLLSGAAPRAVVTYGDQERIAELLP------  903
          RV  +V  R+A   G   LP   G++     L+    R VVT+ D  R  + LP
Lrg  783  RVDAEVTERAARAHGASPLPTGTGVELFERALAAGDSRLVVTHADPARSDDRLPLADGDL  842

Lnm  904  -APRPSAAQSGRTG-------SPDSPD----SPDG-------------DDIAIIGVAGRY  938
           AP  +A  +   G       +PD+      +PDG             D IA+IG+AG+Y
Lrg  843  AAPDTAAPDTAAPGLAAPGLATPDTAAPGFAAPDGHGPDAVTPGEHGPDAIAVIGLAGQY  902

Lnm  939  PEAEDLEAFWRNLAEGRDCVGEVPADRWDHAAYYDPERGKEGRTYGRRGGFLDGVDRFDA  998
          P+A D++AFWR LAEGRDC+ EVP +RWDHAA +DPE G+ GRTYGR GGFLDG+DRFD
```

```
Lrg  903   PQAADVDAFWRLLAEGRDCITEVPRERWDHAAIHDPEHGRPGRTYGRWGGFLDGMDRFDP  962

Lnm  999   ASFGISRREAELMDPQERLFLTVGRQAVENAGYRPEELARTRVGVFAGVMWNHYQLCTDG  1058
           A FGISRR+AE MDPQERLFLT   Q ++ AG+         VGVFAGVMWNHYQL   G
Lrg  963   AFFGISRRDAERMDPQERLFLTTCWQTLQEAGHPASRTTAAPVGVFAGVMWNHYQLV-QG  1021

Lnm  1059  SAEPVAPTALHCSVANRLSYCLDLSGPSMAVDTACSSSLTSLHLAVESIRRGECALAVAG  1118
           + + V PTA+H +VANR+SY L+L+GPSMAVDTACSSSLT++HLAVES+RRGEC +A+AG
Lrg  1022  AEDGVQPTAMHAAVANRVSYTLNLAGPSMAVDTACSSSLTAIHLAVESLRRGECEMALAG  1081

Lnm  1119  GVNVAAHPQKYLQLAQGRFLSSDGRCRAFGADGDGYVPGEGVGAVLLKPLADALADGDHV  1178
           GVNVAAHPQKYLQLAQGRFLS DGRCR+FGA G GYVPGEGVGAVLLKPLA A ADGDH+
Lrg  1082  GVNVAAHPQKYLQLAQGRFLSDDGRCRSFGAGGTGYVPGEGVGAVLLKPLAKAEADGDHI  1141

Lnm  1179  HAVIKGSFLNHSGRTSGFTVPSPAAQATLIADALDRSGVAADSVGYIEAHGTGTALGDPI  1238
           H VI+ + LNH+GRTSGFTVPSP +QA LI  ALD +G+    +GY+EAHGTGTALGDPI
Lrg  1142  HGVIRATRLNHTGRTSGFTVPSPTSQAALIRAALDAAGLPPSGIGYLEAHGTGTALGDPI  1201

Lnm  1239  EIEGLRQAFADAGLAPGSCAIGSVKSGIGHLESAAGIAAVTKVLLQMRHRELVPSLHSEQ  1298
           EI+GLR+AFA AG   G CAIGSVKS IGHLESAAGIA VTK LLQ++HR+L PSLH+E
Lrg  1202  EIDGLRKAFAGAGTGSGGCAIGSVKSNIGHLESAAGIAGVTKALLQLKHRQLAPSLHAEV  1261

Lnm  1299  PNPHIDFAATPFAVQRTRAPWVPRPGSTVLRAGVSAFGAGGSNAHVLLESA-PPAPATPV  1357
            NP ID A TPF +QR    W      +  RAGVSAFGAGGSNAH++LE    PA
Lrg  1262  VNPAIDLATTPFRLQRELTDWPAPADGSPRRAGVSAFGAGGSNAHLVLEEYRAPADRPAR  1321

Lnm  1358  AGPQLFVFSAKDERTLREVVRRQLRHLDGPG-PVGSSADEATALLTGEVAALLDVPVDAV  1416
            G +L V SA+D   LR    R    L G   P  + A E   +LTG VA +L VP DAV
Lrg  1322  GGRELIVLSARDADALRVYAERIRAVLHGAAEPSTAGAAELRRVLTGAVAQVLGVPDDAV  1381

Lnm  1417  DVRENLADLGVDRLALAELGRRVEGRLP---AGVPLSGQASVTELAASAALAARPDALPL  1473
           D  E L DLGVDR  LA + + ++   P     +P+ G  S+ +LA       RP A+PL
Lrg  1382  DPAEPLVDLGVDRTGLALVRKVLDEHRPGLGGPLPVEGDRSIDQLAGQLGTEPRP-AVPL  1440

Lnm  1474  ADIAHTLRVGRSPLAVRLAVVCGEPEELRRRLAAFLDGDEPGEGVFTGRADDDKEPVRLE  1533
           A +AHTLRVGR  L+ RLAV+  +  EL   L  +L G  P  G + GR        R E
Lrg  1441  AALAHTLRVGRDQLSSRLAVLAADHAELLAALDRYLSGAAPEAGQYWGREGAGTAEPRPE  1500

Lnm  1534  RAAELFRLGRLSELARAWADGAAVEWDDCRAGDGVRPRRVPLPAHPLDERSYWIGGWR  1591
             AEL R GRL E+A AW+ GA V W DC A  G  PRRV LP  PL E  +W+GGW+
Lrg  1501  -LAELVRAGRLEEVAAAWSAGADVPWADC-APTGPAPRRVSLPVPPLREERHWLGGWQ  1556

Lnm  1605  TAHAPTAVPAAP---------------VDQADLPEVREERPGL-DPQEVLWAVVDAVRTR  1648
           TA AP  VP  P                   D+PE      GL DP EV   VV  +
Lrg  1845  TAAAPVEVPVTPQAAPPAEVPVPVPAAARPVDVPEA-----GLPDPAEVERLVVATLCGI  1899

Lnm  1649  LYLERDEVDHRLSFNEMGVDSVGAVEIVEQLGARFALEMDPVTLFDHPTVPRLAEHVREL  1708
           +Y   DE+D RLSF+E GVDS+GAVEIV  L  RF L++D V ++DHPTV RL  HV E
Lrg  1900  VYATEDEIDRRLSFSESGVDSIGAVEIVRSLNQRFGLDIDSVAVYDHPTVARLTAHVLET  1959

Lnm  1709  HRQSPAPRPQA-------------------------------APA----------------  1722
             Q+ A    A                                 AP
Lrg  1960  AEQARALHRSALTQAPAPAPAPASAPEPAPAPVAAAPEPPPAPEPPAPEPPAAPPVRLAP  2019

Lnm  1723  --------------------------------PAQPAAPEAAALPAPAPAPAPAPAPASK  1752
                                          PAQP  P   A P      AP    A  S+
Lrg  2020  PAAAAPVAPLAPATPGQTSLRPLRGTPVQQPAQPLQPAQPAQPVRLAPLAPRAEGAEPSR  2079

Lnm  1753  PEPAASPDACDDIAVIGMSGRFPGAEDLDAFWENIAAGRDSFTEVPAQRWDVGPVFDADR  1812
           P PA   D  D IA+IG++GRFP AEDLDAFW N+AAGR S +EVP  RW  G  FD DR
Lrg  2080  P-PA---DDADGIAIIGVAGRFPDAEDLDAFWANLAAGRTSISEVPEARWGTG-WFDPDR  2134

Lnm  1813  LVPDRTYSKWAAMLPEVGRFDAAFFNHSPLEAEVMDPQQRLFLEQSWAALEHAGYAVGAD  1872
            VPDR+YS+WAA+LP++G FD  FF  SP+EAE MDPQQR FL+Q+W ALE AGYA
Lrg  2135  RVPDRSYSRWAALLPDLGGFDPRFFQLSPMEAEAMDPQQRQFLQQAWTALEDAGYA-APG  2193

Lnm  1873  DRTSCGVFVGCAPGDYSTLLTEAGRADTGHAFLGTTSSLLPARIGYFLNLDGPTMAVDTA  1932
            R  CGV+VG + GDY  LL  AG+ADTG AFLG   ++L AR+ Y L+L GPTM VDTA
```

```
Lrg  2194  KRLRCGVYVGASGGDYYHLLRAAGQADTGQAFLGNNMAILAARVAYLLDLSGPTMTVDTA  2253

Lnm  1933  CSSSLVAVHLAADSIRRGECAMALAGGVALMVTPQLHVRASKVGMLSPRGTCVPFDASAD  1992
           CSSSL AVHLA ++IR G+C +A+AGGVA+M TPQ+ V +S+VGMLSP G  VPFDA AD
Lrg  2254  CSSSLTAVHLACEAIRGGDCELAVAGGVAVMTTPQMQVWSSRVGMLSPTGRSVPFDAGAD  2313

Lnm  1993  GTVLGEGVGAVVLKRLDRAVADGDHIHGVIKATGVNGDGRTNGITAPSALSQAALIADVH  2052
           G VLGEGVGAVVLK L  A+ADGD I  VIKA+GVNGDGRTNGITAPSA SQA L+  VH
Lrg  2314  GIVLGEGVGAVVLKSLRAALADGDRIQAVIKASGVNGDGRTNGITAPSATSQAELLRAVH  2373

Lnm  2053  RRAGVGADDIGYVEAHGTGTALGDPIEVRALTEVFRRSTDRSGYCGIGTVKANIGHTTMA  2112
           RRAGV A DIGYVEAHGT T LGDPIEV+AL +V   + D  G+  +G+VKANIGHTT A
Lrg  2374  RRAGVEAGDIGYVEAHGTATNLGDPIEVKALNQVL-GAADGPGFTALGSVKANIGHTTTA  2432

Lnm  2113  AGIAGLLKTLLALRHSELPPAPAFDTPNPKTELDSSPFFVVRDRQEWEPGPGGQRIATVS  2172
           AGIAGLLK +LALRH  LPP P F   NPK +L      VVR+   WEPG  G R+ TVS
Lrg  2433  AGIAGLLKVVLALRHRALPPLPGFAEANPKLDLSGGRLRVVRELTPWEPGANGVRVGTVS  2492

Lnm  2173  SFGFSGTNAHAVLAQAPEPQARPEEPDQER-LFAVSARDGAALDRLLLRLADS-DLD-GV  2229
           SFGFSGTN H VLA+ P   A    P   R L  +SAR   AL R+   LA++ + D  +
Lrg  2493  SFGFSGTNCHVVLAEPPARPAPERRPRAARHLVPLSARTPEALTRVAADLAEALEHDPAL  2552

Lnm  2230  TPADLAFTLGVGRAHLPVRAAVIARNVPELRRRLRLLQSGAQAPGCFRTGQGAAAGDLDE  2289
            PAD+++T  +GRAHL VRAA++     EL  +LR L  G ++PGC      A +  LDE
Lrg  2553  EPADISYTRALGRAHLTVRAALLTGGREELLEQLRKLADGQESPGCHL----ADSAGLDE  2608

Lnm  2290  QTRAELAGRARSGSPAERVAALERLAAAYAAGQDLDWQSLSYGDRPRRVPLPTYPFGGDR  2349
                    A S  P      L R AAAY  G   DW +L+ G   RRV +PTY F  +
Lrg  2609  A--------AGSDDP------LVRAAAAYVRGDTPDWAALTAGG--RRVAVPTYRFAREH  2652

Lnm  2350  HWITLP-DTDRTAVPATAPATSRVDLPGQSPQSTPHPLLGAVSGAPGDPDGARFPVPVPA  2408
           +W T P  TD    P TA       LP  + Q+ P P        P  PDG    V V
Lrg  2653  YWATGPAATDHVVPPGTADDHR---LPAPTRQAAPEP------DRPAGPDGGADGVAVVR  2703

Lnm  2409  AHW-VLDHHRIGDRPVLPGAAGLDLAVAAARRCGLRGTVRLHGVQWLRLIDAEAAGTLRL  2467
               V   HR+  RP LPG A L LA   A         RL  V+WLR  +      LRL
Lrg  2704  PEDPVATDHRVAGRPTLPGTAALALAAGLAGL-----PYRLSAVRWLRPCELSEPRRLRL  2758

Lnm  2468  TLTSDGEGYRFAL-STGDDGTVCSRGSLTVRPDAQDAAAPAASSETLDVAEIAARCPYEV  2526
                  G  F L + G DG     G   +    ++ AA  A+   LD+   A RCP
Lrg  2759  AGEESAAGRSFRLEAEGADGPYVRGGFAAL---TEEEAAAFAAEPPLDLRLTADRCPAAR  2815

Lnm  2527  PAERFYDDFRSGGIAYGPSFRVLEKITFGDDEVLGTLRATPDSGGFALHPALLDGAQQTI  2586
            A+  Y  FR+ G+AYGPSFR LE++  GD E LGTLR      G+    A+LD   Q +
Lrg  2816  SADEVYGAFRAAGLAYGPSFRRLEQVRVGDGETLGTLRPADGPAGWQALAAVLDAGLQVL  2875

Lnm  2587  A-ALEGGNDATLVPFSVETVEVVDATAVPAFAHVVRAGKHRYTVRLADRSGRVCVRYEGL  2645
           A  L+      L+PF+V+ V V+ +  +  ++H  R G+ R+TV L D SG +CVR++G+
Lrg  2876  APLLDADGPQALLPFAVDRVTVLRSPGLARYSHARRTGQDRFTVSLTDGSGALCVRFDGV  2935

Lnm  2646  ALRAQHNPVDSM--------MYRPVWRPAPLPQPGNAPAGGRTVVVHTADSTTLAAALAA  2697
           +LRA   P  +         ++RPVW  A  P P  A  GG  ++ H   +  LA AL
Lrg  2936  SLRAVPRPAAAAAAVPEGPGIFRPVWEDA-GPAPAEAAGGGTVLICHPEAAGALAGALTD  2994

Lnm  2698  RTGAGLVALSGAQDAAPDPYAVLEQPLETVYFVARTGDAEGPAEADRTALDLFRLVKRML  2757
                 V   G  D       V E P + VYF+        P E DR+   + RLV+R+L
Lrg  2995  VHRNCHVRTVGHHDID----TVTEVP-DRVYFLTDPAPVHRP-EQDRSVPAMLRLVQRLL  3048

Lnm  2758  AVGRARDRIALRIVLAGAVPADPEDMTETVRPHAAGVLGLARAIESECPRWSVACVDVGA  2817
           A+G A   +A+R VL GAV        E +RPHAAG+LGL     +E PRW+V CVD G
Lrg  3049  ALGAAHTGLAVRAVLFGAVAV---TGGEPLRPHAAGLLGLCATTAAEYPRWTVGCVDAGT  3105

Lnm  2818  DGGTVGAERAAERIVAEPGTEPLVLLRGEERLERVFEPLRPAAPRGTEPFREGGVYVIVG  2877
                G  + A  +V EP  + L+ LR  +RL R      PA   G  P+REGG YVI+G
Lrg  3106  TPAAPG--QLAALLVREPAADRLIALRDGQRLRRTLLAASPAG--GEPPWREGGAYVILG  3161

Lnm  2878  GAGGIGFALSRLLARIARARLVWIGRSPEGPEHRAKAEEIAALGGQVLYVQADVADEAAL  2937
           GAGG+G AL R LAR  RARL  IGR  + P   A   EI  LGG+ +Y++AD AD A L
```

```
Lrg  3162  GAGGLGRALGRHLARTHRARLALIGRRAQDPAIDAALAEITELGGEAVYLRADAADPAQL  3221

Lnm  2938  RRGLASVHTRFGQVDGAVHAALDLRDRTIALMDEEDFLAGLAPKVAGVTAFARVFGAEPL  2997
            R +A+   RFG ++GAVHAALDLRDRT+   D + F   LAPKVAG  AFA     +PL
Lrg  3222  ARAVAAARDRFGTLNGAVHAALDLRDRTLLHADPQTFGEVLAPKVAGTAAFADALRGDPL  3281

Lnm  2998  DFMLVFSSAVSFVEAGGQANYAAASTFEDAYVQWLDRRHDYPVSVVNWGFWGSVGAVADD  3057
           D + VFSSAVSF ++ GQA YAAASTF+DAY QWLD R  YPV V+NWGFWGSVG VAD+
Lrg  3282  DLLAVFSSAVSFTDSPGQAAYAAASTFQDAYAQWLDSRVPYPVQVLNWGFWGSVGVVADE  3341

Lnm  3058  RMRAAFARLGVGSVEPAEGMAVLRRIIAGRLPQTLAMKADRAALPAMGIR  3107
           R     A  GVGS+EPAEG+A L R++A  LPQT+ +KAD   L  +G+R
Lrg  3342  RYGERLAAFGVGSIEPAEGLAALDRVLAAGLPQTVVVKADARGLARLGVR  3391


Lnm  3275  FARAQEAFAAVEAFSRDLLRRTFPRLDGVPRPGERITADELASRLGVVRRHRRLFDAALS  3334
            ARA+  FAA++  + DLLR  F  L  +P   E  T +  A RLG V R R +  A L
Lrg  3398  LARARAGFAALDVVAADLLRAEFAALPELPPLEEPSTLEAFAGRLGAVGRDRSVLGAVLR  3457

Lnm  3335  ILRSCGAVTGDAD-TLTFAEPSAAPGARVEGAEVAALYPEMSGHVTLLERTLGALGEVLA  3393
           +L   GA T D D TLTF      P AR+   E AA +P M  H+TLL+  +  +  +L+
Lrg  3458  VLERAGAATLDRDGTLTFRRALLDP-ARLPVVEFAAAHPAMVPHLTLLQACVAGVPGILS  3516

Lnm  3394  GRRNPMDVLFPKGSVALVEPIYKGQPIADHYNRLLADEVADAARRVRAQEGRPVRVLEIG  3453
           GR    +VLFPKGS ALVEP+Y   P A+H++RL+A EV  +A+R+   E RP+R++EIG
Lrg  3517  GRTAATEVLFPKGSPALVEPVYADGPGAEHFHRLMAAEVVGSAQRLTGGE-RPLRIVEIG  3575

Lnm  3454  AGTGASSRTVLAALAAADAGAHYCYTDISPAFLRHGEREFGPTYPQLAFHTLDISRDPVE  3513
           AGTG+++R VLAA AAA     Y YTD+SPAFLRHGE   G   P + +  LDI RDP
Lrg  3576  AGTGSATRHVLAACAAAGVPTAYRYTDVSPAFLRHGEA--GHRAPGMRYELLDIERDPAA  3633

Lnm  3514  QGIEAASYDVILGTHVLHATPDMERTLRNIRTLLRPGGLVLVNEITRFSEFLTLTFGLTT  3573
           QG E    DV+L T+VLHAT D+ RTL N+RTLLRPGG++ VNE+TR SEF+TLTFGLT
Lrg  3634  QGFEPGCADVVLATNVLHATRDIGRTLANVRTLLRPGGVLAVNEVTRSSEFVTLTFGLTE  3693

Lnm  3574  GWWMYEDAQCRLPHSPLLAPVQWRQSAAAAGLRTVRTGGLPGVPADELEQSLVVAERPVE  3633
           GWW +ED Q RLP S LL P QWR     AG R     G+PG P DELEQ LV +ER +
Lrg  3694  GWWRFEDPQRRLPDSALLGPAQWRACLTEAGFRVTGVRGIPGTPEDELEQCLVTSERELN  3753

Lnm  3634  DSGDASPDGAADEQSPESVRSYVTGVFAEVLKYRAEDLDPAVTLENFGVDSLVSLNIVDR  3693
            + + +P   A +     VR YV  VFAEVLK+RA +LD   T E +G+DSLV  NIV R
Lrg  3754  VTAEQTPVPTAAQ-----VRGYVRQVFAEVLKFRAAELDDHATFETYGIDSLVGQNIVYR  3808

Lnm  3694  LEQDLGDLPQTLLFEYTSIDSIAEYLSAEHGERLARVLGGAPAAAQAQPSAPVPAPVSVD  3753
           +EQDLG LP TLLFE+ +ID +A +L  +  ERL  +LG A        ++PV  PV
Lrg  3809  MEQDLGALPATLLFEHLTIDQLANHLRTDRAERLTALLGPAAPPVAPPVASPVAPPVEAA  3868

Lnm  3754  VPAPIPAPIPA-------PVSVPVDVQEPESEPEPAAAVRTPAGDDPADPLDIAVIGVVG  3806
            PA      P          + PV V EP + P PAA           +P DIAVI V G
Lrg  3869  PPAQPGPVAPVAPVAPVVQATPPVPVAEP-ARPGPAAH---------GEPADIAVIAVSG  3918

Lnm  3807  RYPQSPDLEAFWRNLSEGRSCITEIPSERWDWRRNFDPDKSRKHRSYSRWGGFLEDIEMF  3866
           RYP +PD+E FWRNL +GR  +TE+P++RWDWR  FD  + R  RSYSRWGGFL+DI+ F
Lrg  3919  RYPGAPDVETFWRNLEDGRDAVTEVPADRWDWRPTFDAQRGRGDRSYSRWGGFLDDIDKF  3978

Lnm  3867  DAPLFGILPRDAADIDPQERLFLESCWELLETAGYLGTYTHEPQTGVFAGLMYGEYGLLA  3926
           D   F ILPRDAADIDPQERLFLE+CW+LL+ AGYLG  THE  TGVFAG+MYG YG LA
Lrg  3979  DPAFFNILPRDAADIDPQERLFLETCWDLLDRAGYLGGSTHETMTGVFAGVMYGSYGRLA  4038

Lnm  3927  AATDWPEGRYATGHSAYWSMANRVSYTFDLQGPSLAVDSACSSALSAIQLACESLRRGES  3986
           A T W  G+ +  HSAYWS+ANRVSY FD QGPS AVDSACSS+L+A+ LA ESLRRGE
Lrg  4039  A-TGWAHGKLSGAHSAYWSVANRVSYHFDFQGPSFAVDSACSSSLTAVHLAVESLRRGEC  4097

Lnm  3987  RMAIAGGTNLILHPAHFAALCARNMLSAADACRVFDDGADGFVPGEGAGAVLLKPLAQAE  4046
           RMA+AGG NLILHPAH  +L A NML+   AC+VFD+ ADGFVPGEG GAVLLKPLA AE
Lrg  4098  RMAVAGGVNLILHPAHHVSLSALNMLAGDGACKVFDERADGFVPGEGVGAVLLKPLADAE  4157

Lnm  4047  ADGDTIWGVVKGAFSNAGGKVSGYTVPNPNAQARLVERTLRRSGVHPRTVSYVEAHGTGT  4106
```

```
            DGD I  V+KG+  NAGGK  GYTVPNP AQA L+   +RRSGV PRT+  +EAHGTGT
Lrg  4158  RDGDEILAVIKGSTVNAGGKTGGYTVPNPQAQAALIAEAVRRSGVDPRTIGSLEAHGTGT  4217

Lnm  4107  ALGDPIELGGLTKAFRAAGAT-GDGYCAVGSVKSNIGHLEGAAGIAAVTKVLLQLKHRAL  4165
           ALGDPIE+  LT+AF   GA  G+  CAV SVK+ IGHLEGAAGIA +T+ LLQL+H  +
Lrg  4218  ALGDPIEIAALTRAFEELGADPGEFRCAVSSVKAAIGHLEGAAGIAGLTRALLQLQHGRI  4277

Lnm  4166  APTIHLDRLNPKIDFAGSPFGPQRTAEPWDRPVAGVDGAERSWPRRAGISSFGAGGANVH  4225
              ++L+ +NP+IDFAGSPF P R    W  P  G        PRRAG+S+FGAGGAN H
Lrg  4278  TRCVNLENVNPRIDFAGSPFYPPRETAAWPAPADGS-------PRRAGVSAFGAGGANAH  4330

Lnm  4226  MILEEYTGQDPRDAEDTMGAAGAEEPELFVLSALDRETLARHAGRVADFVAGPEGARVRL  4285
           ++LEEY    PR A  T      E  +LF+LSA  R  L R+A RVA+ +A PEGA + L
Lrg  4331  VVLEEYR---PRTAPATPRLPDGE--QLFLLSARTRGQLVRYAERVAELLATPEGAELPL  4385

Lnm  4286  ADLAHTSRVGRRELPERLAVTAASHAQLAARLREFAATGVAGEGVSTGTARKGGAGSGL-  4344
           A LA TS++GRRE+ ERLAV A   +QLA RL +F A G    GV  G+A     G  L
Lrg  4386  AALARTSQIGRREMAERLAVLATDTSQLADRLHDF-ARGAESAGVVVGSAGGDSGGWSLL  4444


Lnm  4345  ----GAQELTAALAGRRWADAVEHWTLGGRVDWRTADAGRLVRKVAFPTYPFNRSRHWI  4399
               GA  +   LA R+       WT G  VDW+        R+V  P YPF RSRHW+
Lrg  4445  DDEDGAALVATILAKRQLPKLARLWTAGVPVDWQLCWTAPHPRRVQLPPYPFERSRHWL  4503


Lnm  5385  PAARAEIAVIGIAGVFPGSADTDEFWEHLAGGVDLVRPVPKDRTAIRANPATRELRGGFL  5444
           P   A IAV+GIAG  PGSAD DEFW HLA G  LV PVP DRT +R +P TRELRGGFL
Lrg  5224  PYGPAPIAVVGIAGRLPGSADLDEFWRHLAAGDHLVGPVPADRTDLRQDPETRELRGGFL  5283

Lnm  5445  DSVDTFDARLFGISPNEAALMDPQQRLFLQTAWRVFEDAGYRPADLAGAPCGLFVGVATH  5504
           +++  FDA  FGIS  EA LMDPQQRLFL+  WR  E+AGY P++LAG+  GLF GV+T
Lrg  5284  ENIADFDAAFFGISATEAGLMDPQQRLFLEVVWRAVENAGYPPSELAGSATGLFAGVSTT  5343

Lnm  5505  DYDDLLKENGVAVQAHTATGIAHSVLANRVSYLFDLNGPSEAVDTACSSSLVAIHRALRA  5564
           DYDDL++ NGVAVQAHTATG++H+VLANR+S L DL GPSEAVDTACSS+LVAIHRA+RA
Lrg  5344  DYDDLMRTNGVAVQAHTATGLSHAVLANRISRLLDLRGPSEAVDTACSSALVAIHRAVRA  5403

Lnm  5565  IQDGECELAVAGGVNVILTPGLLESFTQSGMLSPDGRCKTFDADADGYVRGEGVGAVLLK  5624
           I DG+C+LA+AGGVN  L+PGL  +FT+SGMLSPDGRCKTFDA ADGYVRGEG GAV LK
Lrg  5404  ILDGDCDLAIAGGVNATLSPGLFTAFTKSGMLSPDGRCKTFDAAADGYVRGEGAGAVELK  5463

Lnm  5625  PLARAEADGDHIYAVVKGTAVNHGGRSNSLTAPNPESQARVVAAAVREAGVEPDTITYIE  5684
            L RA+ADGDHI+ V++ TAVNHGGRS SLTAPNPE+QA+V+  A R A + PDT+++IE
Lrg  5464  RLDRAQADGDHIHGVIRATAVNHGGRSTSLTAPNPEAQAQVLVQAYRRAALSPDTVSHIE  5523

Lnm  5685  AHGTGTRLGDPIEIEGLKKAFTTLHEERGEAVPDTGRIAIGAVKTNIGHLETASGIAGVL  5744
           AHGTGT LGDP+E EGLK+AF  L  E G        IA+GAVKTNIGHLE ASGIAGVL
Lrg  5524  AHGTGTSLGDPVETEGLKRAFAQLSAEAGLPPVRRPHIALGAVKTNIGHLEAASGIAGVL  5583

Lnm  5745  KVVQSMRHRVLPASLHLRRLSPYLRLDGTPFTVNDRHRPWEPALTPDGRQVLRAGVSSFG  5804
           K + S++HR LPA+LHL  L+PYLRLDGTPF VNDR  PW+      GR V RAGVSSFG
Lrg  5584  KTLLSLKHRQLPATLHLTELNPYLRLDGTPFYVNDRTAPWDGVDDGTGRTVRRAGVSSFG  5643

Lnm  5805  FGGSNAHVVLEAYPART-APAVQDFAPHTVPLSAGDPDDLRGYAARLARHLARTPEADLA  5863
           FGGSNAHVVLE Y   T  P     A    PLSA     LR YAA LARHL   P+A+ A
Lrg  5644  FGGSNAHVVLEEYRDETPTPTELPSAATLFPLSAPTAAALRDYAATLARHLEANPDAEPA  5703

Lnm  5864  RVAYTLQTGRTGHRHRFAVRVRDRDELIGALEAFAAGELPDHAATGTARRDAPSVQSDED  5923
           R A+TLQTGR  H  R  +   DRD L+  L A A G+L D                 D
Lrg  5704  RAAWTLQTGRDAHAERVVLAAADRDRLLARLGAVARGDLID---------------CPPD  5748

Lnm  5924  PALLRKSWCEGADVPWHTWWPKT-PGRVPLPTAPFARTRHWF  5964
           PA   + W +     W + WP T P R+PLP  P A  RHWF
Lrg  5749  PAA--REWLDTGRTAWASHWPATRPRRLPLPGFPLAPVRHWF  5788


Lnm  6153  EDLVQDVIERELGRTADPAKSFVDNGFGSFDMLRVVASLERVFGALRKTLLFDHPTIGAL  6212
           ++L+ D+I    GRT  P  +F D+G  SFDMLR V++LE+ FGA RKTL+FDHPT+ AL
```

```
Lrg  6119  QELLVDLISTISGRTLTPGVTFADDGLTSFDMLRTVSALEKRFGAQRKTLMFDHPTVPAL  6178

Lnm  6213  AAHLAETHGP-------EAASRLSSPPQDRPRPGPAASQ-EPYTGGALVVEKKALAGQPE  6264
             HL   +GP       EA S + + P+    P  AA   E  +GG  V+ K+ L  +PE
Lrg  6179  TGHLLAEYGPATAGGLREALSSVVAGPEQAAAPTKAAVPPEQSSGGWTVLRKRQLPQRPE  6238

Lnm  6265  LASAVAGLESAYGREGGLPGRDIAPLIFLGAGRTAYFNFSVRDDAMLTWSYVGPTEEMPA  6324
           LA+ VA +++ + +EGGL GRDIAPL+FLGA R  YFNFS +   +  WSYVG  E  P
Lrg  6239  LAAVVAEVDARWAKEGGLAGRDIAPLMFLGAEREGYFNFSRKGGVLFAWSYVGSREYFPK  6298

Lnm  6325  LATEFVRYGQAHGLAANIVSLIRLEEVDGVRFTATPFGALQRLEGIKDFSLEGGRMQRLR  6384
           L  ++V Y   HGL  N +S+  +E V G  F+ATPFGA+QRL+ +  F++EG RM+RLR
Lrg  6299  LIEQYVAYADRHGLQPNFLSVEHIESVAGRPFSATPFGAVQRLDDLSTFTMEGTRMRRLR  6358

Lnm  6385  YAVRKFEKAGTCRTEEYAVGSDPRTDQEITTLIDRWSAAKEMVNPYVSTVRDEIGRGILA  6444
           Y V +F  AG C TEEY VGSDP  DQEI  ++ RW   K+MVNPYV+ V +EIGRG LA
Lrg  6359  YMVNRFTAAGECATEEYRVGSDPAVDQEIVAMMSRWGETKQMVNPYVAVVSEEIGRGQLA  6418

Lnm  6445  ARHRMFLTYLDDRMVSAVIVTKIPSEDGYLLDLEFYPEDAPLGSLDHTVVKIIERTAAEG  6504
            RHRMFLT +D  + SA+IVTKIPSE G+LLDLEFYP++APLG L+  +V+IIE+ AAEG
Lrg  6419  ERHRMFLTRVDGELASAIIVTKIPSESGWLLDLEFYPKEAPLGGLEFAIVRIIEKLAAEG  6478

Lnm  6505  CTVFSFGGSFGAKVCESPNAAPEAEAALTELRSRGIFTGDGNLRFKNKFRTENLPLYLCQ  6564
             +FSFG SFG K  ESPN++PE E  L ELRS GIF  +GN +FKNKFR  N  +YLCQ
Lrg  6479  VEIFSFGASFGVKAGESPNSSPEVENGLAELRSVGIFD-EGNFQFKNKFRPANSTIYLCQ  6537

Lnm  6565  PADAERTDVSRLILMIANPEVGGDRAPTAPTAL------------AAPAPREAPAAPAPR  6612
           P D  R+ V+ +ILMIANP    D   TAP AL            +         +
Lrg  6538  PDDERRSAVADVILMIANP----DLDTTAPEALDDLPATSDSTSDSTSTSDSDTVSEPEP  6593

Lnm  6613  QAPAPGQAPAARPKPAAAPPAAVAAQATEVPADARRRERQLADHGW--NALHLASGDVEF  6670
            +    +AP    +PAA  PA  +A    VPA A       +D G+  N + L +  V F
Lrg  6594  VSEPEPKAPV---RPAAPEPAVASAPERPVPARAAAPRPAGSDSGYGRNPITLPTSAVRF  6650

Lnm  6671  DLITDSWAELDRPFVHARTARLHAGAAGRPVGQTLEGLDLLP---FSCVVATTSGRAAEA  6727
           DLITDSWAEL  P V  R ARL   AAG  +G   EGL  +P   F CVV T SGR AEA
Lrg  6651  DLITDSWAELATPAVTERMARLTEAAAGVELGA--EGLPRVPWLDFECVVPTPSGRTAEA  6708

Lnm  6728  ALCRAWPQ-QGVVVHNSLFPTWYFNHLDHGFTPVAARRAAGAD-DGVFRGDLDLGHLNGL  6785
            LCR+WP  +  VVHN LFPT   +  D+GF PV           G FR D+D   L  +
Lrg  6709  LLCRSWPGIRQAVVHNGLFPTLLMSLADNGFEPVELPGCRPVTRTGPFR-DVDPEALRRI  6767

Lnm  6786  LTEHAGRIAFLCVEVSNNAQGGAALSLHNLTGIRETADRHGLQLVLDATRVLDNAALIAA  6845
           L E  G ++ +C+E+S+NAQGG  +SL NL  +R  A   G+ LVLDATR L+NAA +
Lrg  6768  LAERPGGVSMICLELSDNAQGGYPISLANLREVRRIAVAAGIPLVLDATRALENAASVVE  6827

Lnm  6846  HEPGQTGRDPLDVARELLSLADSVTISLSKDFGVDTGGIVATDDPTVAHHLRERIALRGP  6905
           H+ GQ GR    V  +LL+ AD+VT+ LSKDFG+D GG+VAT  P +   LRE+++ RG
Lrg  6828  HDEGQQGRGIWKVTADLLATADAVTMGLSKDFGIDFGGLVATSRPGLVERLREQVSTRGH  6887

Lnm  6906  EAGRATRALAAAALDDLGWAATATGERVRRVADLRQALAAAGAPVAPGTGTHCVLLDTAR  6965
           +   A R   AAAL D G  A   G R   V  L   L+ AG PV      HCVLLDTA
Lrg  6888  QVNLAGRRQIAAALADSGGVAEQVGRRRAAVKMLWNLLSRAGLPVIGPAAGHCVLLDTAA  6947

Lnm  6966  LPALRGHEHPVPAFLAWLYLHTGIRAAAHLDDGPGTSSLVRLALPVGLGQRETAELTARL  7025
           +      EHPVP+ L W++ HTG+R   HL  G G + L+RLA+PVG G  +  E+  RL
Lrg  6948  MKQFADFEHPVPSCLDWIFEHTGVRGGPHLATGAGAAPLIRLAIPVGTGTPDIREIGKRL  7007

Lnm  7026  TALFGAPQQIPELLLAASDGPAALASYHPVEQVPDDIREAMAEGHTAENDNWAVLREHHP  7085
           T L+ +     EL+        A A+YH    VP+DI+ A+ EG  A++DN  VL +
Lrg  7008  TRLYRSGPAPVELIPVDPTAAPAQAAYHTAAVVPEDIKAALREGVRAKDDNLGVLTDFGA  7067

Lnm  7086  GVERVLLRLPAGDGGGDVEVFTAGAGPALLMMLPFNIGAGLFGPQFAALSERYRVIVVHH  7145
            VE  ++ +P    GG+VE F AG GP LL + PFNIGAG+F  QFA LS+RYRV+VVH
Lrg  7068  PVEHRIVGVPQ---GGEVEAFMAGHGPTLLFIHPFNIGAGVFRHQFAGLSDRYRVVVVHA  7124

Lnm  7146  PGVGDTTACEELGYEGIADLCLRALRRLGVQGPVHVAGASFGGITAQTFALRHPESTASL  7205
           PGVG T A  +L   G+A++   ALR LG  GPVH+AGASFGG+TAQT+AL HP+  ASL
```

```
Lrg  7125  PGVGRTNASADLTLHGLAEVHRAALRELGATGPVHLAGASFGGLTAQTYALEHPDEVASL  7184

Lnm  7206  TLIGSSYKLGNRAGEVNRLALVAKEDFDQVQSLSGSSRLDRERAR-FERLLLRCESMDPQ  7264
           TLI SSYK  NR GEVNRL +V +EDFD++ +  G+   D  R R  E +LLR ESMDPQ
Lrg  7185  TLICSSYKCANRVGEVNRLDVVLQEDFDRIAAADGAGAPDEHRRRQLEAVLLRSESMDPQ  7244

Lnm  7265  TGLRYLDVFATAPDLLGRLGDIAVPTLIVQGRHDTVIPQKTAHLLHGAIPDARYHEVPDA  7324
           TGLRYLDVFAT PDLL RL  IAVPTL+VQGR+DTVIPQKTAHLLHGAI D+RY E+ DA
Lrg  7245  TGLRYLDVFATEPDLLSRLPRIAVPTLVVQGRYDTVIPQKTAHLLHGAIADSRYAEIDDA  7304

Lnm  7325  GHFPSLSSSEEFNAVLSAFLEEH  7347
           GHFP+L+  + FN+VL+AFL EH
Lrg  7305  GHFPALTRPDVFNSVLTAFLSEH  7327
```

## LrgK *vs* LnmK

```
Lnm  1    MTITSSLDVRPEIKQA--------VTVRPGMCGPGSLFVGQLGDWTWETVSAQCDTDVFA  52
          M  T +  +RPE+ +A        VTV+PGMCG  SLFV Q+GDWTWETVS  C TD F
Lrg  2    MPRTENPLLRPEVWRAPDGSVGRLVTVKPGMCGHNSLFVSQVGDWTWETVSEVCGTDAFN  61

Lnm  53   ARDASGNPTYLAFYYFRVRGGRELHPGSLTFGDRLTVTSGCYDQGTESVLTLHRIDRAGS  112
          A D  G PTYL+FYYFRVR G +LHPG+LTFGDR+  TS  +  G+ESVLTLHRI R
Lrg  62   AVDDRGRPTYLSFYYFRVRSGGQLHPGALTFGDRIETTSRVFGFGSESVLTLHRIRRVPP  121

Lnm  113  DDAQRP---LDLHEFYERPRDGSLYVENFNRWVTRSAPGSNEDLVKSSPPGFRNDGLPQL  169
              +P   LD  EF+   +   LYV+NFNRWVTRS   SNE L+ S+P  FR+  LP+L
Lrg  122  GAELKPEQGLDPEEFFAARQPDCLYVQNFNRWVTRSRADSNEGLISSAPADFRHAHLPRL  181

Lnm  170  PAAYSPRAVYREARTAHTFRALDEPGFRLLPDTVEVEHPVDIVRDVNGVGLLYFASYFSM  229
          PAA+SPRA Y  AR   TF    +P +  L    EVE+P+DI RD+NGVGLLYFASYFS+
Lrg  182  PAAHSPRAAYGVARARRTFHDPADPEWEELVAADEVEYPIDITRDINGVGLLYFASYFSI  241

Lnm  230  VDKAALALWRRLGRSDRAFLRRVVVDQQMCYLGNADLDSVLTLGARVRVSTETPGEELVD  289
          +D A L +WR  GR+DR FL R V+D Q+CYLGNAD DSVL +  R       P EE  +
Lrg  242  IDGAMLKMWRNQGRADRRFLDRTVLDHQLCYLGNADADSVLRIRLRSWRRRGDPAEERWN  301

Lnm  290  VVISDRDSGRVIAVSTLHTQ  309
           V+ D  S R +AV TLH +
Lrg  302  AVVEDAASDRCLAVCTLHVR  321
```

## LrgL *vs* LnmL

```
Lnm  9   RAQTARLVVEVVTEILPGVDPQLIGGKRHLKDLGADSVDRVEIIAALLDRTRVDAPMSDF  68
         RA  ARLV E V  ILP V  + I G +HLKDLGADSVDRVEII AL+D   V  PM+ F
Lrg  9   RAGVARLVHETVAAILPQVPAERITGDKHLKDLGADSVDRVEIIMALIDTLGVREPMAGF  68

Lnm  69  SDLPDIDSLIDFL  81
         S LPDID+LIDFL
Lrg  69  STLPDIDALIDFL  81
```

## LrgM1 *vs* LnmM

```
Lnm  11   GIEALNVWCGLARLTAADLFAGRGLDPERLDNLMMSERSIGLPIEDPVTNAVNAARPLIE  70
          GIEALN+  GLA+L+ A+LF GRGLD ER+ +LMMS RSI LP EDP+TNAVNAA P++
Lrg  6    GIEALNIHAGLAQLSTAELFEGRGLDRERIGHLMMSARSIALPCEDPITNAVNAAAPIVG  65

Lnm  71   ALSPEERARIELVVTSTESGVDYSKSLTSYVHKYLGLNRHCRLIEVKQACFGATAAVQTA  130
           L+P ++ RIEL++TS+ESGVDYSKS+ SYVH+YLGL+R CR++EVKQAC+ AT A+Q A
Lrg  66   RLAPRDKERIELILTSSESGVDYSKSIASYVHEYLGLSRKCRVMEVKQACYAATGALQIA  125

Lnm  131  VGYLASGISPGAKALVIATDVAVVDEKAEYSEPAAGHGAAAMLLSDRPRVLAMDLGAFGN  190
           GYLASG+SPGAKALVI TDV++VD +A Y+EPA G GAAAMLL D PRVLA+DLGAFG
Lrg  126  AGYLASGVSPGAKALVIGTDVSMVDARAGYAEPATGTGAAAMLLGDDPRVLALDLGAFGA  185
```

```
Lnm  191  YSYETLDSARPSPRFDIADVDRSLFAYLDCLKNAYADYAARVTDVDFTRDFDHLVMHTPF  250
          +S+ET+DSARP P +DIADVD SLF YLDCL N++ADY  +V DVD +  FD L +HTPF
Lrg  186  HSHETMDSARPMPDYDIADVDGSLFTYLDCLSNSFADYCGKVADVDLSTTFDQLALHTPF  245

Lnm  251  AGLVKAGHRKMMREQGVTGP-RIDEDFARRVAPSLIYPGSVGNLCSGSVYLALASLLDSG  309
          AGLVKAGHRK+MRE     P  ++ DFA RV+PSL+YP  VGNLCSGSVYLALASL+D+
Lrg  246  AGLVKAGHRKLMREHARAAPDAVEADFAARVSPSLVYPSQVGNLCSGSVYLALASLIDNA  305

Lnm  310  VVTAPSRVGLFSYGSGCSSEFFSGIVDEQSAATVAEQGIGKRLEARARITFDEYLAVLEH  369
               +RVGLFSYGSGC+SEFFSG+VDE S A +AE  I  RL AR  + F EY  +L
Lrg  306  PYRGTARVGLFSYGSGCASEFFSGLVDEGSRAALAELDIAGRLNARVPLDFAEYTELLAE  365

Lnm  370  NLECLVPVENRTVDPAEWEPLLDRVGDRPEILTFTGVKDYHRQYAW  415
          N  CLVPVE+R ++   +   LD    R  +L + G + YHR Y W
Lrg  366  NSRCLVPVEDRKIEVERYRRFLDARPGREPLLAYRGTEGYHRTYEW  411
```

LrgN *vs* LnmN

```
Lnm  14   DGAAAGLRLFCFAHAGGGSSFFHPWRRALGPGVDVRPVVLPGRERRARETSHTRMGPLVE  73
          +G    +RLFC  HAGGG +FFHPWR AL PGV+VRPVVLPGRE R RE  +  M   +
Lrg  2    NGTEPEIRLFCLPHAGGGGAFFHPWRAALAPGVEVRPVVLPGRESRIRELPYVTMEQAIG  61

Lnm  74   GLVTELAPQLDLPYVLFGHSLGSIVAYETARALLERGSRPPLALLVSGRRGPFVPDHRRP  133
           L   LAPQLD PY LFGHS+G+ V YE AR  L  G   P+ L VS RR P +P  R
Lrg  62   PLAELLAPQLDRPYALFGHSMGAAVGYELARRFLALGLPAPVRLFVSARRAPHLPARRAS  121

Lnm  134  VHNLPEDEFLAEVSRLGGTPSEVLRQRDLLRHFLPPLRADHEVNETYRPVPPPGPALTCP  193
             L +  FLAEVSRL GTPS+VL Q +L+R FLP LRAD E+N+TY P+  P P L CP
Lrg  122  YAGLDDAAFLAEVSRLNGTPSDVLEQPELVRLFLPTLRADFELNDTYTPL--PAPRLDCP  179

Lnm  194  VFAFTGDADPLADPHAVARWREVTSGDFRLRVFPGDHFYLKGAPDDLMSALRAAM  248
          + AF G  DP AD   +  W +VT+G FR R F GDHFYLK    DL+  +RA +
Lrg  180  ISAFVGRDDPEADARELKAWEQVTAGAFRFREFDGDHFYLKDRAADLLDEIRADL  234
```

LrgX *vs* LnmX

```
Lnm  3    DTLLELPDDFSRVLAIVAHPDDIEFGAGPAVAQWTAQGREVAYLLVTRGEAGISDLEPAQ  62
          + L  LPDD+SR LA+VAHPDDIEFG   AVA WTA G+ V+YLLVT+G+AGI  L PA+
Lrg  4    EELQTLPDDWSRALAVVAHPDDIEFGTSSAVAAWTAAGKSVSYLLVTKGQAGIDGLAPAE  63

Lnm  63   CGPVREAEQRKAAAELGVHEVDFLDHYNDGTIEYGPGLRRDLARAVRRHRPELIVTFNHH  122
             VREAEQR +A  +GV EV+FLDH  DG IEYG GLRRD+A A+RRHRPE ++ FN
Lrg  64   SAVVREAEQRASAKIVGVGEVEFLDH-RDGEIEYGLGLRRDIAAAIRRHRPEFVLGFNGR  122

Lnm  123  DTWASGAWNTPDHRAVGLAALDAVADAANRWIFPELLDEGLEPWRAGK-VAIAGSPHATH  181
          +T ++G WNTPDHR    A LDAV DA NRWIF +L   GLEPW   K VA+A SP  TH
Lrg  123  ETTSTGKWNTPDHRHTAHALLDAVGDAGNRWIFEDL---GLEPWGGVKYVAMANSPQPTH  179

Lnm  182  AVAVDDDSRDRAVRSLAAHDRYLGSLSDDPPQERARFILGHLLAATAPRFGGRDGVAFQI  241
          AV V  D  +  + SL AH  YLG L  +PP    R  +         RFGGR  VA +I
Lrg  180  AVDV-TDHLEAGIASLEAHSAYLGGL--NPPVTSVREPMTAFAELVGERFGGRPAVALEI  236

Lnm  242  V  242
          +
Lrg  237  I  237
```

LrgZ *vs* LnmZ’

```
Lnm  1    MTQMRIQATFVVDVWDGTDD-EPVDGGPVTGRVELTKTYTEGDVKGSATGHMVTTQ-GPG  58
          MTQ+   + +    W+ ++  EP DG P+  R ++ + +  GDV+G+    ++  +
Lrg  1    MTQI-ANSAWETSSWEESNYFEPADGPPLI-RADVKRVF-RGDVEGTGEAVLLCCRPDEK  57

Lnm  59   GAAYVAQERVTGIMGGRTGTFVLEHRATQVPGTD-PVTWAGIVPGSGTGELAGVSGEGSL  117
           A YV+ E +   + GR+GTFV++H A+   G D P T   +VP SGTG L G++G  +
Lrg  58   SAGYVSTEHIVATLAGRSGTFVVQHGASM--GDDEPQTLGFVVPNSGTGGLTGLTGTCAF  115
```

```
Lnm  118  GH  119
          GH
Lrg  116  GH  117
```

LrgW1 *vs* LnmW

```
Lnm  5    TFSYLTDQLRSHAALHPDRTALVIDGCPDLLYGEWDRRSEALARGLLAAGTSRGTRIGIF  64
          T   +TD LR  A  HPDRTAL IDG   L YG+W RR +  A GLLAAG ++G RIG+
Lrg  6    TVVRVTDLLRRRAEHHPDRTALDIDGTDALNYGDWQRRVDRTAHGLLAAGVTKGRRIGLL  65

Lnm  65   FGGMDWAGYAVAYLGALKAGATVLHLPLALPADELERRALQCELAGIVHGRTAPPTTSAV  124
          +GGMDW  YAVAYL  L  GAT +HL   L   E+ERR  +C    ++HG +  P  S
Lrg  66   YGGMDWTDYAVAYLAVLSVGATAVHLSDRLGEPEIERRLTECRATAVIHGGSLRPPASFT  125

Lnm  125  AWTGTLDELSAPGETPVDLVHSPADAAEIVYSSGTTGLARGVVVSHQNLATAGGPPSVMA  184
           W+  + EL +  ETPVD+  +P D A+++Y+SGTTGLA+     H NL    GP  ++
Lrg  126  GWSAEVAELDSGDETPVDVPLAPEDIADVLYTSGTTGLAKAFTNPHGNLTFGRGPEGLLQ  185

Lnm  185  HDEPTPMVASVNLGITASATTVSMVLNATPTTLVLAPPGDADRLCALIEHHAASTVMMTP  244
           + PTP++A + LG T+SATTV+++   +P+ LVLA   D +R+  LI  H   +VM+TP
Lrg  186  FENPTPLLAPMPLGTTSSATTVAIIAVTSPSALVLAAVDDVERMAELISRHRIGSVMITP  245

Lnm  245  NLAVQMTRDGALGRYDLTSVTTVATASAFLHPPLARALLAAMPRARVIGAYSASQAKPAV  304
           +A++M       R+D+  V  VA ASA L P L+R LL   P A +  AYS S+A PAV
Lrg  246  WIAMRMLAARIGERHDVGCVERVAIASAPLAPALSRGLLKLFPAAELNTAYSQSEAVPAV  305

Lnm  305  TIGTFDPARPMSAGRPAPGTHVLITDEHGAELPAHRVGRIWLRADGAPPRNRLDAGPEAT  364
           + TFDPARP + GR A GT V I D  GAELP   VG I LR+  AP R  LDA  +A
Lrg  306  VVNTFDPARPSTLGRAARGTEVRIADALGAELPLGEVGEIQLRS-AAPGRRYLDARRDAE  364

Lnm  365  GVPEGGWCDTGDLGHVDDEGELYLFDRETDAVPTPAGLVSSLRVESVLLEH--EAVADAA  422
           V   GW  TGDLGH+D+EG L+LFDR  D +      VSS+ VE+ L EH     A
Lrg  365  -VRIDGWIRTGDLGHLDEEGWLHLFDRGDDVLDGGGVRVSSVAVEAALYEHPAVREAAVV  423

Lnm  423  VVAAGPAGVAAAIVPAAGATHDPKLLAATLAAHAKDSLAPHEIPERVLVVDELPRNDLGK  482
             +GPA V     PAA         A  L A   + L PH++P  V   + LPR   GK
Lrg  424  AAGSGPAAVVVLEDPAA---------AGELPAFLAERLEPHQLPVLVEARESLPRGITGK  474

Lnm  483  VVKRLIRDRL  492
          V+KR++R  L
Lrg  475  VLKRILRQEL  484
```

LrgW2 *vs* LnmW

```
Lnm  11   DQLRSHAALHPDRTALVIDGCPDLLYGEWDRRSEALARGLLAAGTSRGTRIGIFFGGMDW  70
          D LR  A +HPD+  + I+G   L YGEW +R+ A+ARGLL  GTSRG RI + FGG+DW
Lrg  14   DLLRLRAEIHPDQIVVNINGERTLSYGEWYKRANAVARGLLDRGTSRGERIALLFGGLDW  73

Lnm  71   AGYAVAYLGALKAGATVLHLPLALPADELERRALQCELAGIVHGRTAPPTTSAVAWTGTL  130
            YA+AYLG + AGAT +H+   + A E  RR  QC++ G+V G           W  T+
Lrg  74   IDYAIAYLGIVNAGATAVHMHRDISAAEFNRRIAQCQVTGLVRGHDVVVPEGFEGWAATV  133

Lnm  131  DELSAPGETPVDLVHSPADAAEIVYSSGTTGLARGVVVSHQNLATAGGPPSVMAHDEPTP  190
          DE+ +   TPV +   P D A+I+Y+SGTTG A+ +   H NL    GP       +P P
Lrg  134  DEVDSGDPTPVKVELRPDDLADILYTSGTTGTAKAIATPHGNLTFGRGPEGFKQLGKPKP  193

Lnm  191  MVASVNLGITASATTVSMVLNATPTTLVLAPPGDADRLCALIEHHAASTVMMTPNLAVQM  250
          ++A + LG T+SATT+++ L   P TLVL P  D DR+  LIE +   +VM TP + +QM
Lrg  194  LLAPIPLGTTSSATTMAIALT-NPATLVLCPVDDVDRMGELIEQYQIVSVMFTPWIGIQM  252

Lnm  251  TRDGALGRYDLTSVTTVATASAFLHPPLARALLAAMPRARVIGAYSASQAKPAVTIGTFD  310
                 +DL+ V T+ATASA L P  A AL+  MP A+V   Y+A +A PAV   TFD
Lrg  253  VAGKIHETHDLSCVETLATASAPLPPATASALMRMMPNAKVTSVYAAREAVPAVIAATFD  312

Lnm  311  PARPMSAGRPAPGTHVLITDEHGAELPAHRVGRIWLRADGAPPRNRLDAGPEATGVPEGG  370
```

```
           +RP   GRP  G+ +L+ D  G  +    +G IWLR  GAP R  L+ G E
Lrg  313  VSRPFCVGRPGEGSELLVADADGNPVATGEIGEIWLRC-GAPKRLFLE-GAEREEQLTDD  370

Lnm  371  WCDTGDLGHVDDEGELYLFDRETDAVPTPAGLVSSLRVESVLLEHEAVADAAVVAAGPAG  430
          W  T DLG++D EGEL+LFDR  DAV     LVS++  E+ L E   V  AAV+    AG
Lrg  371  WTRTRDLGYLDAEGELHLFDRAADAVTVDGELVSTIHTEAALYECPGVEQAAVLGVPAAG  430

Lnm  431  VAAAIVPAAGATHDPKLLAATLAAHAKDSLAPHEIPERVLVVDELPRNDLGKVVKRLIRD  490
               + A     D   L A  AA A + L PH+IP R  +VD LPR  +GKV+K  +R
Lrg  431  TDRVELAAVLVLADDDGLPAVRAALA-ERLEPHQIPTRFQLVDALPRGVMGKVLKHQLRR  489

Lnm  491  RLTAS  495
          +L  S
Lrg  490  QLAGS  494
```

LrgC3 *vs* LnmA

```
Lnm  22   HPKFAELRETDPLARVRLPYGGEGWMV-TRYDDVRAANSDPRFSRAQ---------IGED  71
          +P + +LR+  P         G+G +V +R+ D  AA  D    +           + +D
Lrg  34   YPLYHQLRDAAPAL-----LTGDGTLVLSRHADCNAALRDRSLGKGDEWLKLQLKDVSKD  88

Lnm  72   TPRTTPLARRSDTILSLDPPEHTRLRRLLSKAFTARRMGAMQSWLEELFAGLLDGVE-RT  130
            R      +   IL+ +PP+HTRLRR++S AFT R + A++  +     GLLD +  R
Lrg  89   DLRGVMELMQRSMILT-NPPDHTRLRRIVSSAFTGRHVEALRDGVTRRVDGLLDRLAARP  147

Lnm  131  GHPADIVRDLAQPFTIAVICRLLGVPYEDRGRFQ---HWSEVIMSTTAYSKE--EAVSAD  185
          G  AD++ +LA P  ++ I  LLG+P  DR       H   ++M   +   E    V+A
Lrg  148  G--ADLMTELAMPLPVSTIGDLLGIPEADRAELVPVIHELGLLMEPASGPAEINRGVAAQ  205

Lnm  186  ASIRAYLADLVSARRAAPHDDLLGVLVSARDDDDRLTEDELITFGVTLLVAGHETSAHQL  245
          A + +YL  L++ +R  P DDLL  L S     D L E E+I   + L  AG+  +A+ +
Lrg  206  AHLASYLGGLIAEKRQRPQDDLLSRLAST--SADALDETEVIATALLLFGAGNTPTANLI  263

Lnm  246  GNMVYALLTHEDQLSLLREQPELLPRAVEELLRFVPLGNGVGNARIALEDVELSGGTVRA  305
          GN + AL+   +Q   L E P LLP AVEE+LRF        +    LE    +G  +R
Lrg  264  GNGLDALVRFPEQRQRLTEDPGLLPSAVEEMLRFD--SPSQFDVFTVLEPHSFAGTELRP  321

Lnm  306  GEGVVAAAVNANRDPRAFDDPDRLDITREKNPHLAFGHGAHYCLGAQLARMELRVAIGGL  365
          G+GV+     AN DP  FDDPD  D+ R++N HL+F  G H+CLGA LAR++  V  G L
Lrg  322  GQGVMMMLGAANHDPERFDDPDAFDVGRKENGHLSFAAGIHHCLGAHLARLQAEVVFGRL  381

Lnm  366  LERFPGLRLAVPA  378
          L RFP L  A PA
Lrg  382  LARFPKLEPAAPA  394
```

**Alignments of LrgS, LrgH and LrgY with proteins from syringomycin BGC**

LrgS *vs* SyrB1

```
Syr  20   GAFLHEIFSDRARQFPERTAVSDAARTLSYAQLDALSTKLAARLRDEGVTYGTRVGMYLP  79
          G  LH++F  +A + P+R AVS A R L+Y +L+A +  +AARLR  G      +G+ +
Lrg  4    GKTLHQLFEVQAARTPDRVAVSGADRALTYRELNAEADAVAARLRQAGAGPDRLIGLCVD  63

Syr  80   RSVDLVTSLLGILKAGGTYVPVDPQYPGKRVEHIVRDSELSLIIGDAANLPKISSLR--V  137
          RS DLV  LLGILKAG  YVPVDP YP +RV  ++ DS++S ++  +    +++     V
Lrg  64   RSADLVVGLLGILKAGAAYVPVDPAYPAERVAFLLDDSQVSAVVSVSRVAERLADCAAPV  123

Syr  138  LALDELLSAPALQPAAQDTRIDPNNSTAYIIYTSGSTGEPKGVQVSHGNVSRLLESTQRA  197
          + LD     PA   A +++R    +  AY+IYTSGSTG PKGV V H N  RL E T
Lrg  124  VWLDRDTEPPAAPAAVEESR---ESDLAYVIYTSGSTGVPKGVLVEHRNAVRLFEQTAEL  180

Syr  198  YGFNAQDVWSMFHSIGFDFSVWEIWGALAHGGQVAVVPYDISRSPAALRQWLADQRITVL  257
           G+ A DVW++FHSI FDFSVWE+WGAL HGG++ V   +  RSP  L + LAD+ +TVL
```

```
Lrg  181  VGYRADDVWTLFHSISFDFSVWELWGALLHGGRLVVAGTETVRSPELLHKLLADEGVTVL  240

Syr  258  SQTPSAFRGLDEADRGNTAPL-ALRYVVLGGEALPASVLRPWVERHGDQKPALINMYGIT  316
          +QTPSAFR L  A   +TA L ALR VV GGE L   +L PW  R+GD++PAL+NMYGIT
Lrg  241  NQTPSAFRRLVGA---STAKLPALRLVVFGGERLDVKLLEPWFARYGDERPALVNMYGIT  297

Syr  317  EATVHTTFKRVLAQDLETAAMVSLGKPLDGWRLHLLDANQAPVAAGTTGELYIEGAGVAQ  376
          E TVH T + +   DL+   +  +G+PL G  LHLLD +  PVA GT GELY+ G GVA+
Lrg  298  ETTVHVTARPITRADLDEPGVSPIGRPLPGVTLHLLDEDGGPVADGTPGELYVGGTGVAR  357

Syr  377  GYLNREALNVERFVEL---PGAVRAYRTGDLMTLESNGEYRYAGRCDEQLKISGFRIEPG  433
          GY  R  L  ERF  +     AVR YR+GD     ++GEY Y GR D+Q+KI GFRIEPG
Lrg  358  GYHRRPELTAERFRTVGTGADAVRLYRSGDRAVRTADGEYLYVGRADDQIKIRGFRIEPG  417

Syr  434  EIEASLQTSPSVAAAHVGVHDYGDGDLRLVAYVVPGQGVDAWTEQAR---SEVAALMAEN  490
          EIEA L   P +A+A V   D+G+GD+RL AY+VP  G +   E+     +EV+A  A
Lrg  418  EIEALLADDPRLASAIVVPQDHGEGDIRLTAYLVPRPGAEIDDEELGRLVAEVSARAAGT  477

Syr  491  LPGYMRPSVYVPLAELPVTHHGKIDKQQLPSPAAGTALSG-AADVKGLSEQEHFVLKVWS  549
          LP +MRPS Y  + E+P T  GK+D+  LP+     A +G A     L+  +  V  + +
Lrg  478  LPEHMRPSAYRLITEVPTTAQGKVDRSALPALPHRQAPAGPAGGGVELTPTQRQVDAIVT  537

Syr  550  EDLGLKNIGVNDDFFDSGGTSLALIRSLSKLKTHYKINLDPGILADGATAKVLA  603
          E L    IG++DD F+ G TSLA +R ++ +   +K++L  G   D AT + L+
Lrg  538  EVLARPGIGLDDDLFEHGATSLAFMRVIASVNRRWKLSLT-GAELDAATVRQLS  590
```

## LrgH *vs* SyrB2

```
Syr  1    MSKKFALTAEQRASFEKNGFIGPFDAYSPEEMKETWKRTRLRLLDRSAAAYQDLDAISGG  60
          ++  F LT E+RASF++ G+ GPF  Y  EEM+  W+  RLRL+DRS A YQD  A SG
Lrg  4    VTGGFTLTPEERASFQERGYFGPFKVYEIEEMQRRWRIERLRLMDRSNAVYQDEAAQSGN  63

Syr  61   TNIANYDRHLDDDFLASHICRPEICDRVESILGPNVLCWRTEFFPKYPGDEGTDWHQADT  120
          TNI+NYDRHLD +FLA HICRPEI DRV S+LGP+VLCWR+EFFPKYPGDEGTDWHQADT
Lrg  64   TNISNYDRHLDSEFLADHICRPEIVDRVASVLGPDVLCWRSEFFPKYPGDEGTDWHQADT  123

Syr  121  FANASGKPQIIWP-ENEEFGGTITVWTAFTDANIANGCLQFIPGTQNSMNYDETKRMTYE  179
          FANASG PQI+WP E+++FGGTITVWTAFT+AN  NGCLQFIPG+   MNYDETK+M Y
Lrg  124  FANASGVPQILWPDEHKDFGGTITVWTAFTEANEDNGCLQFIPGSHTRMNYDETKKMHYT  183

Syr  180  PDANNSVVKDGVRRGFFGYDYRQLQIDENWKPDEASAVPMQMKAGQFIIFWSTLMHASYP  239
          PD+ N V K GVRRGFFGYDYR+LQ+D ++KPDE+ AV M M+ G+ I+FWSTLMHAS+P
Lrg  184  PDSINQVDKGGVRRGFFGYDYRELQVDADFKPDESQAVSMVMRPGEAIMFWSTLMHASWP  243

Syr  240  HSGESQEMRMGFASRYVPSFVHVYPDSDHIEEYGGRISLEKYGAVQVIGDETPEYNRLVT  299
          HSG++ EMR+GFA RYVP+ V VYPD++ IEEYGG +SLE+YGAV V G+   ++NR+VT
Lrg  244  HSGKTDEMRLGFAGRYVPTSVRVYPDTEQIEEYGGTVSLERYGAVLVGGENRYDHNRMVT  303

Syr  300  HTTRGKKF  307
           TT G  F
Lrg  304  RTTLGHPF  311
```

## LrgY *vs* SyrC

```
Syr  77   DHAPLLVHASANRERPPVVLALPCGIPFDLCRDWFDALSERFFVVTWETRGLFGACEAFD  136
          D A L V+A+ +   P V++A  CG+P +LC  W + L   + +VTWETRGLF     FD
Lrg  121  DGARLPVYAAGDPAAPAVLIASACGMPAELCSRWLELLGGEYRIVTWETRGLFTDEPDFD  180

Syr  137  QIAVDTDAQVADMISVMNHFGLSTAHLMGICAGAVIALSAAAAHAERVNSLSLWHGDYNL  196
          ++  DT+AQ  D+ +VM+H G+ TAHL+G C G+V+AL+AA    ER++SLSLWHG Y L
Lrg  181  KLEWDTEAQAGDVFAVMDHLGIRTAHLLGFCGGSVVALAAARRSPERIDSLSLWHGAYEL  240

Syr  197  GDNDLRAAHQQNFEWLMESAAQDRDEAADLQAMFLDQATLATTPESIAHVVLYPYVNARV  256
          G    +  H +N + LM  AA+DRD AA + A+FL    L  TP  +AH+VLYP+  + +
Lrg  241  GPESPKLDHHRNIQALMAMAAEDRDTAAAVHAVFL-STMLGGTPPDLAHLVLYPFATSEL  299
```

```
Syr  257  VLSLCRLNDALNKTELAPRLTRITAPTLVVAGDADSTTHIGGSAHIAASIKDATLHVERN  316
              CRLN A+   ++ P L  +  PTLVV  + D+T +  GSA +A  + +ATL VE
Lrg  300  FYRYCRLNGAITDIDVNPLLVGVDHPTLVVTSEDDTTANPLGSAEVARRLPNATLSVEPT  359

Syr  317  GSHLAFFASSQQSKQTAFSFLEE  339
          G H++ F    +  + A  FL E
Lrg  360  GDHISLFKLGGRLGELALGFLRE  382
```

**Figure S1. Confirmation of *Streptomyces argillaceus* mutants.** Generation of *Streptomyces argillaceus* mutants in *lrg* and flanking genes. Each panel includes a scheme representing construction of a mutant and its confirmation by Southern hybridization or PCR. (A) *S. argillaceus* ΔlrgG; (B) *S. argillaceus* ΔlrgR3; (C) *S. argillaceus* ΔlrgR4; (D) *S. argillaceus* Δorf18; (E) *S. argillaceus* ΔlrgT1; (F) *S. argillaceus* ΔlrgP1; (G) *S. argillaceus* ΔlrgC1; (H) *S. argillaceus* Δorf52; (I) *S. argillaceus* ΔlrgC3; (J) *S. argillaceus* ΔlrgO.

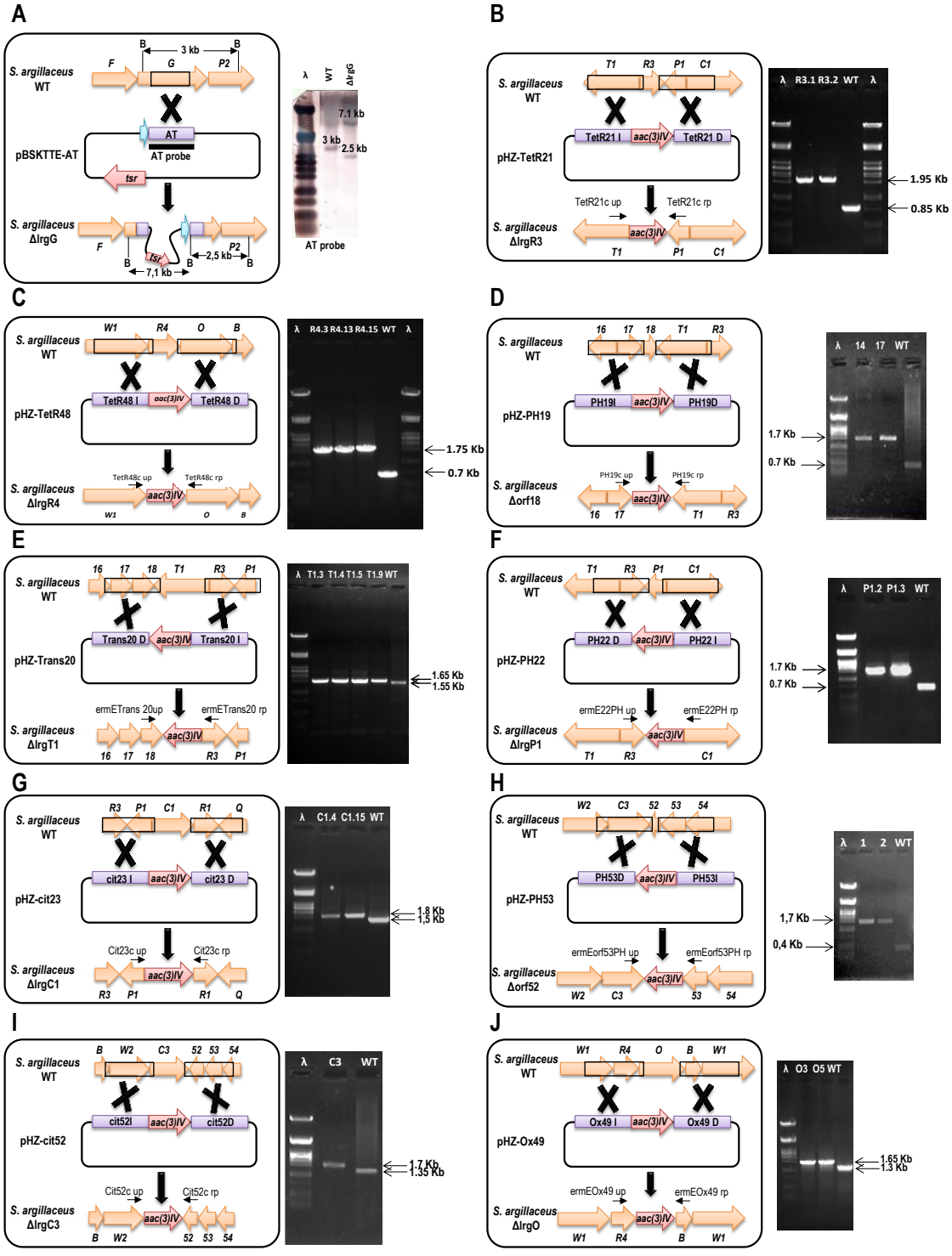

**Figure S2. UPLC analyses of *S. argillaceus* WT-R2 cultivated in different media.** Chromatograms (maxplot) of ethyl acetate extracts from *S. argillaceus* WT (black line) and *S. argillaceus* WT-R2 (red line) cultivated in (A) LF1 and (B) R5A media.

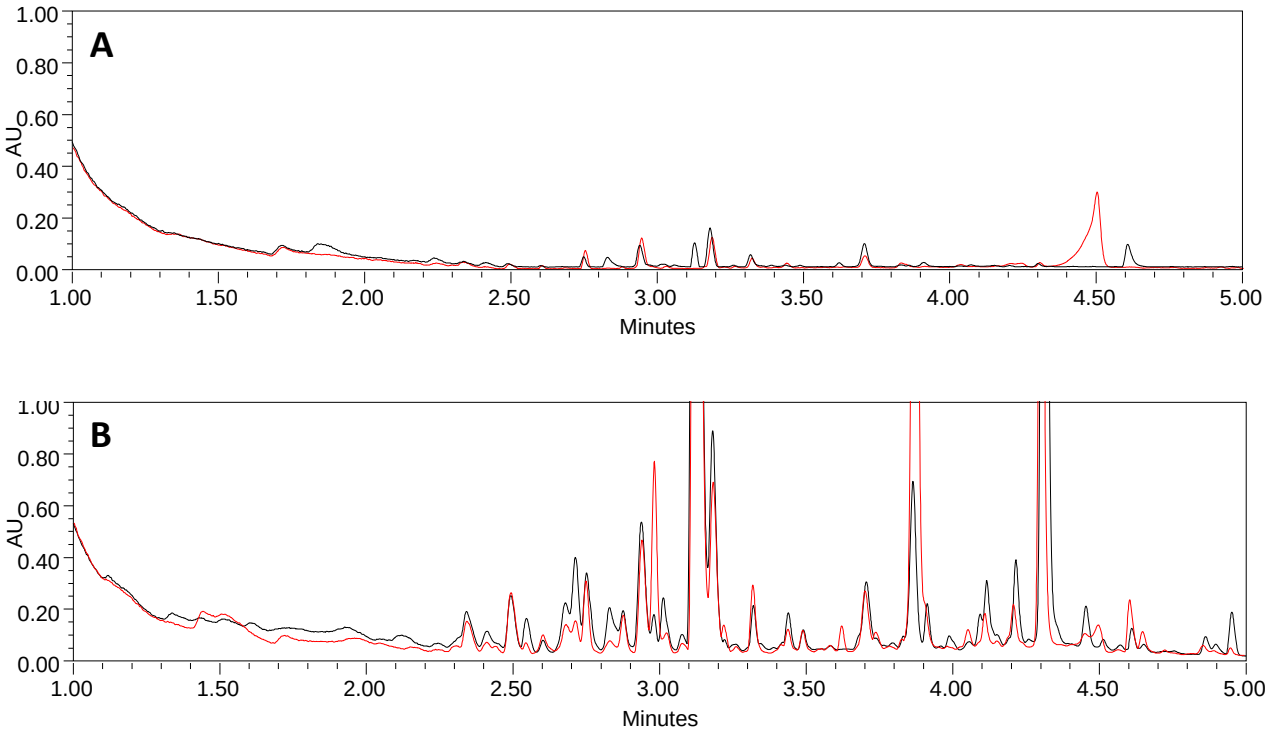

**Figure S3. UPLC analyses of extracts from *S. argillaceus* mutants.** A-G: Metabolite profiles of mutants (in red) in comparison with the wild type strain (in black): (A) *S. argillaceus* ΔorfI8-R2; (B) *S. argillaceus* ΔlrgT1-R2; (C) *S. argillaceus* ΔlrgP1-R2; (D) *S. argillaceus* ΔlrgR3-R2; (E) *S. argillaceus* ΔlrgC1-R2; ; (F) *S. argillaceus* ΔorfI52-R2; (G) *S. argillaceus* ΔlrgC3-R2. LRG intermediates accumulated by mutants are indicated by dots. H-J: Metabolite profiles of complemented mutants (in red) in comparison to the corresponding mutant (in black): (H) *S. argillaceus* ΔlrgC1-R2-C1; (I) *S. argillaceus* ΔlrgC3-R2-C3; and (J) *S. argillaceus* ΔlrgO-R2-O. Peaks corresponding to largimycin A2 (LRG A2) are indicated.

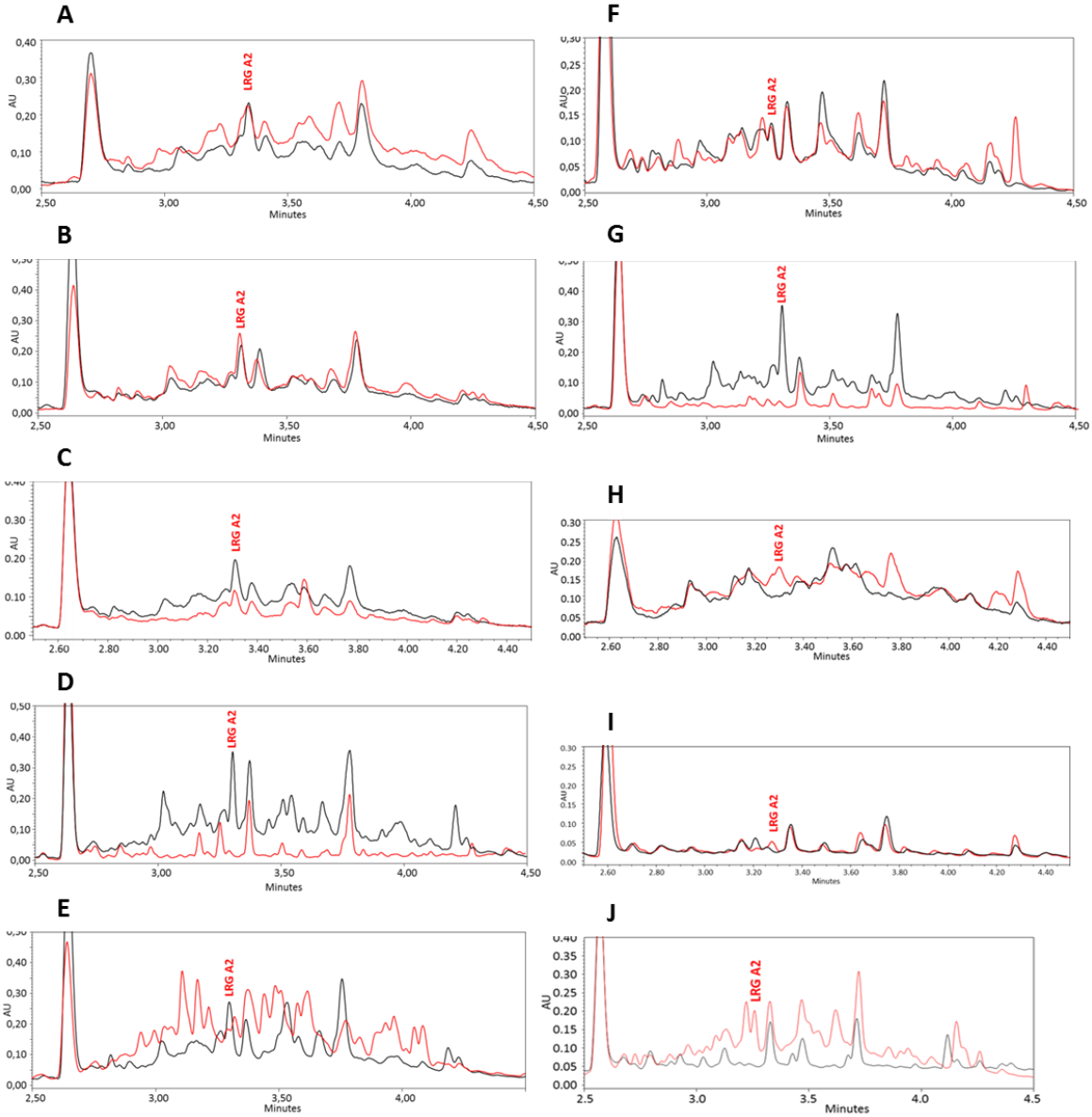

**Figure S4. Spectroscopic data of LRG A1 (1)**: (A) UV-DAD spectrum; (B) HRMS spectrum; (C) 1H NMR spectrum in DMSO-d6, 500 MHz; (D) COSY spectrum; (E) NOESY spectrum; (F) Edited HSQC spectrum; (G) HMBC spectrum.

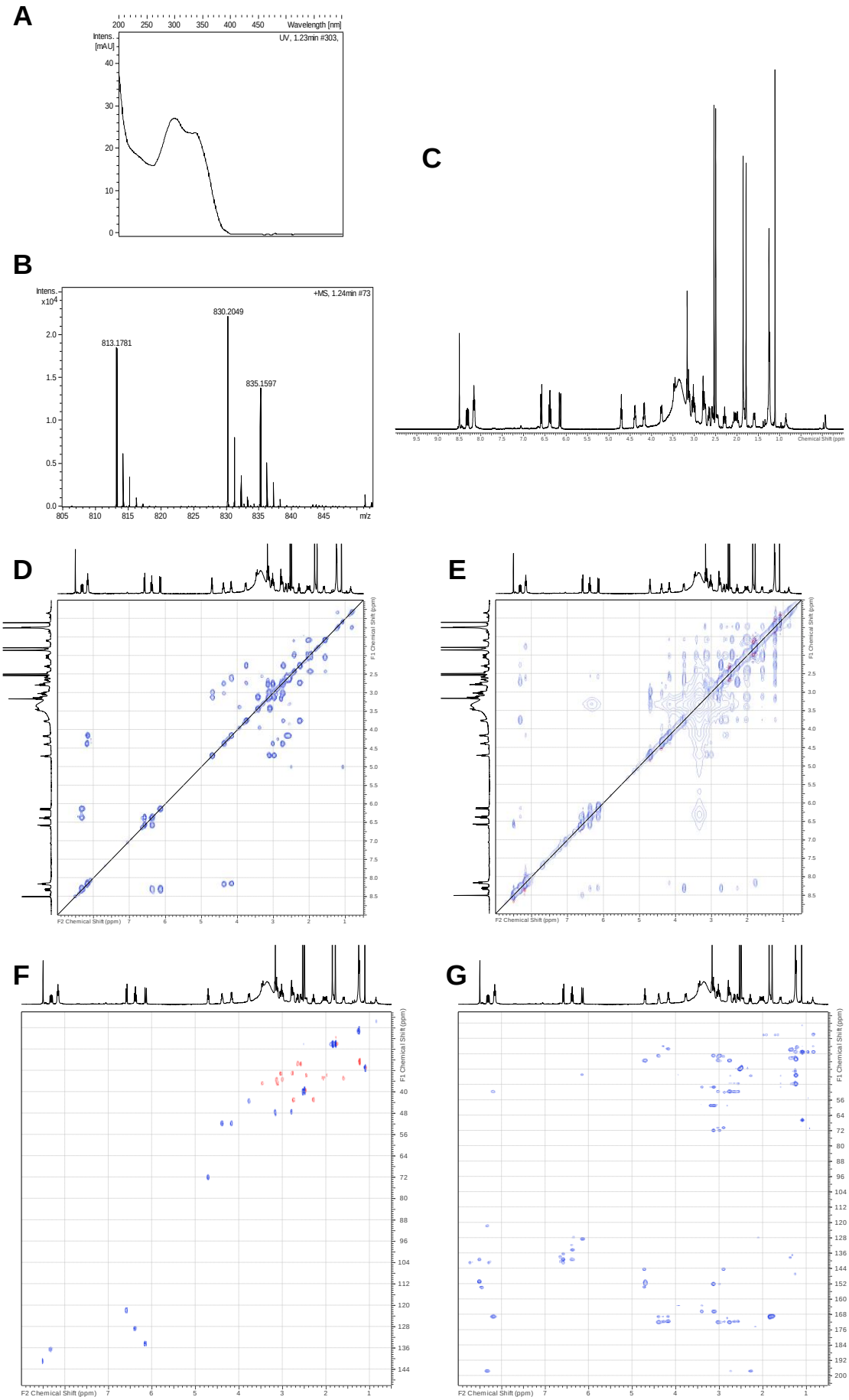

**Figure S5. Spectroscopic data of LRG A2 (2)**: (A) UV-DAD spectrum; (B) HRMS spectrum; (C) 1H NMR spectrum in DMSO-$d_6$, 500 MHz; (D) COSY spectrum; (E) NOESY spectrum; (F) Edited HSQC spectrum; (G) HMBC spectrum.

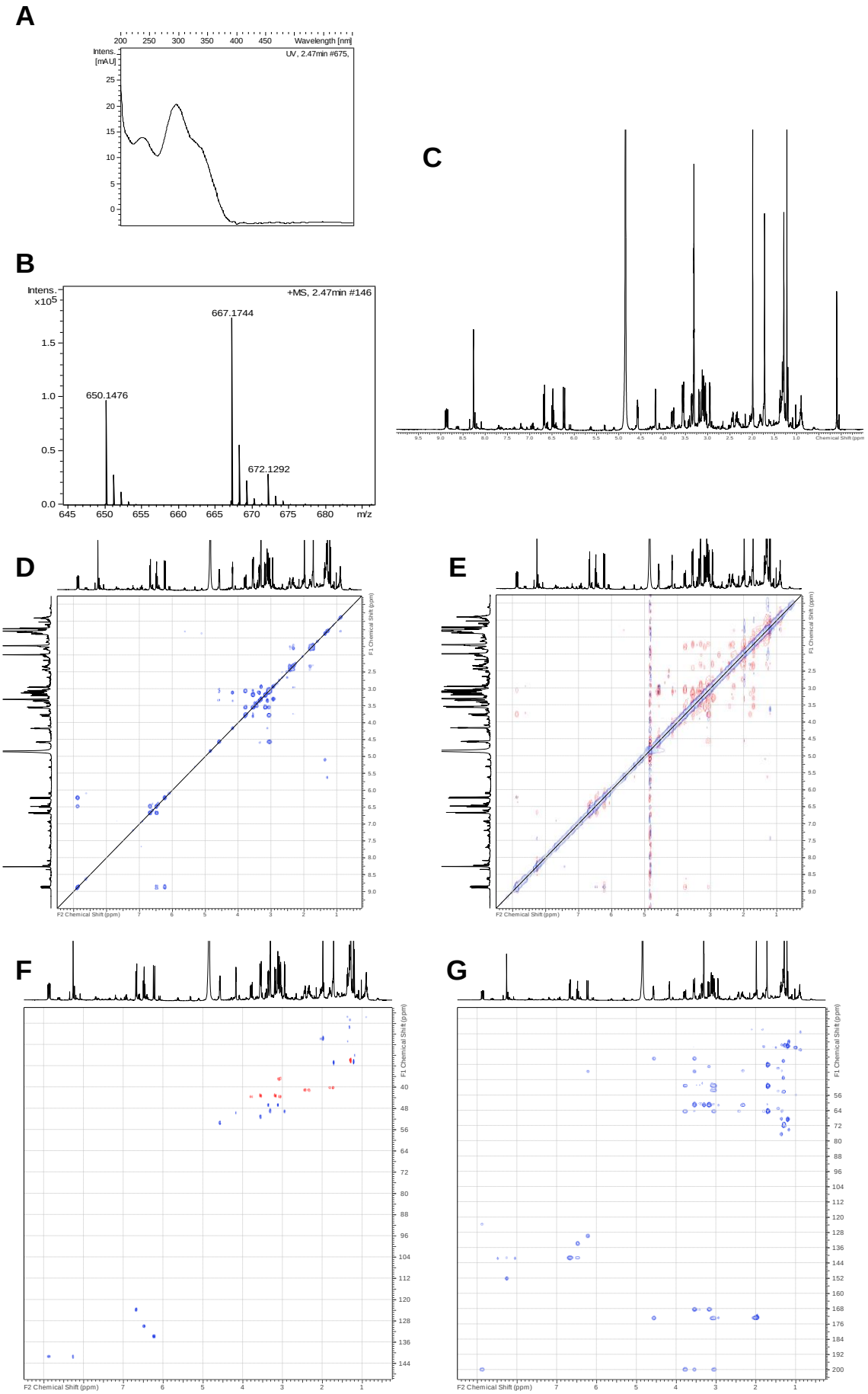

**Figure S6. Spectroscopic data of LRG A3 (3):** (A) UV-DAD spectrum; (B) HRMS spectrum. **NMR spectra of LRG A1 (1) in $CD_3OD$**: (C) 1H NMR spectrum, 500 MHz; (D) COSY spectrum; (E) NOESY spectrum; (F) Edited HSQC spectrum; (G) HMBC spectrum.

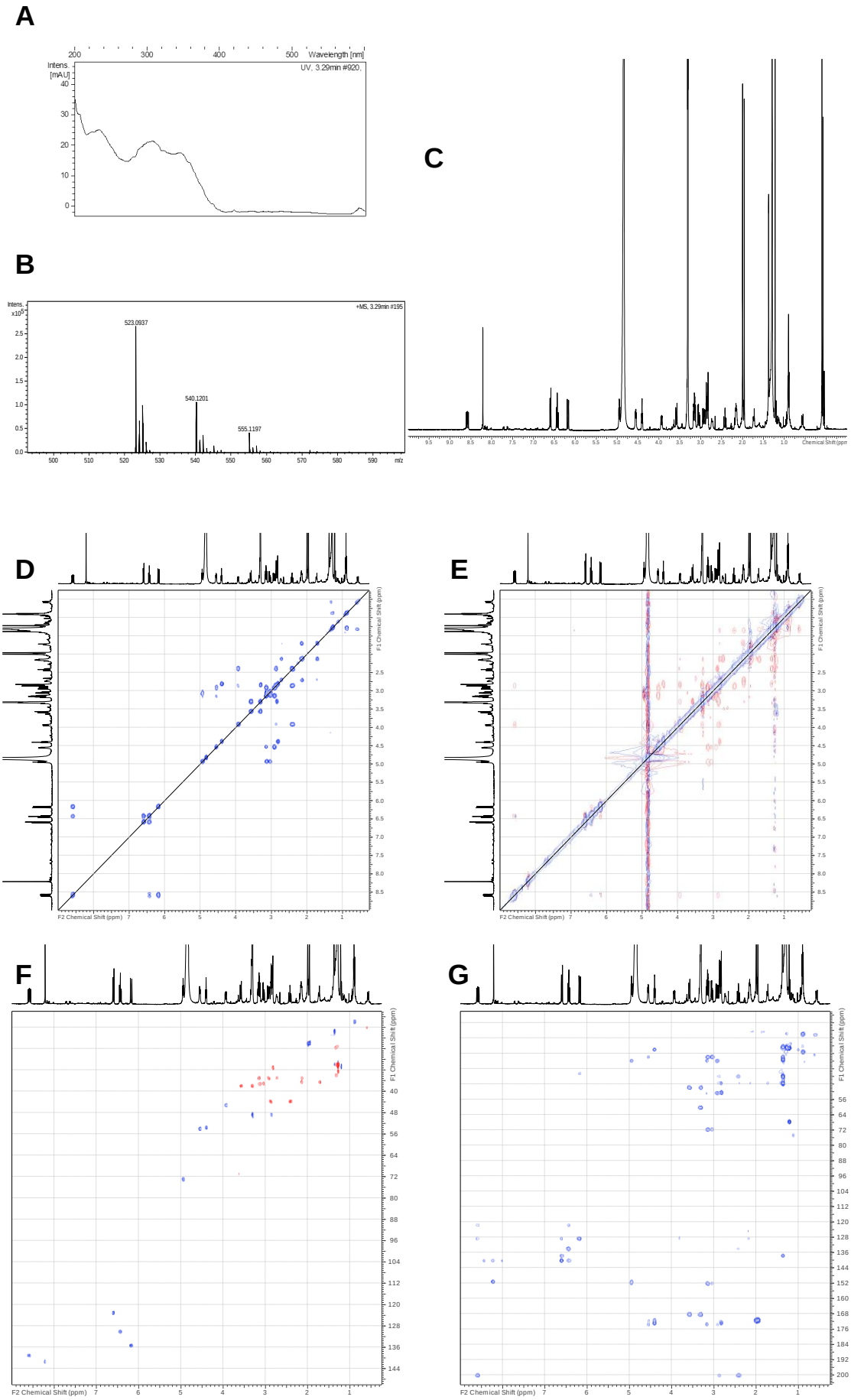

**Figure S7. Spectroscopic data of LRG A4 (4)**: (A) UV-DAD spectrum; (B) HRMS spectrum; (C) 1H NMR spectrum in $CD_3OD$, 500 MHz; (D) COSY spectrum; (E) NOESY spectrum; (F) Edited HSQC spectrum; (G) HMBC spectrum.

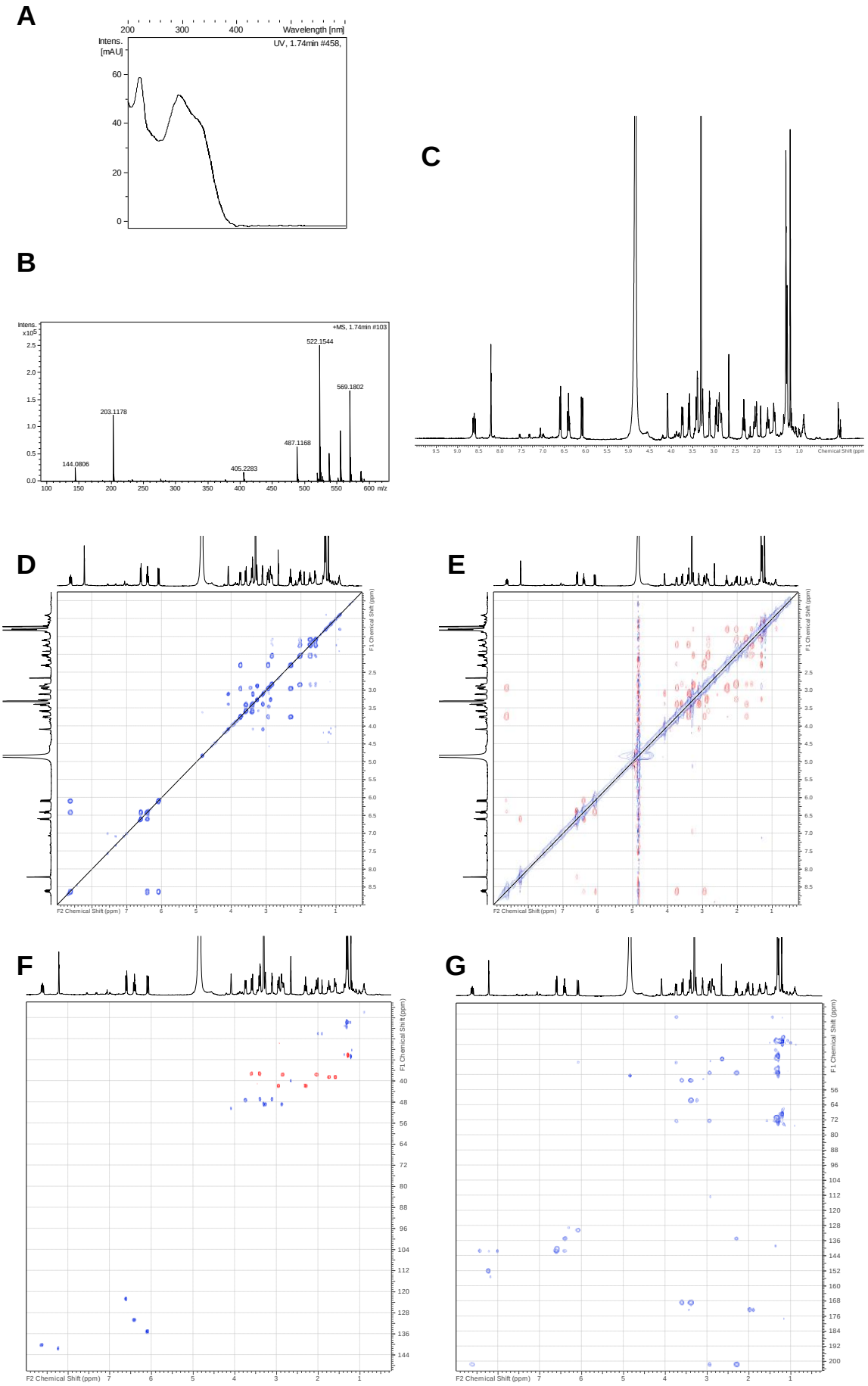

**Figure S8. Spectroscopic data of LRG O1 (5)**: (A) UV-DAD spectrum; (B) HRMS spectrum; (C) 1H NMR spectrum in DMSO-d6, 500 MHz; (D) COSY spectrum; (E) NOESY spectrum; (F) Edited HSQC spectrum; (G) HMBC spectrum.

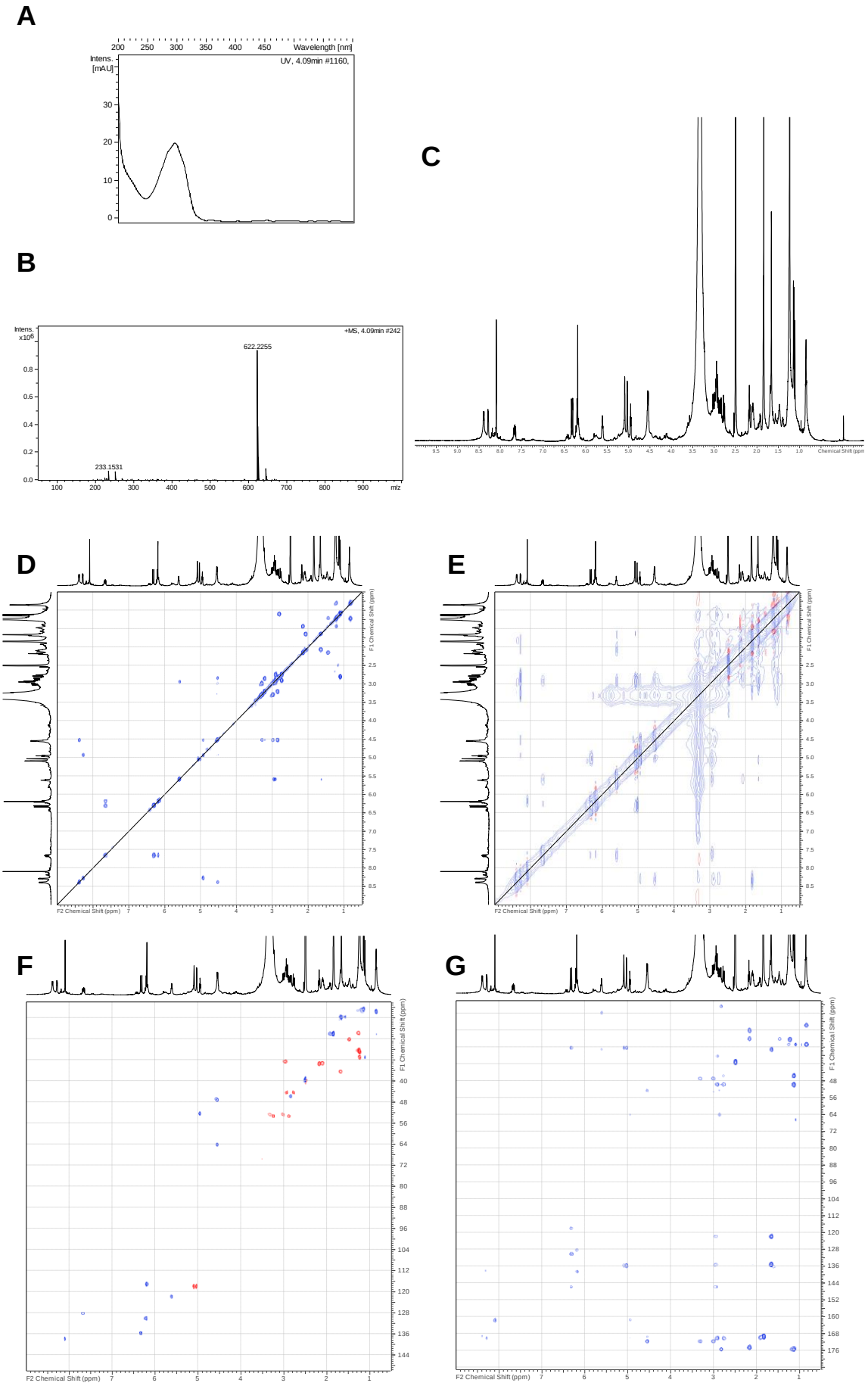

**Figure S9. (A) Energy-minimized molecular models of LRG A1 (1), LRG A2 (2), LRG A4 (4) and LRG O1 (5)**. The CysNAc moieties have been omitted for the sake of clarity. The measured distances (in Å) related to the observed key NOESY correlations are indicated. **(B) Determination of the oxime double bond stereochemistry.** Comparison of the experimental $^{13}C$ chemical shift for C-18 with the predicted one for both geometric isomers. The model compound used for validation is highlighted in the box.

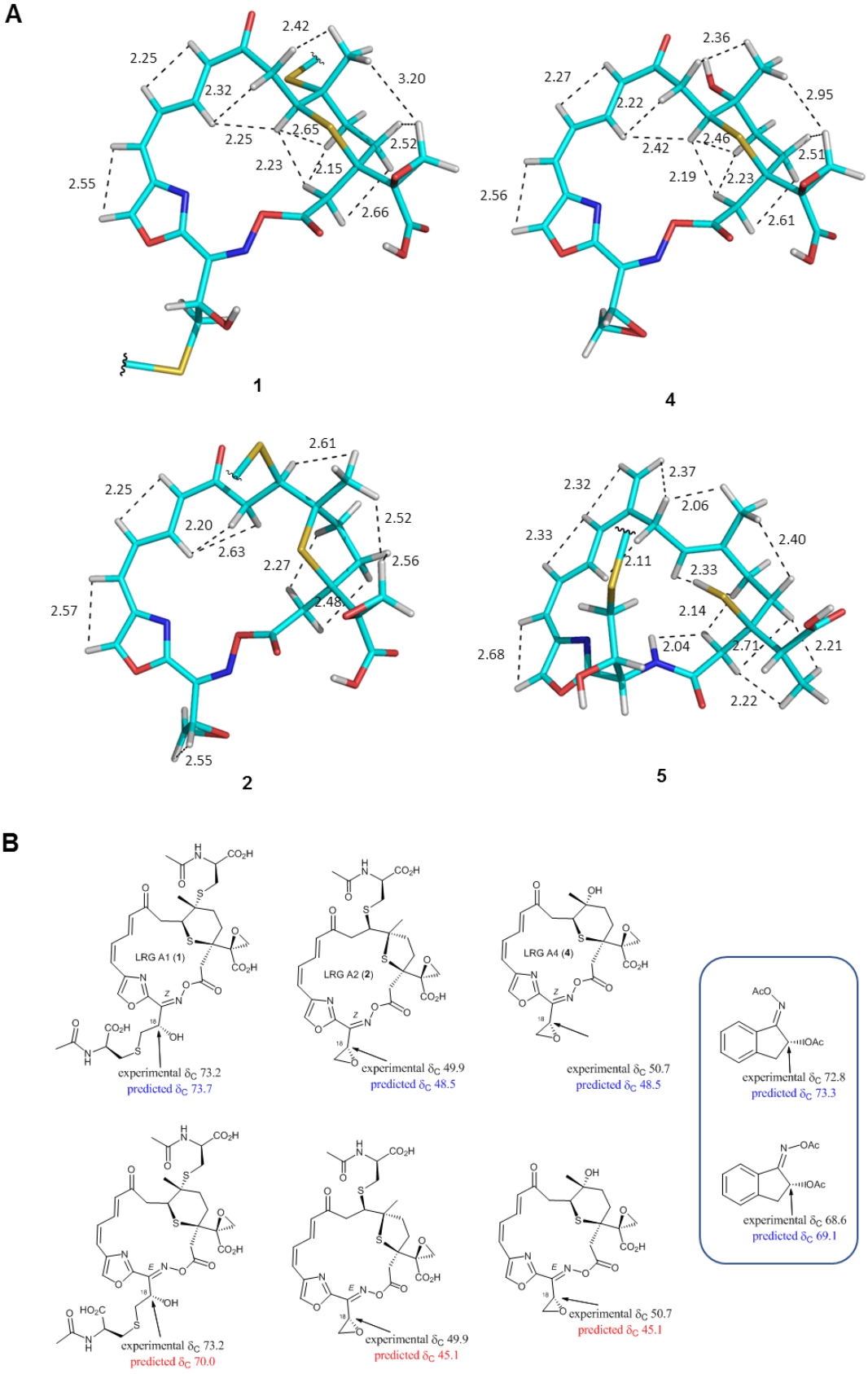

**Figure S10. Mechanism of formation of a CysNAc-S-conjugate after nucleophilic attack of mycothiol over an alkyl chloride.** The mycothiol S-conjugates obtained after nucleophilic attack over the episulfonium intermediate ("activated largimycin", Fig. 6) follow the same amidase hydrolysis fate.

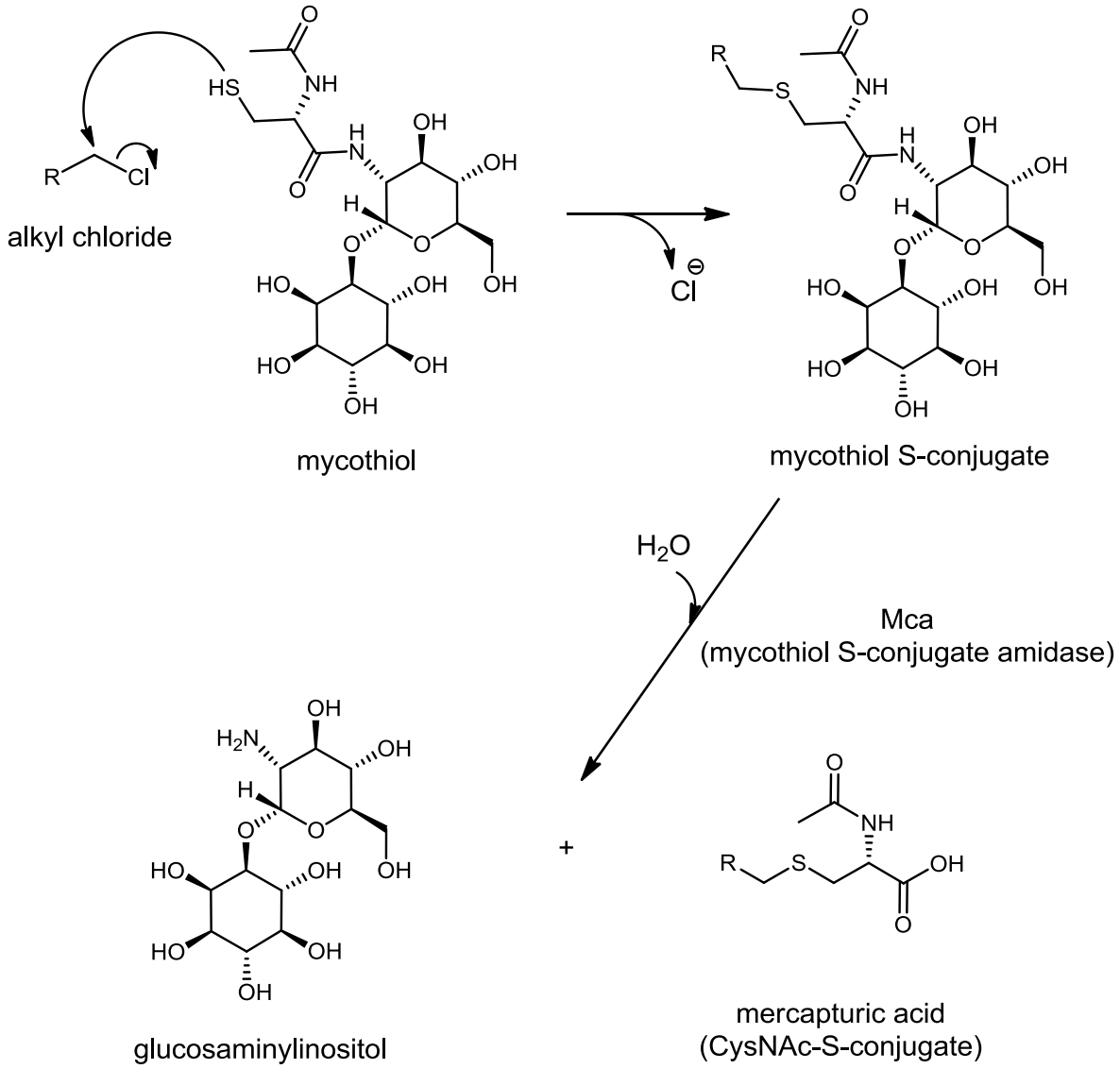

**Table S1. Oligonucleotides used for PCR.**

| PRIMER | SEQUENCE 5'-3' | PRIMER | SEQUENCE 5'-3' |
|---|---|---|---|
| **PRIMERS DESIGNED TO GENERATE MUTANTS** | | | |
| **MutAT831_A** | AAAGGATCCGATCGCCAACGTCACC | **orf22PHI rp** | AACCTGCAGATGATCTCTTCGTCTCGG |
| **MutAT831_B** | AAAGAATTCCCCGATCTCGAACATGTCAC | **orf22PHD up** | TAAGGATCCAAGTACGTCGGCATGCTG |
| **TetR21I up** | TATAGATCTTCAGCCAGGTCCTTCCGG | **orf22PHD rp** | CATGATATCCGATGATCGAGACGCTGA |
| **TetR21I rp** | ATAGAATTCTCACCGCCGAGTAGATGG | **Cit23 I up** | CCAAGATCTCAAGTAGGGAACGCAATCG |
| **TetR21D up** | ATAGGATCCTTGGTCGCCGAACTCG | **Cit23 I rp** | TATAAGCTTGAACGGACCCGCTCGAAG |
| **TetR21D rp** | AAAGATATCATGGGGGTCGGTCAGGTA | **Cit23 D up** | AAAGATATCCAGTCCCGATGACCGTGC |
| **TetR48I up** | TATCTCGAGTTCGTAGCGTGGGATTTC | **Cit23 D rp** | TAATCTAGAACCTGGTGCCACCCCGAA |
| **TetR48I rp** | ATAGAATTCTCGAAGATGGCCTGCTCA | **Cit52 I up** | AAAGAATTCCCAGAAGGTCGCCGACTA |
| **TetR48D up** | ATACATATGCCGACGAACTGGTGGTCT | **Cit52 I rp** | TATAAGCTTTCCTCCACCGGCTGCATC |
| **TetR48D rp** | AAAGATATCAGCTCCAGCACCCACTTC | **Cit52 D up** | TAAGATATCTCGACAGCGTCCCGGTCA |
| **PH19I up** | TAAGAATTCCTCGAGGTGCACCCCACG | **Cit52 D rp** | CGATCTAGACTACCTGGGCTCGATCCTG |
| **PH19I rp** | CATCTGCAGAGACGGCCATCCATCTAG | **orf53PHI up** | CCCGAATTCGCTGTACTGGTTCTTCTG |
| **PH19D up** | TAAGGATCCGAGGCGCTCAAGGCCGT | **orf53PHI rp** | TATAAGCTTACACGGCGCGCATCAGGA |
| **PH19D rp** | CCATCTAGAGGCAGGGCGTAGTCGAGC | **orf53PHD up** | CAAGATATCACCGTCGAGGACTGAGCC |
| **orf20TranspI up** | CGCAGATCTAGAATTGAGCAGTGGCAC | **orf53PHD rp** | ACATCTAGAATGATGCCCAACGCCAAG |
| **orf20TranspI rp** | ATAGGTACCATGGTGGCGTCGAGGATT | **Ox49I up** | TAAGAATTCGACGACGTGGAGCGGATG |
| **orf20TranspD up** | TAAGGATCCCACCGCAGGCCAAGAAC | **Ox49I rp** | TAACTGCAGCGATCACGTCCCAATCCA |
| **orf20TranspD rp** | ACCGATATCCCTTGTTGAAGTGGCTGA | **Ox49D up** | CCCGATATCTCTACACCCCCGAACTGCT |
| **orf22PHI up** | TAAGAATTCCGTCACGCGTCATCGACT | **Ox49D rp** | ACGTCTAGAGCCCTCCAGGAAGAGCCG |
| **PRIMERS DESIGNED TO EXPRESS GENES** | | | |
| **Reg24 up** | TTTGGATCCTGCGAGGCAACAGTATG | **ermECit23 rp** | ATAGCTAGCGGTTATGCCCTGGCCTC |
| **Reg24 rp** | TATGGATCCTCAGACCCTCGCCACGTA | **ermECit52 up** | AAAACTAGTCTGACCACCCGGCACCC |
| **Reg831_A** | ATTGGATCCTTCTAGGCGATTCGAG | **ermECit52 rp** | TAAGCTAGCGGTGTGCGGGGTCGTCA |
| **Reg831_B** | AAAGAATTCCGATGAACAGCAGTACG | **ermEOx49 up** | AAAACTAGTGTGCTGCGTACGGTCGGA |
| **ermECit23 up** | GCCACTAGTGTGAGAAAAGACGACGGG | **ermEOx49 rp** | ATAGCTAGCCGAACGGGACAAGGACGA |
| **PRIMERS DESIGNED TO VERIFIED MUTANTS** | | | |
| **TetR21c up** | AAAGGATCCCTACTCTAGCAGCGGCAA | **ermEorf22PH up** | CCCACTAGTTCTGAACAGGTTTCGGCC |
| **TetR21c rp** | ATAGAATTCTGACCCGGCAGCGCTC | **ermEorf22PH rp** | CCAGCTAGCGGTCACTTGTGGATCTTC |
| **TetR48c up** | ATAGAATTCCTGGACCCTACCGTGCA | **Cit23c up** | ATCAGGTTGTCAGTTCCGTG |
| **TetR48c rp** | TATAAGCTTGACCGCCCACTTTTGAGC | **Cit23c rp** | TCACGCGTCATCGACTCC |
| **PH19c up** | GCCGTGAAACGTGCTCTG | **ermEorf53PH up** | CCCACTAGTACCCGCAACTGGAAGACC |
| **PH19c rp** | AAGAACCTGGAGTAGCCGC | **ermEorf53PH rp** | CATGCTAGCTCAGTCCTCGATCACGG |
| **ermEorf20Transp up** | CAAACTAGTCTGTTAGCACCCCCGCAG | **Cit52c up** | CCTGCAACGGAGAGAAAA |
| **ermEorf20Transp rp** | CCAACTAGTCTCAGCCAGGTCCTTCCG | **Cit52c rp** | CGTGATCGAGGACTGACC |

**Table S2. Comparison of specifity-conferring codes of adenilation domains of hybrid NRPS/PKS from leinamycin (*lnm*)-type gene clusters.**

| Adenylation domain (Ser) | Clade[1] | Stachelhaus code[2] | 8 angstroms signature[3] |
|---|---|---|---|
| LrgI | VII* | DLFNAALVWK | LFCTFDLSVFDGNSALAGDIALGGPTETTVWSNV |
| Scan_P | VII | DLFNAAMVWK | LFCTFDLSVFDGNSALAGDIAMGGPTETTVWSNV |
| M1013 | VII* | DLFNAAMVWK | LFCTFDLSVFDGNSALAGDIAMGGPTETTVWSNV |
| CB01373_T | VII | DLFNAAMVWK | LFCTFDLSVFDGNSALAGDIAMGGPTETTVWSNV |
| **Consensus** | | **DLFNAAxVWK** | **LFCTFDLSVFDGNSALAGDIAxGGPTETTVWSNV** |
| **Adenylation domain (Cys)** | | | |
| LnmI | I | DLFNFSLVWK | LWCTFDLSVFDGNSFLSGDISLGGATEATVWSNW |
| CB01635_I | I | DLFNFSLVWK | LWCTFDLSVFDGNSFLSGDISLGGATEATVWSNW |
| CB02959_H | II | DLFNFSLVWK | LWCTFDLSVFDGNSFLSGDISLGGATEATVWSNW |
| GnmB | IX | DLFNFSLVWK | LWCTFDLSVFDGNSFLSGDISLGGATEATVWSNW |
| Sf56_B | IX | DLFNFSLVWK | LWCTFDLSVFDGNSFLSGDISLGGATEATVWSNW |
| CB02891_O | III | DLFNFSLVWK | LWCTFDLSVFDGNSFLSGDVSLGGATEATVWSNW |
| Sast_F | XI | DLFNFSLVWK | LWCTFDLSVFDGNSFLSGDVSLGGATEATVWSNW |
| Snov_X | XIII | DLFNFSLVWK | LWCTFDLSVFDGNSFLSGDVSLGGATEATVWSNC |
| Saes_V | XIV | DLFNFSLVWK | LWTTFDLSVFDGNSFLSGDISLGGATEATVWSNW |
| S110_Y | XII | DLFNFSLVWK | LWSTFDLSVFDGNSFLSGDVSLGGATEATVWSNY |
| WsmW | XII | DLFNFSLVWK | LWATFDLSVFDGNSFLSGDVSLGGATEAAVWSNY |
| Mtul_K | IV | DLFNFSLVWK | LWATFDLSVFDGNSFLSGDVSLGGATEATVWSNY |
| S109_K | V | DLFNFSLVWK | LWATFDLSVFDGNSFLSGDVSLGGATEATVWSNY |
| Slee_K | V | DLFNFSLVWK | LWATFDLSVFDGNSFLSGDVSLGGATEATVWSNY |
| TSRI0384-2_K | V | DLFNFSLVWK | LWATFDLSVFDGNSFLSGDVSLGGATEATVWSNY |
| Mcnb_U | XV | DLFNFSLVWK | LWATFDLSVFDGNSFLSGDVNLGGATEATVWSNW |
| Mmar_U | XV | DLFNFSLVWK | LWATFDLSVFDGNSFLSGDVNLGGATEATVWSNW |
| MI5_U | XV | DLFNFSLVWK | LWATFDLSVFDGNSFLSGDVNLGGATEATVWSNW |
| Maur_U | XV | DLFNFSLVWK | LWATFDLSVFDGNSFLSGDVNLGGATEATVWSNW |
| CB01201_B | X | DLFNFSLVWK | LWCTFDLSVFDGNSFLSGDVNLGGATEATVWSNW |
| Caci_P | VIII | DLFNFSLIWK | LWCTFDLSAFDGNSFLSGDVSLGGATEATIWSNY |
| Bubo_Q | XVIII | DLFNFSLVWK | LWCTFDLSVFDGNSYLSGDVSLGGATEATVWSNF |
| Sal964_Q | XVII | DLFNYSLVWK | LWCTFDLSVFDGNSYLSGDVSLGGATEATVWSNW |
| CB02613_J | XVI | DLFNYSLVWK | LWCTFDLSVFDGNSYLSGDVSLGGATEATVWSNW |
| **Consensus** | | **DLFNxSLxWK** | **LWxTFDLSVFDGNSxLSGDxxLGGATEATxWSNx** |
| **Consensus "Ser"** | | DLFN**AAxV**WK | L**FC**TFDLSVFDGNS**A**L**A**GD**IAx**GG**P**TE**T**T**V**WSN**V** |
| **Consensus "Cys"** | | DLFN**xSLx**WK | L**Wx**TFDLSVFDGNS**x**L**S**GD**xxL**GG**A**TE**A**T**x**WSN**x** |
| **Consensus** | | DLFN**xxxx**WK | L**xx**TFDLSVFDGNS**x**L**x**GD**xxx**GG**x**TE**x**T**x**WSN**x** |

[1]Pan *et al*. (2017). [2]Stachelhaus *et al*. (1999). [3]Raush *et al*. (2005). *this work. Identical amino acids in each type of binding domain are highlighted by yellow boxes. Amino acids that differ between the "Ser" and the "Cys" binding domains are highlighted by green boxes. Accession numbers of LrgI homologous clusters: *scan*, *Streptomyces canus* ATCC 12647 (WP_059300321); *M1013*, *Streptomyces* sp. (WP_076977086.1); *CB01373*, *Streptomyces* sp. (NZ_NNBK00000000.1). Accession numbers of *lnm*-type gene clusters: *lnm*, *S. atroolivaceus* S-140 (AF484556.1); *CB01635*, *Streptomyces* sp. CB01635 (NZ_NNBL00000000.1); *CB02959*, *Streptomyces* sp. CB02959 (NZ_NNBP00000000.1); *gnm*, *Streptomyces* sp. CB01883 (MF925481); *sf56*, *Streptomyces* sp. NRRL F-5630 (WP_037826198); *CB02891*, *Kitasatospora* sp. (NNBO00000000); *sast*, *Saccharotrix* sp. ST-888 (KJK59202.1); *snov*, *Streptomyces novaecaesareae* NRRL B-1267 (NZ_JNWQ01000000); *saes*, *Saccharotrix espanaensis* DSM 44229 (WP_015102008); *S110*, *Streptomyces* sp. NBRC 110035 (WP_042163782); *wsm*, *Streptomyces* sp. CB02120-2 (MF925482); *mtu*, *Micromonospora tulbaghiae* DSM45142 (SCE73393); *S109*, *Streptomyces* sp. NBRC 109436 (WP_064455881); *slee*, *S. leeuwenhoekii* DSM 42122 (CQR60407); *TSRI0384-2*, *Streptomyces* sp. TSRI0384-2 (NZ_NOWW00000000.1); *mcnb*, *Micromonospora* sp. CNB394 (WP_026267551); *mmar*, *M. marina* DSM 45555 (SCE83186); *MI5*, *Micromonospora* sp. L5 (WP_013476297); *maur*, *M. aurantica* ATCC 27029 (WP_013287043); *CB1201*, *Streptomyces* sp. CB1201 (NZ_NNBJ00000000.1); *caci*, *Catenulispora acidiphila* DSM 44928 (ACU715116.1); *bubo*, *Burkholderia ubonensis* RF25-BP1 (WP_059615564); *sal964*, *Salinispora arenícola* CNH964 (NZ_JAEY00000000.1); *CB02613*, *Streptomyces* sp. CB02613 (NZ_NNBN00000000.1).

**Table S3**. $^{1}H$ NMR (500 MHz) and $^{13}C$ NMR (125 MHz) data of LRG A1 in DMSO-$d_6$ and $CD_3OD$.[a]

| | **LRG A1** (DMSO-$d_6$) | | | **LRG A1** ($CD_3OD$) | |
|---|---|---|---|---|---|
| **Position** | **$\delta_C$, type** | **$\delta_H$ (*J* in Hz)** | **Position** | **$\delta_C$, type** | **$\delta_H$ (*J* in Hz)** |
| 1 | 166.6, C | | 1 | 169.6, C | |
| 2 | 37.1, $CH_2$ | 3.46, d (14.2)<br>3.12, d (14.2) | 2 | 38.2, $CH_2$ | a. 3.58, d (14.1)<br>b. 3.32, d (14.1) |
| 3 | 49.6, C | | 3 | 51.3, C | |
| 4 | 34.0, $CH_2$ | a. 2.46, br d (12.9)<br>b. 1.99, t (12.9) | 4 | 35.3, $CH_2$ | a. 2.73, br d (12.2)<br>b. 2.15, m |
| 5 | 35.1, $CH_2$ | a. 2.07, t (13.1)<br>b. 1.60, br d (13.1) | 5 | 36.7, $CH_2$ | a. 2.16, m<br>b. 1.73, br d (13.2) |
| 6 | 47.8, C | | 6 | 49.0, C | |
| 7 | 43.6, CH | 3.77, dd (12.5, 3.2) | 7 | 45.5, CH | 3.94, dd? (12.5, 3.5) |
| 8 | 43.1, $CH_2$ | a. 2.75, dd (13.0, 3.2)<br>b. 2.29, t (13.0) | 8 | 44.0, $CH_2$ | a. 2.88, dd (12.9, 3.8)<br>b. 2.42, t (12.9) |
| 9 | 197.6, C | | 9 | 201.3, C | |
| 10 | 134.4, CH | 6.15, d (16.2) | 10 | 135.5, CH | 6.18, d (16.3) |
| 11 | 136.6, CH | 8.32, dd (16.2, 11.4) | 11 | 139.1, CH | 8.59, dd (16.3, 11.3) |
| 12 | 128.8, CH | 6.38, t (11.4) | 12 | 130.2, CH | 6.43, t (11.4) |
| 13 | 122.0, CH | 6.58, d (11.4) | 13 | 123.2, CH | 6.60, d (11.4) |
| 14 | 139.5, C | | 14 | 141.7, C | |
| 15 | 141.2, CH | 8.50, s | 15 | 141.5, CH | 8.21, s |
| 16 | 151.1, C | | 16 | 152.6, C | |
| 17 | 152.2, C | | 17 | 153.5, C | |
| 18 | 72.2, CH | 4.71, t (7.1) | 18 | 73.2, CH | 4.95, dd (7.4, 6.4) |
| 19 | 35.7, $CH_2$ | a. 3.14, m<br>b. 3.00, dd (13.5, 7.1) | 19 | 37.6, $CH_2$ | a. 3.15, m<br>b. 3.04, dd (13.9, 6.3) |
| 20 | 17.3, $CH_3$ | 1.25, s | 20 | 17.8, $CH_3$ | 1.38, s |
| 21 | 59.2, C | | 21 | 61.7, C | |
| 22 | 47.8, $CH_2$ | a. 3.17, d (5.5)<br>b. 2.80, d (5.5) | 22 | 49.0, $CH_2$ | a. 3.31, m<br>b. 2.86, d (5.6) |
| 23 | 172.1, C | | 23 | 172.9, C | |
| 1' | 171.7, C | | 1' | 174.1, C | |
| 2' | 52.1, CH | 4.17, dd (14.0, 7.6) | 2' | 53.9, CH | 4.40, t (6.6) |
| 3' | 29.7, $CH_2$ | a. 2.66, dd (12.3, 8.4)<br>b. 2.58, dd (12.3, 5.6) | 3' | 31.5, $CH_2$ | 2.83, m |
| 4' | 169.2, C | | 4' | 172.6, C | |
| 5' | 22.3, $CH_3$ | 1.79, s | 5' | 22.2, $CH_3$ | 1.96, s |
| 2'-NH | | 8.16, d (7.5) | | | |
| 1'' | 172.2, C | | 1'' | 174.7, C | |
| 2'' | 52.1, CH | 4.39, dd (13.0, 7.8) | 2'' | 54.3, CH | 4.56, dd (7.9, 4.5) |
| 3'' | 33.2, $CH_2$ | a. 3.04, dd (13.5, 4.4)<br>b. 2.77, m | 3'' | 35.3, $CH_2$ | a. 3.16, m<br>b. 2.94, dd (13.9, 6.3) |
| 4'' | 169.3, C | | 4'' | 172.8, C | |
| 5'' | 22.3, $CH_3$ | 1.86, s | 5'' | 22.4, $CH_3$ | 2.00, s |
| 2''-NH | | 8.18, d (7.2) | | | |

[a] $^{13}C$ chemical shifts determined from the indirect dimension of HSQC and HMBC spectra

**Table S4**. $^{1}H$ NMR (500 MHz) and $^{13}C$ NMR (125 MHz) data of LRG A2 and LRG A4 in $CD_3OD$.[a]

| | **LRG A2** ($CD_3OD$) | | | **LRG A4** ($CD_3OD$) | |
|---|---|---|---|---|---|
| **Position** | **$\delta_C$, type** | **$\delta_H$ (*J* in Hz)** | **Position** | **$\delta_C$, type** | **$\delta_H$ (*J* in Hz)** |
| 1 | 168.4, C | | 1 | 169.0, C | |
| 2 | 43.4, $CH_2$ | a. 3.56, d (15.4)<br>b. 3.19, d (15.4) | 2 | 37.5, $CH_2$ | a. 3.59, d (14.1)<br>b. 3.40, d (14.1) |
| 3 | 61.4, C | | 3 | 51.6, C | |
| 4 | 41.3, $CH_2$ | a. 2.45, ddd (12.9, 5.5, 2.0)<br>b. 2.34, td (12.9, 5.7) | 4 | 37.8, $CH_2$ | a. 2.85, br d (14.0)<br>b. 2.04, t (14.0) |
| 5 | 40.4, $CH_2$ | a. 1.82, td (12.8, 5.7)<br>b. 1.73, m | 5 | 38.8, $CH_2$ | a. 1.75, t (14.0)<br>b. 1.59, dd (14.0, 3.2) |
| 6 | 64.7, C | | 6 | 72.8, C | |
| 7 | 51.3, CH | 3.56, dd (6.5, 4.9) | 7 | 47.5, CH | 3.74, dd (12.5, 3.5) |
| 8 | 43.8, $CH_2$ | a. 3.79, dd (18.5, 6.5)<br>b. 3.07, br d (18.5) | 8 | 42.1, $CH_2$ | a. 2.96, dd (13.0, 3.5)<br>b. 2.31, t (13.0) |
| 9 | 199.8, C | | 9 | 201.7, C | |
| 10 | 133.9, CH | 6.23, d (16.2) | 10 | 135.2, CH | 6.09, d (16.3) |
| 11 | 141.5, CH | 8.87, dd (16.1, 11.4) | 11 | 140.3, CH | 8.61, dd (16.3, 11.3) |
| 12 | 130.1, CH | 6.48, t (11.4) | 12 | 130.8, CH | 6.41, t (11.4) |
| 13 | 123.8, CH | 6.68, d (11.4) | 13 | 122.8, CH | 6.60, d (11.4) |
| 14 | 141.4, C | | 14 | 141.0, C | |
| 15 | 141.6, CH | 8.26, s | 15 | 141.6, CH | 8.22, s |
| 16 | 152.3, C | | 16 | 152.2, C | |
| 17 | 148.7, C | | 17 | n. d., C | |
| 18 | 49.9, CH | 4.17, dd (4.0, 2.4) | 18 | 50.7, CH | 4.09, dd (5.0, 2.4) |
| 19 | 46.9, $CH_2$ | a. 3.36, dd (5.5, 2.4)<br>b. 3.12, m | 19 | 47.1, $CH_2$ | a. 3.40, m<br>b. 3.11, dd (5.5, 4.4) |
| 20 | 30.9, $CH_3$ | 1.72, s | 20 | 18.0, $CH_3$ | 1.33, s |
| 21 | 61.4, C | | 21 | 62.0, C | |
| 22 | 49.3, $CH_2$ | a. 3.31, m<br>b. 2.95, d (5.8) | 22 | 49.0, $CH_2$ | a. 3.27, br d (5.5)<br>b. 2.87, m |
| 23 | 172.7, C | | 23 | n. d., C | |
| 1’ | 173.1, C | | | | |
| 2’ | 53.7, CH | 4.57, dd (7.2, 4.5) | | | |
| 3’ | 37.1, $CH_2$ | a. 3.11, m<br>b. 3.04, dd (13.7, 7.2) | | | |
| 4’ | 172.7, C | | | | |
| 5’ | 21.9, $CH_3$ | 1.99, s | | | |

[a] $^{13}C$ chemical shifts determined from the indirect dimension of HSQC and HMBC spectra

**Table S5. $^{1}H$ NMR (500 MHz) and $^{13}C$ NMR (125 MHz) data of LRG O1 in DMSO-$d_6$.[a]**

| | **LRG O1** (DMSO-$d_6$) | |
|---|---|---|
| **Position** | **$\delta_C$, type** | **$\delta_H$ (*J* in Hz)** |
| 1 | 170.0, C | |
| 2 | 44.6, $CH_2$ | a. 2.93, d (14.6)<br>b. 2.77, d (14.6) |
| 3 | 50.2, C | |
| 4 | 36.7, $CH_2$ | 1.68, m |
| 5 | 33.5, $CH_2$ | 2.10, m |
| 6 | 135.2, C | |
| 7 | 122.0, CH | 5.61, t (6.7) |
| 8 | 32.8, $CH_2$ | 2.97, m |
| 9 | 145.8, C | |
| 10 | 135.8, CH | 6.32, d (15.6) |
| 11 | 128.3, CH | 7.67, dd (15.5, 9.2) |
| 12 | 130.2, CH | 6.21, t (11.4) |
| 13 | 117.2, CH | 6.18, d (11.4) |
| 14 | 138.6, C | |
| 15 | 138.0, CH | 8.10, s |
| 16 | 161.1, C | |
| 17 | 52.6, CH | 4.95, t (5.0) |
| 17-NH | | 8.28, d (5.5) |
| 18 | 64.5, CH | 4.55, m |
| 19 | 53.5, $CH_2$ | a. 3.24, m<br>b. 2.88, m |
| 20 | 118.1, $CH_2$ | a. 5.09, br s<br>b. 5.03, br s |
| 21 | 16.1, $CH_3$ | 1.66, s |
| 22 | 46.0, CH | 2.84, q (6.9) |
| 23 | 13.1, $CH_3$ | 1.14, d (6.9) |
| 24 | 175.3, C | |
| 1’ | 171.6, C | |
| 2’ | 47.4, CH | 4.55, m |
| 3’ | 52.9, $CH_2$ | a. 3.32, m<br>b. 3.02, m |
| 4’ | 169.2, C | |
| 5’ | 22.3, $CH_3$ | 1.84, s |
| 2’-NH | | 8.39, d (7.9) |

[a] $^{13}C$ chemical shifts determined from the indirect dimension of HSQC and HMBC spectra

**SUPPLEMENTARY REFERENCES**


- Altschul, S.F., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic. Acids. Res. 25, 3389-3402.DeLano, W.L. (2002). The PyMOL Molecular Graphics System. (Palo Alto, CA, USA). DeLano Scientific. p. Available: http://www.pymol.org.
- Fernández, E., Weissbach, U., Sánchez Reillo, C., Braña, A.F., Méndez, C., Rohr, J., and Salas J.A. (1998). Identification of two genes from Streptomyces argillaceus encoding glycosyltransferases involved in transfer of a disaccharide during biosynthesis of the antitumor drug mithramycin. J. Bacteriol. 180, 4929-4937.
- Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., and Hopwood, D.A. Practical *Streptomyces* genetics (The John Innes Foundation, Norwich, UK. 2000).
- Lombó, F., Braña, A.F., Méndez, C., and Salas, J.A. (1999). The mithramycin gene cluster of *Streptomyces argillaceus* contains a positive regulatory gene and two repeated DNA sequences that are located at both ends of the cluster. J. Bacteriol. *181*, 642–647.
- Malmierca, M.G., González-Montes, L., Pérez-Victoria, I., Sialer, C., Braña, A.F., García Salcedo, R., Martín, J., Reyes, F., Méndez, C., Olano, C., et al. (2018). Searching for glycosylated natural products in actinomycetes and identification of novel macrolactams and angucyclines. Front. Microbiol. 9, 39.
- Martín, J., Crespo, G., González-Menéndez, V., Pérez-Moreno, G., Sánchez-Carrasco, P., Pérez-Victoria, I., Ruiz-Pérez, L.M., González-Pacanowska, D., Vicente, F., Genilloud, O., et al. (2014). MDN-0104, an antiplasmodial betaine lipid from Heterospora chenopodii. J. Nat. Prod. 77, 2118-2123.
- Menéndez, N., Nur-e-Alam, M., Fischer, C., Braña, A.F., Salas, J.A., Rohr, J., and Méndez, C. (2006). Deoxysugar transfer during chromomycin A3 biosynthesis in *Strepomyces griseus* subsp. *griseus*: new derivatives with antitumor activity. Appl. Environ. Microbiol. *72*, 167–177.
- Pérez-Victoria, I., Martín, J., and Reyes, F. (2016). Combined LC/UV/MS and NMR strategies for the dereplication of marine natural products. Planta Med. 82, 857-871.
- Sambrook, J. and Russell, D.W. (2001). Molecular Cloning: A Laboratory Manual, 3rd edn. (Cold Spring Harbor Laboratory).
- Sun, Y., Zhou, X., Liu, J., Bao, K., Zhang, G., Tu, G., Kieser, T., and Deng, Z. (2002). *Streptomyces nanchangensis*, a producer of the insecticidal polyether antibiotic nanchangmycin and the antiparasitic macrolide meilingmycin, contains multiple polyketide gene clusters. Microbiology *148*, 361–371.
- Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H.U., Bruccoleri, R., Lee, S.Y., Fischbach, M.A., Müller, R., Wohlleben, W., et al. (2015). AntiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res. 43, W237–W243.