The lipoprotein modification enzyme Lgt as a potential target for novel antibiotics

Simon Legood

Université Paris Cité

École doctorale Bio Sorbonne Paris Cité (ED562)

Unité Biologie et Génétique de la Paroi Bactérienne (Institut Pasteur)

The lipoprotein modification enzyme Lgt as a potential target for novel antibiotics

Par Simon LEGOOD

Thèse de doctorat de Microbiologie

Dirigée par Nienke BUDDELMEIJER

Présentée et soutenue publiquement le 13 décembre 2022

Devant un jury composé de :

Jean-Emmanuel HUGONNET , *Professeur des Universités, Sorbonne Université*..... *Président du jury*

Laure BEVEN - *Professeur des Universités, Université de Bordeaux*..... Rapporteur

Arnaud CHASTANET - CR-HDR, *Université Paris-Saclay, INRAE*.....Rapporteur

Martin PICARD - DR-HDR, *Université Paris Cité, IBPC*..... .....Examinateur

Anne-Marie WEHENKEL - CR-HDR, *Université Paris Cité, Institut Pasteur* .....Examinatrice

Nienke BUDDELMEIJER - CR-HDR, *Université Paris Cité, Institut Pasteur*..... Directrice de thèse

Pasteur-Paris Université – Oxford University International PhD Program

Résumé étendu en français

L'enzyme de modification des lipoprotéines Lgt comme cible potentielle pour de nouveaux antibiotiques

La nécessité de développer de nouveaux antibiotiques est évidente. Avec des taux élevés de décès liés à la RAM, la crise de la RAM est un problème mondial. Divers mécanismes sont utilisés pour s'attaquer à ce problème : depuis les vaccins bactériens et des modifications du microbiote jusqu'à l'augmentation de la sensibilisation et à l'amélioration de l'administration des antibiotiques existants. Une autre option consiste à développer de nouveaux antibiotiques. Ceux-ci pourraient être dirigés contre des cibles existantes, bien que la résistance puisse se développer rapidement contre ces inhibiteurs, ou contre de nouvelles cibles. Des rapports suggèrent que l'antibiotique idéal aura plusieurs cibles ou plusieurs voies métaboliques, car cela réduit l'apparition de la résistance.

Parmi les agents pathogènes qui causent le plus de mortalité en raison de la RAM, la majorité est constituée de bactéries didermiques appartenant à la catégorie des bactéries Gram-négatives. Il existe un certain nombre de processus essentiels chez toutes les bactéries, comme la synthèse et la réplication de l'ADN ou la synthèse de protéines. Certaines voies essentielles uniques existent chez les bactéries Gram-négatives en raison de l'importance de leur membrane externe. Ces voies essentielles comprennent le système Bam qui insère les protéines dans la membrane externe, la voie de translocation du LPS qui transporte et insère la molécule de lipopolysaccharide de la membrane externe jusqu'à sa destination et enfin la voie de modification des lipoprotéines qui est nécessaire pour modifier les lipoprotéines impliquées dans les voies Bam et LPS. Aucune de ces trois voies n'est la cible d'antibiotiques. Comme la voie de modification des lipoprotéines est nécessaire aux deux autres voies essentielles, nous avons cherché à savoir si elle pouvait être une bonne cible pour de nouveaux antibiotiques. Cette voie est constituée de trois enzymes, Lgt, Lsp et Lnt, suivies du mécanisme de translocation Lol. Comme il s'agit de la première enzyme de la voie et d'une protéine extrêmement bien conservée, nous avons choisi d'étudier Lgt plus en détail.

Pour remettre en question l'hypothèse selon laquelle Lgt est une bonne cible pour de nouveaux antibiotiques, nous avons utilisé trois axes d'étude. Le premier consistait à étudier la conservation de Lgt, le second à comprendre plus en détail son caractère essentiel et le dernier à mettre au point un test *in vitro* pour le dépistage des inhibiteurs.

Pour affiner nos efforts, nous avons sélectionné des pathogènes clés de la RAM à étudier plus en détail. Nous avons constaté que Lgt est bien conservé au niveau de la séquence, à l'exception du motif clé HGGL décrit dans la litérature. Ce motif contient le résidu essentiel H^{103} qui a été considéré comme un résidu clé dans l'activité catalytique de Lgt. Cependant, en accord avec les rapports précédents, H^{103}

n'est pas largement conservée avec la tyrosine ou le tryptophane observés dans les pathogènes sélectionnés [78, 79, 81]. Cela soulève la possibilité que H^{103}, bien que clairement essentielle pour l'activité, ne soit pas la base catalytique proposée pour la réaction. L'emplacement de H^{103} dans la fente latérale, un canal de sortie possible pour les substrats, peut jouer un rôle plus structurel comme une porte, permettant ou bloquant l'entrée du substrat. D'autres études sont nécessaires avant de pouvoir conclure que les modèles précédents de l'activité enzymatique de la Lgt sont corrects. Les modèles de calcul utilisent un peptide lipobox raccourci et non la lipoprotéine complète. Par conséquent, le rôle de la région périplasmique étendue de la lipoprotéine ne peut pas être évalué. Étant donné que la liaison protéine-protéine prédite de L6-7, du bras-1, du bras-2 et de certains domaines de tête sont absentes, une modélisation plus poussée avec la lipoprotéine complète pourrait faire la lumière sur les interactions possibles lipoprotéine-Lgt.

Une séquence conservée nouvellement décrite, appelée « motif du cou », située sous le domaine de tête de l'enzyme, est hautement conservée au sein des firmicutes et des protéobactéries, et se trouve à proximité du motif de signature Lgt dans le noyau catalytique proposé de l'enzyme. Le rôle du motif du cou doit être exploré davantage et peut fournir une spécificité d'espèce pour l'enzyme.

Nous avons cherché à savoir si la *lgt* de *A. baummannii, P. aeruginosa* et *H. pylori* pouvait complémenter deux souches de déplétion, Δ*lgt*^{P} et Δ*lgt*^{C}. Dans la souche Δ*lgt*^{P}, la *lgt* de type sauvage est présente sur un plasmide multicopie et dans la souche Δ*lgt*^{C}, elle est présente en une seule copie sur le chromosome. Dans les deux souches, le gène *lgt* de type sauvage est sous le contrôle de P_{ara}. Les gènes *lgt* complémentaires sont présents sur le plasmide pAM238 sous le contrôle du promoteur P_{lac}.

Nous avons observé que si la croissance était entièrement rétablie dans le modèle Δ*lgt*^{P}, cela pouvait être dû à l'expression résiduelle du gène *lgt* de type sauvage sous le contrôle de P_{ara} sur un plasmide à nombre de copies élevé. Par conséquent, la capacité des gènes *lgt* à rétablir la croissance par complémentation pourrait être interprétée davantage comme un soutien apporté au gène *lg*t de type sauvage résiduel et comme le fait que les cellules ne dépendent pas entièrement de l'enzyme hétérologue. Dans le modèle Δ*lgt*^{C}, le problème de l'expression résiduelle de Lgt est en grande partie éliminé grâce à la copie unique de *lgt* de type sauvage présente sous le contrôle de P_{ara}. Dans cette souche, nous observons que seule la *lgt* de *H. pylori* peut compléter entièrement la croissance et la viabilité. Ceci est surprenant car *H. pylori* est la plus éloignée des trois enzymes du point de vue de l'évolution. Les études de complémentation devraient inclure des Lgt provenant d'un plus grand nombre d'espèces, en particulier des bactéries Gram positif d'intérêt clinique telles que *Streptococcus spp*. Jusqu'à présent, toutes les études de complémentation réussies chez *E. coli* ont porté sur des enzymes Lgt présentant H^{103} et non une variation de ce résidu clé.

L'avènement d'un logiciel de prédiction structurelle à haut niveau de confiance, tel qu'AlphaFold2, nous a permis d'analyser les différences dans la structure des Lgt des principales espèces pathogènes. Alors qu'il est prévu qu'elles soient très similaires sur la majorité de l'enzyme, le domaine de la tête exposée au périplasme a montré une grande variabilité. Les entérobactéries ont un grand domaine de tête avec deux hélices alpha qui dépasse largement le plan de la membrane et *M. tuberculosis* possède un domaine de tête encore plus grand. Les firmicutes et les agents pathogènes protéobactériens plus éloignés ont un domaine de tête plus petit qui peut reposer plus près du plan de la membrane. Nous avons émis l'hypothèse que le domaine de tête pouvait jouer un rôle dans la localisation des composants de la voie de modification des lipoprotéines à proximité les uns des autres. Cependant, la prédiction de ColabFold a jugé peu probables les interactions entre Lgt, Lsp et Lnt. Le domaine de tête de Lgt d'*E. coli* avait une faible prédiction de liaison protéine-protéine, mais les domaines de tête des protéobactéries éloignées *H. pylori*, le firmicute *S. aureus* et *M. tuberculosis* avaient une prédiction de liaison protéine-protéine plus élevée. Cela suggère que le domaine de tête périplasmique peut avoir une signification fonctionnelle. Pour explorer cette hypothèse de manière expérimentale, nous avons cloné les domaines de tête de *H. pylori, S. aureus* et *M. tuberculosis* dans le gène *lgt* de *E. coli* et l'avons exprimé dans les souches de déplétion Lgt ; Δlgt^{P} et Δlgt^{C}. Nous avons constaté que seul le domaine de tête de *H. pylori* pouvait compléter la croissance des souches déplétées. Cela suggère que les domaines de tête peuvent jouer un rôle fonctionnel. Cependant, bien que les protéines aient été produites et que la production de protéines ne semble pas être en corrélation avec la viabilité et la croissance, nous n'avons pas encore déterminé la localisation des protéines ou leur stabilité. Comme le noyau de l'enzyme est bien conservé, les domaines de tête peuvent fournir un site d'inhibition à spectre plus étroit, tandis que le noyau de l'enzyme peut fournir une inhibition à spectre plus large.

Le deuxième axe de l'étude était d'explorer l'essentialité de Lgt. Les molécules inhibitrices de la voie de modification des lipoprotéines telles que celles qui inhibent Lsp ou la machinerie Lol perdent leurs capacités inhibitrices dans *E. coli* lorsque la lipoprotéine abondante Lpp n'est pas produite. Il a été démontré que la capacité de Lpp à se lier au peptidoglycane tout en étant mal localisé à la membrane interne est léthal. Nous avons donc cherché à vérifier si cela était vrai pour l'inhibition de Lgt. Comme nous n'avons pas encore d'inhibiteur de Lgt, nous avons utilisé deux modèles de déplétion pour étudier la perte de Lgt, Δlgt^{P} et Δlgt^{C}. Nous avons observé que dans le modèle Δlgt^{P}, la production de Lgt à partir de P_{ara} dans des conditions réprimées, en dessous de niveaux détectables, était probablement suffisante pour rétablir la croissance après une longue phase de latence. Ces révertants de croissance ont été explorés plus avant et nous avons observé que, bien que le phénotype de croissance révertant permette une croissance permanente de type sauvage dans des conditions réprimées, il n'y avait pas de mutations sur le chromosome bien que la présence du plasmide soit indispensable. Il est probable

que le rétablissement de la croissance ne provient pas simplement de la production résiduelle de Lgt à partir de P_{ara}, car la phase de latence prolongée n'est pas observée lorsque Δlgt^{P} est cultivée dans de faibles niveaux d'inducteur de P_{ara}, le L-arabinose. La suppression de *lpp* dans des conditions où *lgt* n'était pas activement exprimé était également possible chez Δlgt^{P}.

Cependant, dans Δlgt^{C}, où il y a une seule copie de Lgt sous le contrôle de P_{ara}, la croissance et la viabilité n'ont pas été rétablies dans des conditions restrictives et la délétion de *lpp* n'a pas rétabli la croissance. Cela suggère qu'un autre mécanisme de mort cellulaire se produit. Il convient de noter que Lpp a rétabli la croissance et la viabilité à de faibles niveaux d'expression de *lgt* dans ce contexte, ce qui montre que, bien que Lpp ne soit pas la seule cause de la mort cellulaire, la présence de la lipoprotéine abondante est un facteur déterminant. Au cours de cette étude, nous avons observé de profondes modifications morphologiques de la cellule dans des conditions de déplétion de *lgt* qui différaient en présence et en absence de Lpp. Lorsque l'expression de *lgt* est réprimée, les cellules deviennent plus arrondies aux pôles, de multiples pôles sont observés et leur surface augmente avant de finir par lyser. En l'absence de Lpp, les cellules s'allongent et ne s'arrondissent pas aux pôles mais finissent tout de même par lyser. Nous avons conclu que bien que la Lpp soit un facteur de mort cellulaire en l'absence de modification des lipoprotéines, d'autres lipoprotéines essentielles ou clés ont un rôle plus important. En l'absence de Lpp, l'expression résiduelle de *lgt* provenant de P_{ara} dans Δlgt^{P} est suffisante pour permettre la modification d'un nombre suffisant de lipoprotéines clés, telles que LolA, ou BamB-E par exemple, pour maintenir la survie des cellules. Cependant, lorsque cette expression résiduelle est réduite dans Δlgt^{C}, même en l'absence de Lpp, la transformation des lipoprotéines est insuffisante pour maintenir la survie.

Par conséquent, en accord avec un rapport précédent qui montrait que la délétion de *lpp* ne réduisait pas l'inhibition de Lgt par un inhibiteur peptidique, nous avons également observé que Lpp n'est pas la cause principale de la mort dans des conditions de déplétion de Lgt.

Ce que cette étude a mis en évidence, c'est l'importance de cette voie pour le maintien de la viabilité mais aussi de la forme de la cellule. Les changements morphologiques de la cellule lorsque Lgt est réduit sont sévères et un rapport précédent a montré qu'une souche dépourvue de Lgt est plus sensible à la destruction par le sérum. Ces données suggèrent qu'une inhibition incomplète de Lgt pourrait avoir un effet suffisant sur la cellule pour que l'inhibition totale ne soit pas nécessaire dans un environnement infectieux. Comme nous l'avons vu, l'inhibition partielle peut être suffisante pour avoir un effet sur la condition physique et donc soutenir le système immunitaire, ce qui facilite l'élimination de l'infection. Il a déjà été démontré que la déplétion de Lgt réduit la CMI d'une série d'antibiotiques et notre observation que Lgt a un effet serveur sur la morphologie cellulaire pourrait indiquer que Lgt

est une bonne cible dans le cadre d'une thérapie combinée. Avec l'avènement des tests de criblage pour Lnt ou la machinerie Lol et la description de certains inhibiteurs pour Lsp et LolCDE, la possibilité d'une thérapie combinée avec des cibles multiples dans la même voie est envisagée.

Le lipoprotéome complet d'*E. coli* n'a pas encore été défini de manière exhaustive et des dizaines de lipoprotéines d'*E. coli* n'ont pas de fonction assignée. Il en va de même pour de nombreuses autres espèces de bactéries et, par conséquent, l'impact total de l'inhibition de la voie de modification des lipoprotéines n'est pas connu. Certaines études ont cherché à étudier les lipoprotéomes des bactéries en appliquant la protéomique et la lipidomique par MALDI-TOF MS et ces méthodes ont permis de découvrir de nouvelles modifications post-translationelles. Ces travaux doivent être poursuivis. Une méthode de détermination et de définition des lipoprotéomes des bactéries permettrait d'affiner et d'améliorer les outils de prédiction sur lesquels on compte beaucoup. En identifiant les lipoprotéines, nous pourrions alors mieux comprendre les rôles qu'elles jouent et de faire la lumière sur la spécificité ou la préférence des lipoprotéines pour Lgt. Comme nous avons observé une capacité de complémentation différentielle des enzymes Lgt de différentes espèces dans *E. coli*, nous proposons que cela peut être dû à la spécificité lipoprotéine-Lgt. Par exemple, les lipoprotéines d'*E. coli* peuvent être mieux supportées par la Lgt de *H. pylori* que par celle de *A. baummannii*.

Des études *in vivo* plus poussées sur la déplétion ou l'inhibition de Lgt nous permettraient également de mieux comprendre le rôle des lipoprotéines dans les modèles d'infection et pourraient mettre en évidence d'autres cibles pour de nouveaux antibiotiques.

Nos résultats confirment que Lgt est effectivement essentielle chez *E. coli* et que le gène *lgt* est présent en une seule copie chez 22 bactéries pathogènes clés. Comme nous l'avons vu, les produits génétiques uniques et essentiels utilisés comme cibles d'antibiotiques sont susceptibles de présenter des taux de développement de résistance plus élevés, car la pression de sélection est forte. C'est certainement un problème avec Lgt, car aucune autre protéine n'a été découverte pour jouer le même rôle. Puisque nous avons observé un phénotype de croissance révertant lorsque Δlgt^{P} est cultivé en conditions restrictives qui n'est probablement pas causé uniquement par l'expression basale de P_{ara}, il est possible que d'autres mécanismes cellulaires soutiennent la croissance et la viabilité des cellules en l'absence d'une quantité suffisante de Lgt. Diao *et al.* (2021) déclarent qu'ils n'observent pas de développement de résistance à leur inhibiteur de peptide cyclique dans des conditions de laboratoire, ce qui est un signe prometteur, mais cela soulève la question de savoir si Lgt est la seule cible de leur inhibiteur ou s'il a des effets multiples sur la cellule.

Enfin, nous avons cherché à développer un test *in vitro* pour étudier Lgt et pour cribler les inhibiteurs. Nous avons réussi à purifier Lgt après avoir modifié les protocoles de purification précédemment

rapportés et nous avons constaté que l'enzyme était active dans notre test de retard sur gel adapté. Nous n'avons pas réussi à mettre au point un test quantitatif basé sur la fluorescence, mais cette voie devrait être explorée plus avant. Nous pensons que l'utilisation de la spectrométrie de masse, en particulier le MALDI-TOF-MS, pourrait constituer une meilleure option pour l'étude de la Lgt. Cette méthode peut être adaptée à un débit élevé et nécessite une préparation minimale des échantillons. Son principe de fonctionnement a été démontré mais elle n'a pas encore été utilisée à des fins de dépistage.

D'autres aspects doivent être étudiés plus avant. Certaines bactéries possèdent plusieurs enzymes Lgt, mais on ne connaît pas encore l'ampleur de ce phénomène. Une analyse phylogénétique à grande échelle devrait être menée pour déterminer quels organismes ont des Lgt multiples et quel est l'avantage physiologique de ce phénomène. Les mécanismes de régulation de la voie de modification des lipoprotéines ne sont pas connus et comme il a été démontré que différentes formes de lipoprotéines peuvent être préférées dans certaines conditions, le mécanisme par lequel ceci est contrôlé serait un domaine d'étude intéressant.

Lgt est essentielle et présente dans tous les pathogènes RAM clés, elle possède une cavité centrale conservée qui est considérée comme le noyau catalytique et qui peut fournir un site d'inhibition à large spectre, mais elle possède également un domaine de tête variable qui pourrait être ciblé par des inhibiteurs à spectre plus étroit. La fonction du groupe de tête est encore inconnue. Un crible plus efficace doit être développé pour l'activité de Lgt afin de trouver des inhibiteurs ciblés. En l'absence d'inhibiteurs, il est difficile d'évaluer leur potentiel *in vivo*. Le peptide cyclique G2824 donne un aperçu mais ne suffit pas à répondre à l'hypothèse. Toutefois, compte tenu du besoin urgent de nouveaux antibiotiques, le ciblage de Lgt comme nouvelle cible est prometteur.

Résumé en français

L'enzyme de modification des lipoprotéines Lgt comme cible potentielle pour de nouveaux antibiotiques

La résistance aux antimicrobiens (RAM) est une menace croissante pour la santé publique. Les moyens de combattre cette crise sanitaire sont nombreux et l'un d'eux consiste à développer de nouveaux antibiotiques contre de nouvelles cibles pour combattre les infections.

Dans cette étude, nous avons identifié la voie de modification des lipoprotéines comme une cible potentielle pour de nouveaux antibiotiques en raison de son caractère essentiel dans de nombreux pathogènes clés de la RAM. La voie de modification des lipoprotéines est constituée de trois enzymes, Lgt, Lsp et Lnt. Nous avons émis l'hypothèse que Lgt, en tant que première enzyme de cette voie de modification post-traductionnelle, serait une cible privilégiée pour de nouveaux inhibiteurs. À cette fin, nous avons étudié la conservation de Lgt et constaté qu'elle est généralement bien conservée parmi les principaux agents pathogènes de la RAM au niveau de la séquence et de la structure et qu'elle se présente comme un gène à copie unique et absent chez les espèces eucaryotes. Nous réaffirmons les conclusions selon lesquelles un résidu d'histidine clé censé jouer un rôle primordial dans la catalyse n'est pas largement conservé et que notre compréhension de la réaction doit donc être réévaluée. Nous avons utilisé deux souches d'*E. coli* dépourvues de Lgt et réalisé des études de complémentation avec des Lgt d'*A. baumannii*, de *P. aeruginosa* et d'*H. pylori*. Nous avons constaté que le Lgt d'*H. pylori* était le plus efficace pour complémenter la croissance et la viabilité. Le domaine de tête, qui fait saillie dans le périplasme, est l'une des parties de la structure qui présente des variations. Nous avons observé que l'échange du domaine de tête d'*E. coli* avec celui d'*H. pylori* était viable dans les études de déplétion, mais que les domaines de tête de *M. tuberculosis* et de *S. aureus* ne l'étaient pas. Cela ouvre la voie à l'étude d'un rôle fonctionnel du domaine de tête de Lgt et soulève la possibilité qu'il puisse être une cible pour une inhibition à spectre étroit.

Afin de mieux comprendre le caractère essentiel de Lgt, nous avons utilisé les deux souches de déplétion de Lgt dans *E. coli* et observé que Lgt est essentielle. La lipoprotéine Lpp est considérée comme un déterminant clé de l'essentialité de Lsp et Lnt. Nous constatons que, bien que la croissance, mais pas la morphologie, soit améliorée en l'absence de Lpp, Lgt est toujours essentielle. Les changements de morphologie à de faibles niveaux de production de Lgt nous ont conduits à émettre l'hypothèse que d'autres lipoprotéines importantes pour la biogenèse de l'enveloppe cellulaire sont le facteur clé de l'essentialité de Lgt.

Enfin, nous avons cherché à développer un test *in vitro* pour étudier les activités enzymatiques de Lgt et cribler des inhibiteurs à petites molécules. Bien que nous ayons réussi à purifier l'enzyme et à

adapter un test de retard sur gel à faible débit pour étudier son activité, nous devons optimiser un test quantitatif basé sur la fluorescence compatible avec le criblage à haut débit.

Lgt est essentielle chez les protéobactéries et nous mettons en évidence de nouvelles voies d'étude, telles que l'importance des domaines de tête et le rôle du résidu histidine clé dans l'activité. Nos données préliminaires sont prometteuses pour le développement d'un test *in vitro* à haut débit pour étudier l'activité enzymatique de Lgt.

Mots clés : Lgt, modification des lipoprotéines, résistance antimicrobienne, tests d'activité *in vitro*, essentialité des gènes.

Summary in English

Antimicrobial resistance (AMR) is an increasing threat to public health. Ways to combat this health crisis are numerous and one of these ways is to develop new antibiotics against novel targets to combat infections.

In this study we identified the lipoprotein modification pathway as a potential target for novel antibiotics due to its essentiality in many key AMR pathogens. The lipoprotein modification pathway consists of three enzymes, Lgt, Lsp and Lnt. As the first enzyme in this post-translational modification pathway, we hypothesized that Lgt would be a prefered target for novel inhibitors. To this end we have explored the conservation of Lgt and find that it is generally well conserved across key AMR pathogens at a sequence and structural level and present as a single copy gene. We re-state the findings that a key histidine residue thought to play a primary role in catalysis is not widely conserved and therefore our understanding of the reaction needs to be re-evaluated. We employed two Lgt depletion strains of *E. coli* and performed complementation studies with Lgt from *A. baumannii, P. aeruginosa* and *H. pylori* and found that *H. pylori lgt* was the most efficient at complementing growth and viability. One area of the structure which showed variation was the so-called head domain which protrudes into the periplasm. We observe that exchange of the head domain from *E. coli* with that of *H. pylori* was viable in depletion studies but the head domains from *M. tuberculosis* and *S. aureus* were not. This opens the door to the study of a functional role of the Lgt head domain and raises the possibility that it may be a target for narrow spectrum inhibition.

In order to understand the essentiality of Lgt in greater depth we employed the two depletion strains of Lgt in *E. coli* and observed that Lgt is essential. The lipoprotein Lpp is noted as a key determinant in the essentiality of Lsp and Lnt. We find that although growth, but not morphology, is improved in the absence of Lpp, Lgt is still essential. Changes in morphology at low levels of Lgt production led us to hypothesise that other lipoproteins important for cell envelope biogenesis are the key factor in Lgt essentiality.

Finally, we sought to develop an *in vitro* assay to study the enzymatic activities of Lgt and screen for small molecule inhibitors. Although we successfully purified the enzyme and adapted a low throughput gel-shift assay to study its activity, we need to optimise a quantitative fluorescence-based assay compatible with high throughput sequencing.

Lgt is essential in proteobacteria and we highlight new avenues of study, such as the importance of the head domains and the role of key histidine residue in activity and our preliminary data are promising for the development of a high throughput *in vitro* assay to study Lgt enzymatic activity.

Key Words: Lgt, lipoprotein modification, antimicrobial resistance, in vitro activity assays, gene essentiality

Acknowledgements

Firstly, I would like to thank the members of the jury for accepting to read and evaluate this thesis.

I would like to thank Dr. Nienke Buddelmeijer for providing me with the opportunity to study the subject of this thesis and also for her supervision, guidance and support over the past three years.

To the head of the Biology and Genetics of the Bacterial Cell wall (BGPB) Unit, Dr. Ivo Boneca for his support and for hosting me in his laboratory.

To the members of the BGPB unit for the many helpful and interesting scientific discussions but for making the past three years an enjoyable social experience.

To the Pasteur-Paris University (PPU) International Doctoral Program for the funding, support and training.

To the PPU-Oxford office and Prof. Syma Khalid for the opportunity to spend time at Oxford University and study areas of science beyond my field.

To the various platforms at Institut Pasteur who provided the skills and expertise to support my research to go further.

Finally, to my wife Hannah and daughter Elsie for their love and support during a busy three-years in a foreign country.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Full term
AMR	Antimicrobial resistance
Ara	Arabinose
ATP	Adenosine triphosphate
BSA	Bovine Serum Albumin
CBB	Coommassie Brilliant Blue
CDC	Center for Disease Control
CL	Cardiolipin
CM	Cytoplasmic membrane
CMC	Critical micelle concentration
DAG	Diacylglyceryl
DDM	n-dodecyl-β-D-maltoside
DEPC	Diethyl pyrocarbonate
DHA	Dihydroxyacetone
DHAP	Dihydroxyacetone phosphate
D*lgt*^{C}	Lgt depletion strain
D*lgt*^{P}	Lgt depletion strain
DM	Decylmaltoside
DOLOP	Lipoprotein prediction tool
DPPG	1,2-Dipalmitoyl-sn-glycero-3-phosphoglycerol
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EPEC	Enteropathogenic *E. coli*
G1P	Glycerol-1-phosphate
G3PDH	Glycerol-3-phosphate dehydrogenase
GDH	Glycerol dehydrogenase
GI	Gastrointestinal
Glu	Glucose
GST	Glutathione-S-transferase
HAI	Healthcare-associated infection
HIV	Human immunodeficiency virus
HRP	Horseradish peroxidase
IM	Inner membrane
IMV	Inverted membrane vescicles
IPTG	Isopropyl β-d-1-thiogalactopyranoside
LB	Luria broth
LBA	Luria broth agar
LC-MS	Liquid Chromatography – Mass Spectrometry
Lgt	Prolipoprotein Diacylglyceryl Transferase
Lgt^{ts}	Temperature sensitive Lgt depletion strain
LMNG	Lauryl Maltose Neopentyl Glycol
LMP	Lipoprotein modification pathway
LPS	Lipopolysaccharide
MALDI-TOF	Matrix Assisted Laser Desorption Ionization - Time of Flight
MCS	Multiple cloning site
MDR	Multi-drug resistant
MIC	Minimum inhibitory concentration

Abbreviation	Full term
MRSA	Methicillin resistant *S. aureus*
MSA	Multiple sequence alignments
MTSES	Sodium (2-sulfonatoethyl) methanethiosulfonate
NAD/H	Nicotinamide adenine dinucleotide
NAG-NAM	N-acetylglucosamine and N-acetylmuramic acid
NBD	Nitrobenzofurazan
OD	Optical desnsity
OG	Octylglucoside
OL	Oleic acid
OM	Outer membrane
OMP	Outer membrane protein
OMV	Outer membrane vesicles
PA	Phosphatidic acid
PAE	Predicted aligned error
PAGE	Polyacrylamide gel electrophoresis
PAP	Periplasmic adapter proteins
P_{ara}	P_{BAD} promoter
PBP	Penicillin binding proteins
PBS	Phosphate buffer saline
PC	Phosphatudylcholine
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PGN	Peptidoglycan
P_{lac}	P_{lac} promoter
plDDT	Predicted local distance difference test
POPG	1-palmitoyl-2-oleoyl-PG
PR	Proluciferin reductase
RND	Resistance-nodulation-division
SDS	sodium dodecyl sulphate
SNP	Single nucleotide polymorphisms
STI	Sexually transmitted infection
TBS	Tris-buffered saline
TLR	Toll-like receptor
TM	Transmembrane
URT	Upper respiratory tract
UTI	Urinary tract infection
WGS	Whole genome sequencing
WHO	World Health Organisation
XDR	Extensive-drug resistance

Section I: Introduction

100 years of antibiotics

The discovery of antibiotics to treat bacterial infections in the twentieth century was a major breakthrough in modern medicine. The first antibiotic to be used was Salvarsan in 1910 for the treatment of syphilis but is no longer in clinical use [2]. This was quickly followed by sulfonamides which inhibit upstream processes in DNA synthesis. These early antibiotics were both superseded by penicillin which became available as treatment in the 1940s. Penicillin is a natural compound produced by the fungi *Penicillium* that inhibits the bacterium's ability to synthesise their cell wall. Until the 1970s there was a golden era in antibiotic discovery during which the majority of currently used classes of antibiotics were discovered or developed [3]. Antibiotics not only prevent mortality due to pathogenic infections, but they enable a great number of other clinical procedures. For example, antibiotics are routinely given to patients undergoing surgery or childbirth. However, resistance to antibiotics occurred rapidly.

Antimicrobial resistance

Antimicrobial resistance (AMR) is the process by which microorganisms become resistant to antimicrobial chemotherapies. There are several mechanisms through which resistance occurs; natural barriers such as the cell envelope reduce permeability to antibiotics; efflux systems remove antibiotics from the cell; modifications to the target by genetic or non-genetic means; overexpression of a target or functionally similar proteins; transfer of resistance genes between bacteria; or direct modification of the antibiotic by the bacteria (Figure 1) [4]. Once the organism is resistant to a particular antibiotic, the antibiotic is no longer effective as a treatment and alternative clinical approaches are necessary. However, due to the emergence of multi-drug resistance (MDR) and extensive-drug resistance (XDR) alternative treatments are not always available.

Antimicrobial resistance occurs more readily when good clinical practice is not followed. Over prescription of antibiotics when they are not strictly necessary increases exposure of pathogens to antibiotics, increasing the opportunity to develop resistance. This is true also for being prescribed antibiotics for too long a course. Although commonly stated that a course of antibiotics should be completed to reduce resistance, increasing discussion around this concept suggests that short course can be as effective as long course, thereby reducing the exposure of antibiotics to the pathogen unnecessarily [5, 6].

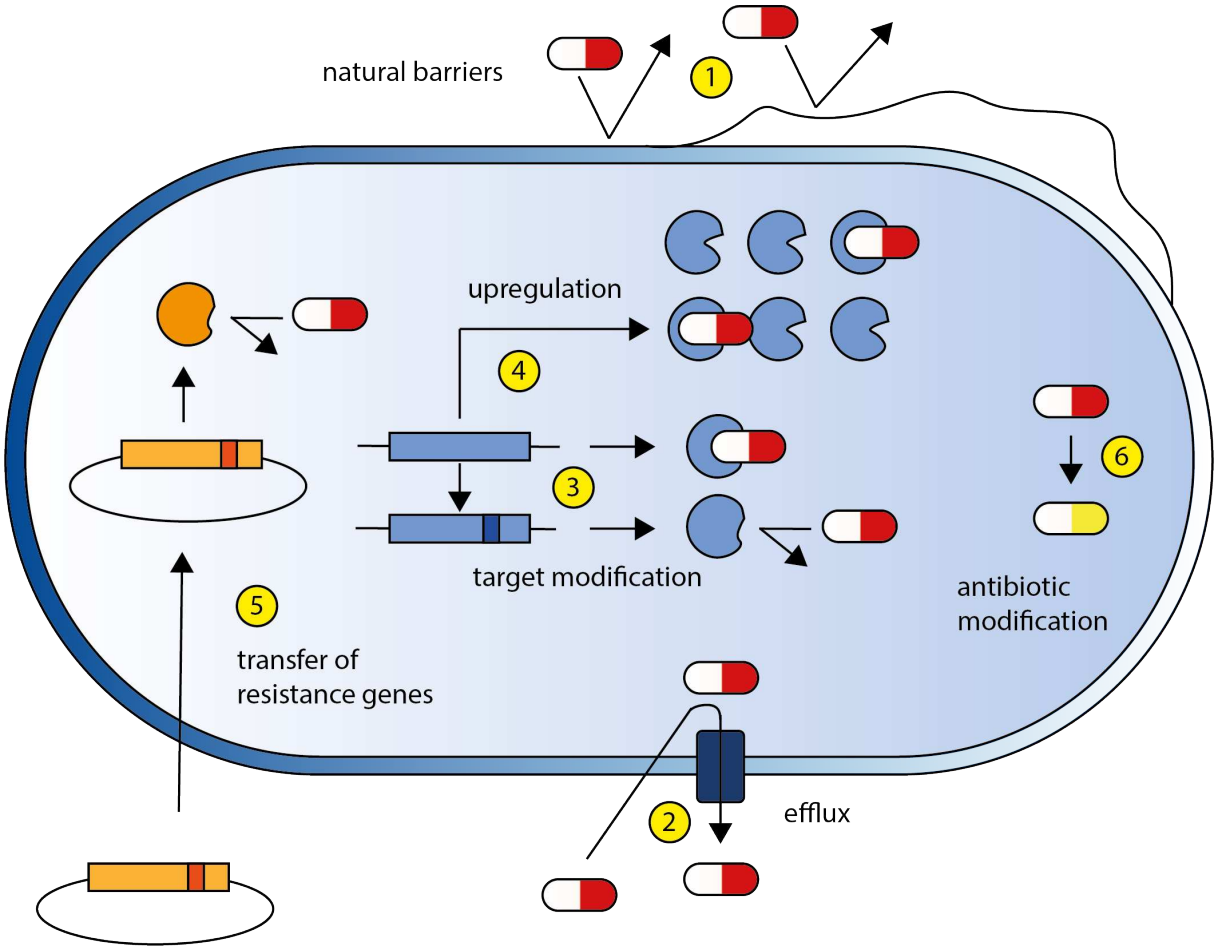


Figure 1. Mechanisms of AMR. *There are multiple ways that bacteria resist antibiotics 1) natural barriers; 2) efflux of antibiotics from the cell reducing the concentration below inhibitory levels; 3) target modifications by genetic and non-genetic means; 4) overexpression of a target or functionally similar protein; 5) transfer of resistance genes between bacteria; 6) direct modification of the antibiotic by the bacteria.*

Clinical prevalence of AMR

Projections on the future impact of AMR have estimated that deaths could reach around 10 million per annum by 2050 [7]. At the time of this report (2016) estimated annual deaths from AMR were reported to be 700,000. However, the latest estimates of global deaths associated with AMR infections in 2019 was just under 5 million of which roughly a quarter were directly attributable to AMR. To be classified as directly attributable to AMR the patient would have survived if the pathogen was not resistant to the treatments used [8]. This places direct AMR related deaths to be greater than HIV or malaria. AMR infections have a profound financial impact too with estimated costs of $40 billion per annum which could have a cumulative cost of $100 trillion by 2050 [7] . The urgent need to tackle AMR is increasingly evident.

Global initiatives to tackle AMR

To hone the efforts of scientists, clinicians and industrial partners to where the need is most urgent, the World Health Organisation (WHO) [9] has published a list of priority pathogens (Table 1). Prior to this advice from the WHO, Rice (2008) published a list of pathogens which pose a particularly problem due to AMR both in the developed and developing world. These organisms are collectively referred to as ESKAPE pathogens (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa and Enterobacter species*) and concur with the WHO assessment with the addition of *Klebsiella pneumoniae* [10]. A recent report from the Center for Disease Control (CDC) in the USA list AMR pathogens in order of the seriousness of the threat posed. Much of this list concurs with the ESKAPE and WHO Priority Pathogens with the addition of *Clostridiodes difficile* [11]. These lists have been widely adopted by scientists as a way of targeting those pathogens which pose the greatest problems. Table 1 provides an overview of these pathogen lists with a ranking related to the number of deaths caused in 2019 [8]. The list of bacterial pathogens can subdivide based on their membrane composition. Sixteen of the twenty-two bacteria have cell envelopes composed of two membranes (diderm) as opposed to a single lipid membrane (monoderm) (Figure 4). As the majority of Priority Pathogens and all ESKAPE pathogens are diderm, strategies to tackle this group of pathogens are of particular importance. As described later, the diderm cell envelope poses particular challenges to inhibitors but also provides opportunities for inhibition.

Table 1. Key AMR pathogens.

Pathogen	Priority Pathogens[a]	ESKAPE[b]	Threat Report[c]	AMR deaths (rank) [d]	Envelope
Acinetobacter baumannii	Critical	Yes	Urgent	5	Diderm
Campylobacter jejuni	High		Serious		Diderm
Citrobacter freundii				17	Diderm
Clostridioides difficile			Urgent		Monoderm
Enterococcus faecalis			Serious	12	Diderm
Enterococcus faecium	High	Yes	Serious	8	Monoderm
Escherichia coli*	Critical	Yes	Urgent	1	Diderm
Haemophilus influenzae	Medium			18	Diderm
Helicobacter pylori	High				Diderm
Klebsiella pneumoniae		Yes		3	Diderm
Morganella morganii				22	Diderm
Mycobacterium tuberculosis	Critical		Serious	7	Diderm
Neisseria gonorrhoeae	High		Urgent		Diderm
Proteus mirabilis				13	Diderm
Pseudomonas aeruginosa	Critical	Yes	Serious	6	Diderm
Salmonella *spp.*	High		Serious	11,21	Diderm
Serratia marcescens				15	Diderm
Shigella flexneri	Medium		Serious	19	Diderm
Staphylococcus aureus	High	Yes	Serious	2	Monoderm
Streptococcus agalactiae			Concerning	10	Monoderm
Streptococcus pneumoniae	Medium		Serious	4	Monoderm
Streptococcus pyogenes			Concerning	16	Monoderm

*Pathogens present in various reviews into AMR and their stated level of urgency. *Some reviews refer to E. coli specifically, others to a wider Enterobacteriaceae group. [a][9], [b][10], [c][11], [d][8] – AMR deaths (rank) refers to the number of deaths associated with AMR.*

Approaches to tackle AMR

Alongside targeting specific organisms of interest as described above, various other initiatives provide a more holistic strategy. The most notable example of this is the WHO Global Action Plan for Antimicrobial Resistance [12]. There are five key points to the plan. The first is to increase awareness of AMR at both government and population levels. The second is to increase our knowledge of AMR at global and local levels by improving surveillance programs. A notable advance in this area is the recent global study of deaths caused by AMR infections [8] which provided the most comprehensive global review to date. Thirdly, prevention of infection, thereby reducing the need for antibiotics by reducing the risk of acquiring an infection. Fourthly, improved stewardship of the antibiotics we have currently in use. As over prescription and incorrect prescription of antibiotics is prevalent, a simple strategy to improve their prescription by physicians, combined with public awareness and treatment expectations could reduce AMR infections. Finally, an economic plan for increased global investment in new medicines, whether antibiotics or vaccines alongside improved diagnostics will hopefully increase the arsenal for physicians treating patients with bacterial infections.

Vaccines and Non-antibiotic therapies

The role of vaccines and other non-antibiotic treatments in combatting AMR has been increasingly discussed. The advantage of vaccines are three-fold [13]. Firstly, vaccines designed against AMR priority pathogens will reduce the prevalence of disease by priming the immune response ready for infection. Secondly, vaccines against viral pathogens will reduce the prevalence of viral infections, which in turn lead to secondary infections from opportunistic bacterial pathogens requiring treatment. It also reduces unnecessary prescriptions of antibiotics for viral infections [14]. For example, one study revealed three quarters of COVID-19 patients received antibiotics when only 8% had bacterial infections [15]. In Canada there was a 64% reduction in antibiotic prescriptions related to influenzae after a universal influenza vaccine had been administered [16]. Finally, vaccines against animals, particularly livestock pathogens would greatly reduce the over-use of antibiotics in the agricultural sector, which is the largest consumer of antibiotics [17].

There are different types of vaccines or non-antibiotic therapies that could be and are being used. For example, monoclonal antibody therapy and bacteriophages provide highly specific options for therapy. Treatments using the host microbiome to tackle infections have also been proposed. These are based on natural principles of infection prevention or removal and have shown promise. Vaccines are currently used against bacterial infections such as *H. influenzae*, *Meningococcus* and *Salmonella spp.* and have shown to be highly effective at preventing serious disease and generating herd immunity

[18]. Many vaccines have multiple epitopes and therefore single mutations are unlikely to lead to reduced efficacy. However, some incidents of vaccine-resistant infections have been reported. For example, a vaccine against the Hepatitis B virus, targets a specific epitope but single amino-acid changes in this epitope leads to resistance to the vaccine. It is also speculated that an increase in infections of *B. pertussis* could be due to resistance to the whooping cough vaccine [18].

The preventative nature of vaccines means there will be an inherent reduction in infections and therefore the need for treatments. Reducing the use of antibiotics is a keyway to prevent the development of resistance. There are vaccines under development for nearly all the key pathogens listed in the Table 1 [14, 18]. Of those developed, the *H. influenzae* vaccine has been shown to reduce the prevalence of AMR infections and pneumococcal vaccines were shown to reduce penicillin non-susceptible infections by up to 87% [14].

Antibiotics

As antibiotics have shown to be effective, they are still a popular strategy to tackle AMR infections. However, there are considerable challenges facing their development. Most notably is the lack of input from large pharmaceutical companies due to poor incentivization. However, new antibiotics are still being developed.

Current antibiotics target a select few cellular processes that are essential to bacterial physiology. Targeting cell envelope, DNA and protein synthesis machineries account for nearly all approved drugs (Figure 2) [3]. Although there is a steady but slow stream of new antibiotics becoming available to clinicians, frequently, newly approved antibiotics are derivatives of current classes of antibiotics. When they are not of the same class, they do not have novel targets. For example, since 2017 eleven new antibiotics received approval in the USA. Of these eleven, nine were derivatives of current classes of antibiotics where resistance is expected to occur rapidly. Two new chemical classes were approved: Vaborbactam is a beta-lactamase inhibitor and Lefamulin is a 50S ribosome inhibitor [19]. Thus, although they are new chemical classes, they do not have novel targets.

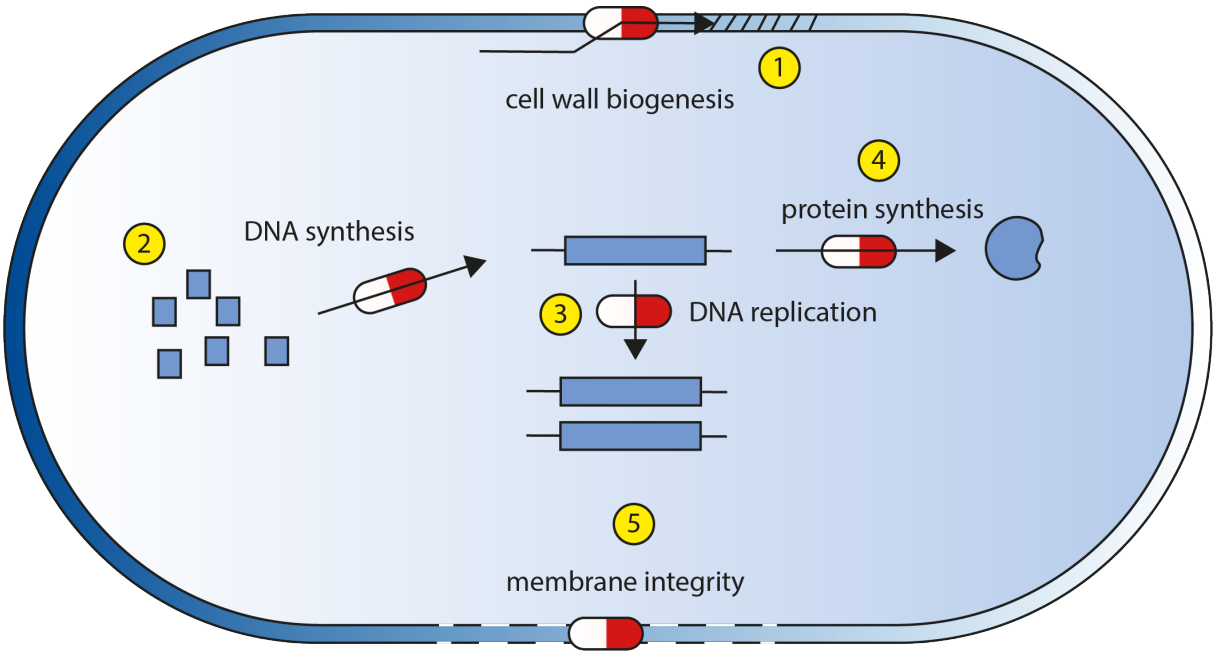


Figure 2. Common targets of antibiotics. *The majority of antibiotics target the protein, DNA or cell wall synthesis machineries. 1) inhibition of cell wall synthesis machinary (e.g. beta-lactams). 2) prevention of DNA synthesis (e.g. sulfonamides). 3) prevention of DNA replication (e.g. fluroquinelones). 4) inhibition of protein synthesis (e.g. tetracyclines). 5) disruption of the cell membrane (e.g. polymyxins).*

Two approaches to antibiotic discovery

There is a strong case to identify potential antibiotic targets in diderm pathogenic bacteria and find inhibitors for them. There are two basic approaches to antibiotic discovery: phenotypic screens and target-based screens (Figure 3). Phenotypic screening typically involves a library of potential inhibitors added to the growth medium of a particular bacteria and bacteriostasis or bacteriocidal is analysed. If a 'hit' is found then the target is identified and the inhibitor may be chemically modified to improve its activity against the bacteria. The advantage of this method is from the outset you know if the compound works on bacterial cell cultures. However, the target is unknown and may be a common target of antibiotics already.

There are many methods to identify a target of an inhibitor discovered in phenotypic screens. Commonly, resistance studies reveal possible target sites. In this process, the bacteria are grown in high levels of antibiotic, above the minimum inhibitory concentration (MIC) and resistant colonies are collected. These resistant strains have their genome sequenced and genes which have mutations are explored further as being possible targets. However, this does not demonstrate beyond doubt the target is the product of the gene that is mutated. Phenotyping studies can also be conducted to narrow down possible targets. For example, microscopy may reveal changes in shape which correlate to reduced activity of cell elongation or division machinery.

Once you have a hypothesis of possible targets, more direct approaches can be taken. Possible targets can be artificially overexpressed in the presence the inhibitor, if the MIC is increased under these conditions, it indicates the protein may be involved in the pathway. If there is a known phenotype related to the hypothesised target, directed analysis could be conducted. For example, if you believe a component of lipoprotein maturation is affected, analysis of the lipidation profile of known lipoproteins would indicate the pathway was affected by the inhibitor. Comparisons of the phenotype of the cell with depletion or deletion strains also indicate that the inhibitor is targeting the specific hypothesised target.

In vitro enzymatic activity assays could also be developed, and the inhibitory effect of the inhibitor compared to negative and positive controls, active and non-active protein, for example. However, this requires an *in vitro* assay to be developed and this can be a complicated process depending on the function and characteristics of the target. Finally, structural determination of the target with the inhibitor may provide sufficient evidence to conclude that the hypothesized target is indeed the target. This can also be explored via *in silico* docking. This method uses predicted or experimentally determined structures and computational models to test possible inhibitors of the active site. It can

therefore also be applied to screen for possible inhibitors. In this case large screening libraries are therefore required less as more targeted molecules can be selected or developed after computational screens.

The second approach, target-based screening, is contrary to phenotypic screening. A target is identified as having potential and an *in vitro* screening assay is developed. If a 'hit' is found it is chemically modified to increase its efficacy on bacterial growth. The advantage of this method is that the target is predetermined and molecules that originally may not be picked-up by phenotypic screening – due to chemical properties preventing entry into the cell, for example – are identified. However, an appropriate high-throughput assay needs to be developed and this is not always straight forward and lead optimization is far from simple. A factor that limits both approaches is the availability of diverse molecule libraries to screen against [20, 21].

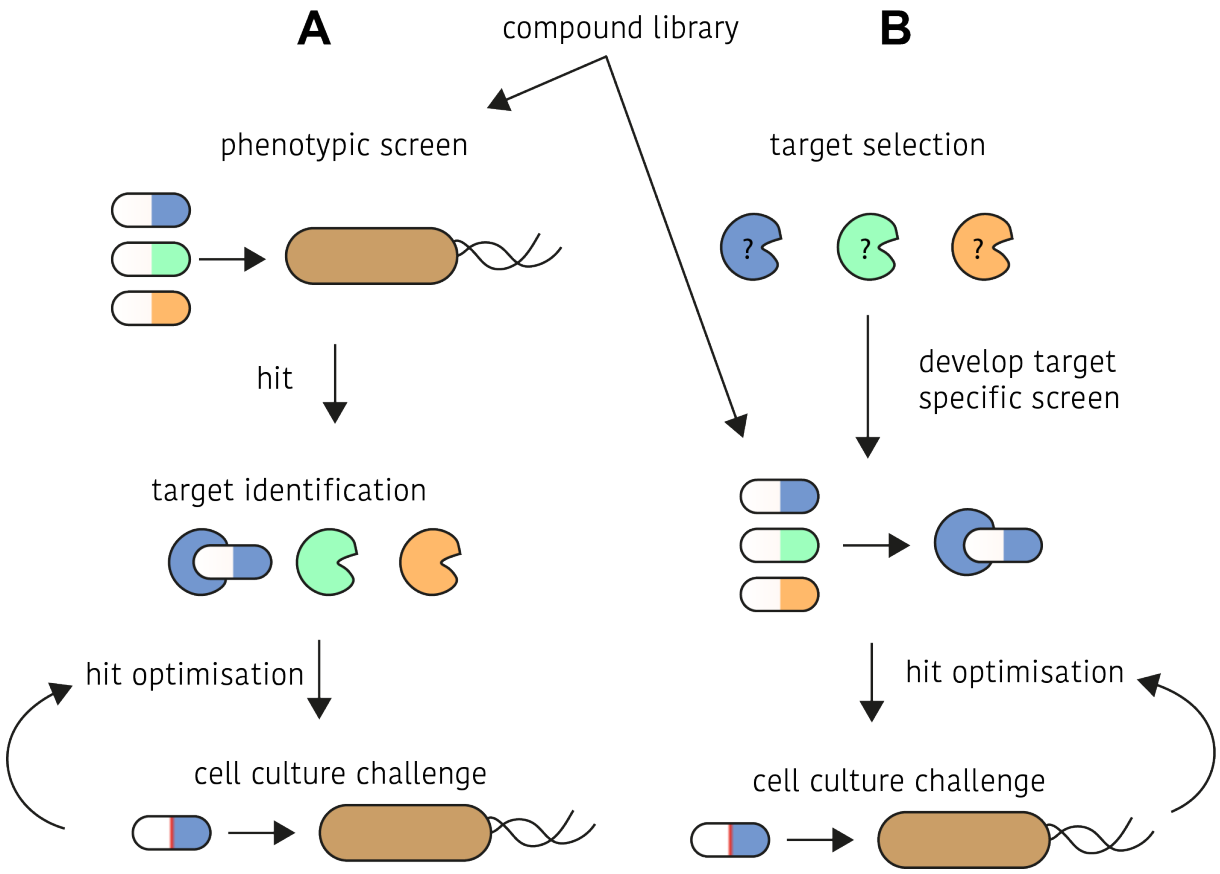


Figure 3. Methods of antibiotic discovery. *A) Phenotypic screening for inhibitors where a bacteriostatic or bactericidal screen is prepared and molecules are assessed relative to controls. B) Target based screens require an in vitro assay before hit modification to improve its activity against live cells.*

Selecting a target

Selecting a target for target-based screening has its challenges. Targeting essential proteins is believed to lead to higher rates of resistance due to high selective pressure of the pathogen and this is observed in many single-target antibiotics. Resistance is reduced if the antibiotic has multiple targets leading to the hypothesis that the most effective antibiotics are multi-target or target pathways [21, 22].

Inhibitors of penicillin-binding proteins (PBPs), β-lactams for example, such as penicillin, target multiple PBPs in the cells and therefore, mutations in a single PBP would not be sufficient to confer resistance. It is believed that a minimum of two PBPs are required for survival [22]. Fluoroquinolones target DNA Gyrase and Topoisomerase IV and Daptomycin and polymyxins target the cell membrane. Each of these multitarget inhibitors see low levels of single step resistance *in vitro* [21].

An example of when this hypothesis contradicts clinical data is Fosfomycin resistance. Fosfomycin inhibits cell wall synthesis enzyme MurA, and therefore is a single target antibiotic. Resistance is high under laboratory conditions but low in a clinical setting. It is hypothesized that the fitness cost due to the resistance mutation is sufficient to enable the host to counter the infection [21].

Essential genes are those which are deemed essential for survival for an organism under defined conditions. This is tested experimentally by the ability to delete, disrupt or silence genes under the chosen conditions. For example, three methods have been employed to study the global essentiality of *E. coli* genes under rich media, laboratory conditions. The first, known as the Keio collection, systematically replaced each known gene with an antibiotic resistance cassette. If a gene could not be replaced, it was deemed essential [23]. However, this does not take into account finer details such as whether or not particular genetic regions encoding specific domains within a particular gene are essential, as opposed to the whole gene. Goodall *et al.* (2018) used a Transposon directed insertion sequencing (TnSeq) library to randomly insert a transposon across the genome [24]. *E. coli* was then grown and the genome sequenced, regions of the genome which did not house the transposon were deemed essential. Finally, a CRISPRi gene silencing library was conducted [25]. Here, inhibition of initiation or elongation of transcription is achieved by repurposing CRISPR systems. The catalytic inactive form of Cas9, dCas9, forms together with a guide RNA, a DNA recognition complex that binds to the target but does not cleave DNA. This leads to inhibition of transcription otherwise known as gene silencing. In each method, with its different techniques, different lists of essential genes are produced. Although there is a considerable overlap, there is not complete agreement. More and more studies have been carried out in different organisms [26] which help to enhance our knowledge of

essential genes. In fact, some prediction tools have been produced for determining gene essentiality [27].

The conditions under which genes are essential is clinically relevant. A gene may be essential under laboratory conditions but not inside host cells. Some genes are therefore conditionally essential due to environment or other genetic factors. For example, lipoproteins LpoA and LpoB of *E. coli* are not deemed essential in the studies discussed previously but have been shown to be essential in the absence of the other [28]. They play a key role in activating peptidoglycan (PGN) synthesis through the stimulation of the enzymatic activity of PBPs, which is an essential process.

This adds to the complexity of antibiotic discovery. Often phenotypic studies are conducted in rich medium conditions, therefore, increasing the chance of lead compounds that may not be effective in a host environment.

Targeting non-essential proteins has also been increasingly discussed. Inhibiting processes like cell adhesion, efflux or virulence factors could attenuate infection. Targeting essential proteins creates high selective pressure for resistance whereas non-essential targets would induce only a mild selective pressure [29].

The basic approaches to developing new antibiotics are: 1) single multi-target compounds, 2) hybrids of two compounds, 3) combination therapy. Multi-target and non-essential targets give credence to the idea of combination therapies. This approach is common in the treatment of TB, for example. This method describes the use of a combination of antibiotics to treat a single infection. These can be congruous (the two drugs each have growth inhibition effects), syncretic (one of the drugs leads to growth inhibition but the other does not) or coalistic (neither drug alone has growth inhibitory effects) [30]. The ultimate goal is to increase the effectiveness and reduce the generation of resistance to antibiotics.

Targeting the bacterial cell envelope

As described, the bacterial cell envelope is the targets of many antibiotics. Alongside compounds which disrupt the membrane, many target the synthesis of the peptidoglycan. However, one pathway in cell envelope biogenesis and maintenance – lipid modification of proteins – is yet to be targeted by a single antibiotic. The lipoprotein modification pathway (LMP) in diderm bacteria contains multiple essential enzymes and the products of the reaction have roles in many other pathways, including those related to reduced susceptibility to antibiotics, and therefore may be a potential target for novel antibiotics.

Due for the urgent need to develop new antibiotics against Gram-negative bacteria, this thesis will focus mainly on this group of bacteria.

The Gram-negative cell envelope

Historically, bacteria have been classified by the colouring stains used to identify them. The most notable example is the Gram stain, developed by Danish bacteriologist Hans Christian Gram in 1884, which separates bacteria stained by the process into two groups: Gram-positive (often from the phylum firmicutes) and Gram-negative (often proteobacteria). Gram-positive bacteria are generally monoderm and Gram-negatives are generally diderm bacteria. This is true for the priority pathogens (Table 1) with the exception of *Mycobacterium tuberculosis* which is not stained by the Gram stain but is diderm (actinobacteria), albeit with a structurally different outer membrane to the other diderms listed (Figure 4). Except for *Mycobacterium tuberculosis* the diderm priority pathogens fall into the taxonomic group of proteobacteria.

The basic structure of the Gram-negative cell envelope forms a three-layered barrier protecting the cytoplasmic space from the extracellular milieu. There are inner and outer phospholipid membranes separated by the aqueous periplasm which houses the peptidoglycan (PGN) often refered to as the cell wall (Figure 4).

Both membranes contain a variety of proteins that are present in three main forms: either associated to the membrane via charged domains, embedded integrally via transmembrane alpha-helix or β-barrel domains, or anchored into the membrane via a lipid domain. Figure 4 provides an overview of the basic structures of the cell envelope in Gram-negative bacteria and highlights essential pathways.

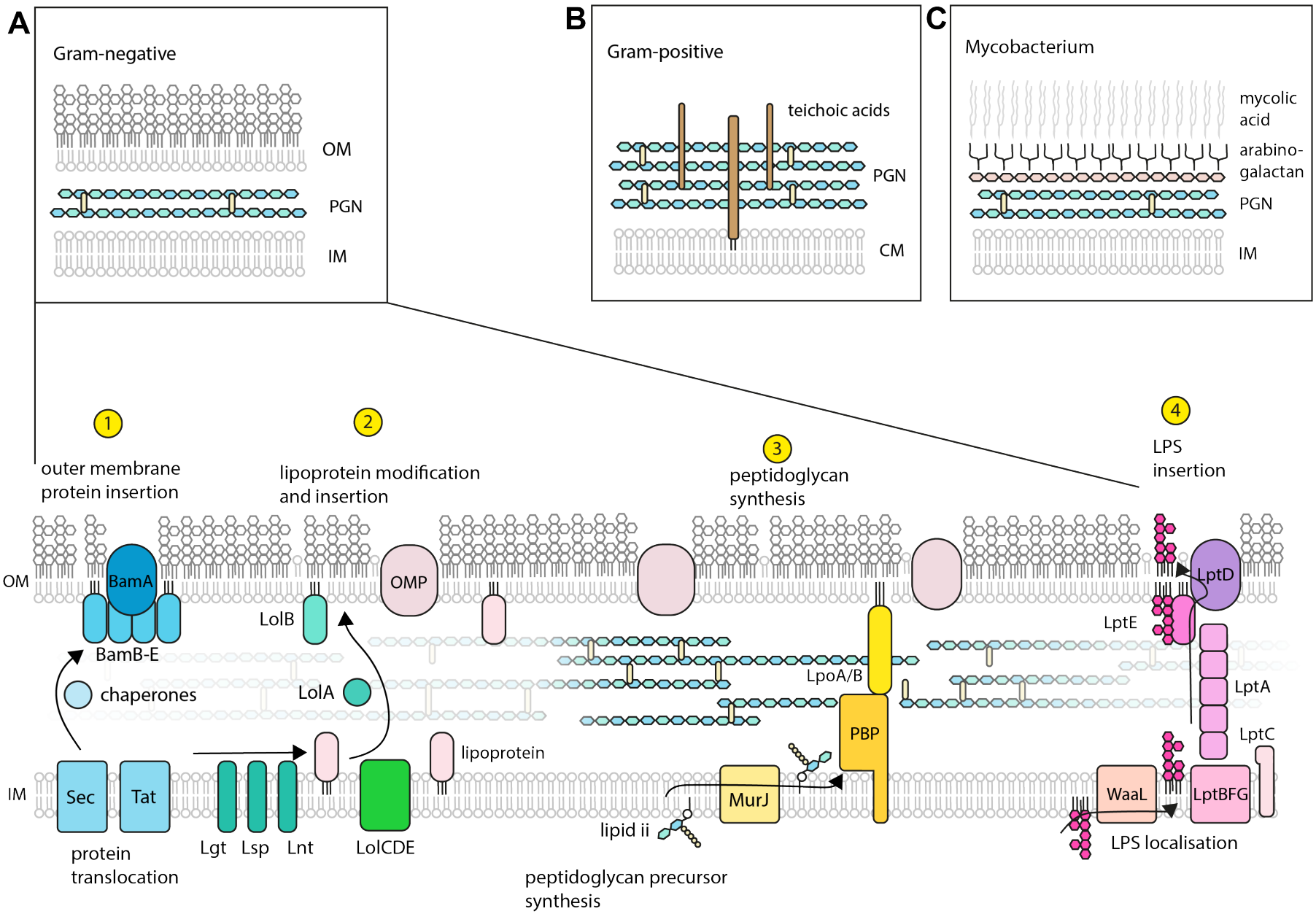

A
Gram-negative
OM
PGN
IM
B
Gram-positive
teichoic acids
PGN
CM
C
Mycobacterium
mycolic acid
arabino-galactan
PGN
IM
1
outer membrane protein insertion
2
lipoprotein modification and insertion
3
peptidoglycan synthesis
4
LPS insertion
OM
BamA
BamB-E
chaperones
LolB
LolA
OMP
lipoprotein
LpoA/B
PBP
LptD
LptE
LptA
LptC
IM
Sec
Tat
protein translocation
Lgt
Lsp
Lnt
LolCDE
MurJ
lipid ii
peptidoglycan precursor synthesis
WaaL
LptBFG
LPS localisation

Figure 4. The bacterial Cell envelope. *The cell envelope varies according to the bacteria. A) Schematic of Gram-negative (E. coli) cell wall. 1) many processes begin with proteins secreted through the Sec or Tat pathways. OMP insertion requires chaperone proteins to direct peptides to the outer membrane components of the pathway which in turn insert them into the outer membrane. 2) lipoprotein modification by Lgt, Lsp and Lnt and localization by the Lol machinery. 3) PGN precursor synthesis occurs in the cytoplasm, but later steps are performed at the membrane with precursors attached to lipid II. Lipid II is flipped across the inner membrane and the NAM-NAG-peptide subunits are synthesized into the new PGN by PBPs. 4) LPS is transported from the inner membrane to the outer membrane by the Lpt pathway. B) overview of cell envelopes from Gram-positive bacteria and Mycobaceria. OM = outer membrane, IM/CM = inner/cytoplasmic membrane, PGN = peptidoglycan, LPS = lipopolysaccharide, PBP = penicillin binding protein, OMP = outer membrane protein.*

Phospholipids

The two membranes of Gram-negative bacteria differ from each other with the outer membrane forming an asymmetric bilayer with phospholipids predominantly on the inner leaflet and lipopolysaccharide (LPS) on the outer leaflet. Whereas the inner membrane is a symmetrical phospholipid bilayer.

Phospholipid biosynthesis is an important pathway in bacteria. Although some bacteria have been shown to utilize phospholipids from their environment, most bacterial species have mechanisms to synthesis their own *de novo*.

The phospholipid composition varies between species and at different stages of the cell cycle (Table 2). In proteobacteria phosphatidylethanolamine (PE) and phosphatidylglycerol (PG) are invariably present with cardiolipin (CL) present in all but a few exceptions [31] (Table 2). The head groups of these phospholipids have different properties. PE for example has no net charge whereas PG and CL have single and double negative charges, respectively (Figure 5). CL is composed of two PG molecules and is believed to allow greater curvature in the membrane [32]. Both PE and PG are used as substrates by the LMP at different steps in the synthesis pathway. PG is utilized in the first step by Lgt, present in all bacteria and PE in the final step by Lnt, present in Gram-negative.

Table 2. Presence of phospholipids in notable bacteria.

Pathogen	PE	PG	CL	Other
Campylobacter jejuni	X	X		
Enterobacteriaceae	X	X	X	
Helicobacter pylori	X	X	X	Lyso-PE
Mycobacterium tuberculosis	X	X	X	PI, PIM, OL
Neisseria gonorrhoeae	X	X	X	Lyso-PE
Pseudomonas aeruginosa	X	X	X	PC, OL, APG
Salmonella spp.	X	X	X	Acyl-PG
Staphylococcus aureus		X	X	LPG, GPL
Streptococcus pneumoniae		X	X	GL

(PE = phosphatidyl-ethanolamine, PG = phosphatidyl-glycerol, CL = cardiolipin, PC = phosphatidyl-choline, OL = oleic acid, Adapted from [31]

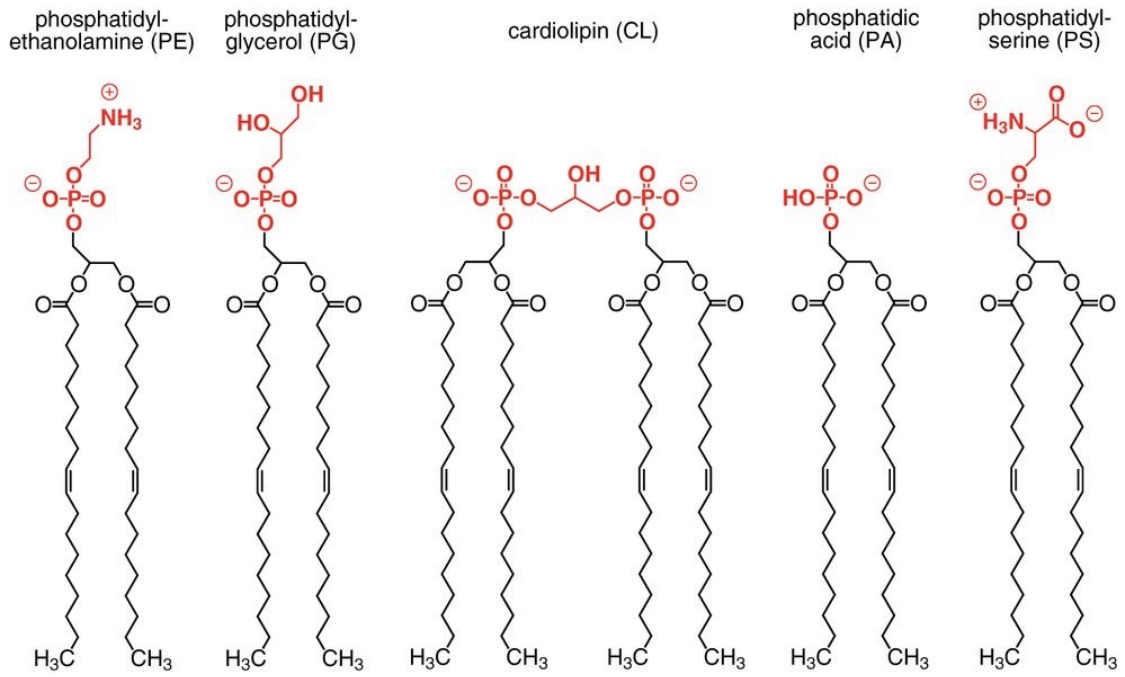


***Figure 5. Common phospholipids**. Displayed here to demonstrate head-groups but the composition of the acyl chain varies. From [33]*

Inner membrane

The inner membrane is a symmetrical phospholipid bilayer. Many important cellular functions occur at this site. For example, multiple steps of peptidoglycan synthesis are performed by integral inner membrane proteins and the precursor itself contains a lipid group (Lipid II) maintaining its presence in the membrane (Figure 4). In fact, PBPs which insert the PGN precursors into the nascent PGN are either membrane associated or integrally bound to the inner membrane via an alpha-helix domain [34]. Electron transport chain also occurs in bacteria at the inner membrane by a number of enzymes either integral, associated to the membrane or as lipoproteins and is the source of energy in the form of ATP for the cell [35]. The inner membrane also houses stress response proteins that signal if there are perturbations in the cell envelope [36]. For example, accumulation of the outer membrane lipoprotein NlpE in the inner membrane is thought to activate the Cpx stress response system [37] triggering multiple downstream effectors.

The inner membrane is also the site of lipoprotein modification. This process (described in detail below) is the process by which proteins become modified with fatty acids which enable them to be anchored into the phospholipid bilayer. The pathway is performed by three integral inner membrane enzymes, Lgt, Lsp and Lnt [38] (Figure 4).

Due to the hydrophobicity of the membrane, proteins cannot simply diffuse from the cytoplasm to the periplasm. There are several ways proteins can cross this membrane to gain access to the periplasm (Figure 4). The secretion or Sec translocon transports unfolded proteins across the membrane and inserts them into the membrane where they may be cleaved and carried across the periplasm by chaperones or remain in the membrane to reside or be further processed, by lipoprotein modification, for example [39, 40]. The twin-arginine secretion or Tat translocon can transport folded proteins across the IM [39] (Figure 4). Further to this, so-called flippases have been reported to carry lipid-attached molecules across the membrane such as cell wall precursors attached to the carrier lipid Lipid II [41]. Recycling of Lipid II by flipping of undecaprenyl pyrophosphate back to the inner leaflet of the inner membrane is performed by UppP in *E. coli* [42], and lysophospholipid flipping from the outer leaflet to the inner leaflet of the inner membrane to be recycled to phospholipids is performed LplT [43]. These all provide ways for important cell components to access the periplasm and to return to the cytoplasm.

Periplasm and the cell wall

The periplasm is an aqueous space containing soluble proteins, periplasmic facing membrane bound or associated proteins, lipoproteins and the peptidoglycan cell wall (Figure 4).

Peptidoglycan precursors are prepared in the cytoplasm and flipped across in lipid-anchored blocks of N-acetylglucosamine-N-acetylmuramic acid (NAG-NAM) with a pentapeptide stem and are incorporated into existing PGN by PBPs. This creates a continuous macromolecular mesh-like structure providing stability to the cell and protecting against internal osmotic pressure. PGN is a dynamic structure with a multitude of enzymes involved in its synthesis, restructuring and degradation [44]. Peptidoglycan biosynthesis is targeted by the majority of antibiotics targeting the bacterial cell envelope. Beta-lactams such as penicillin are competitive inhibitors of PBPs as is the glycopeptide vancomycin [45]. From the most recent analysis of global AMR burden, 70% of deaths attributable to AMR were from pathogens resistant to antibiotics which target the bacterial cell wall machinery [8].

Due to the aqueous nature of the periplasm, there are a number of ways to transport hydrophobic or lipid anchored proteins across the gulf. Lipopolysaccharide (LPS) is an essential component of the outer membrane, but a hydrophobic lipid group prevents its solubilization and is therefore transported by the Lpt system and inserted into the outer membrane by the lipoprotein LptE and outer membrane pore LptD [46] (Figure 4). Outer membrane proteins (OMPs) are translocated across via the SurA, Skp or DegP chaperones [47] and inserted into the membrane by the Bam complex composed of β-barrel protein BamA and four associated lipoproteins BamB-E [48] [49] (Figure 4).

Lipoproteins use the lipoprotein outer membrane localization (Lol) transport system where they are selected for and extruded from the inner membrane by the ABC transporter LolCDE complex [50] to LolA, a chaperone protein. They are inserted into the outer membrane by the lipoprotein LolB although it is hypothesized that there may be an alternative mechanism by which lipoproteins are transported [51]. There are variations of LolCE where a single protein, LolF, is present in non-γ-proteobacteria [52]. Lipoproteins contain a '+2' signal which dictates whether they remain in the inner membrane or are translocated via Lol to the outer membrane. The +2 refers to the amino acid following a fatty-acid modified cysteine residue (C_{+1}) in mature lipoproteins. In *E. coli*, an aspartate residue determines IM localization and any other residue allows for export to the OM [53]. This rule applies to other *Enterobacteriacae* but not more widely in proteobacteria, such as *Pseudomonas* where +3 and +4 residues play a role [52, 54, 55].

Although essential processes, OMP insertion, LPS biosynthesis and surface exposure and lipoprotein modification are not targeted by approved antibiotics.

Outer membrane

Unlike the inner membrane, the outer membrane is asymmetrical. The phospholipid inner leaflet, facing the periplasm, differs greatly from the outer leaflet composed of lipopolysaccharide. LPS is essential in proteobacteria. It is composed of Lipid A, a core polysaccharide section and an O antigen which varies greatly between bacteria [56]. LPS provides stability as well as a permeability barrier to the extracellular milieu and is a key signaling factor to the host immune system via the Toll-like receptor, TLR-4 [57].

The outer membrane houses integral membrane proteins (OMPs) and lipoproteins, some of which are displayed on the cell surface [58]. In *Neisseria*, the surface lipoprotein assembly modulator (SLAM) was discovered to have a role in exposing lipoproteins to the cell surface [59]. Chaperone Skp was shown to have a role in this process suggesting lipoproteins may also chaperoned [60]. As well as being the location of LptE, LolB and BamB-E lipoproteins for essential components of the OM, the outer membrane provides channels for entry of nutrients and exit of small molecules and toxins [61].

In *E. coli*, the most abundant protein is the OM lipoprotein Lpp (Braun's lipoprotein) with an estimated copy number of over 1 million per cell. It primarily plays a structural role [62] where its lipid group is anchored into the inner leaflet of the outer membrane and its coiled structure protrudes into the periplasm as a homotrimer. It is covalently linked to PGN via its C-terminal lysine group by L,D-transpeptidase enzymes [63]. The absence of Lpp greatly reduces the susceptibility of bacteria to known inhibitors of the LMP [64-67]. Lpp will be discussed in greater detail below.

Targeting the cell envelope

There are a number of antibiotics which target the cell envelope. The antibiotic Bacitracin prevents the dephosphorylation of undecaprenylpyrophosphate (C55-PP), a PGN synthesis component found in the inner membrane and the substrate for lipid II biosynthesis [68], and therefore prevents PGN synthesis, leading to cell death. Beta-lactam antibiotics inhibit PBPs and prevent the synthesis of PGN. Vancomycin binds to the D-Ala-D-Ala on the pentapeptide stem preventing access to the PBPs and therefore also preventing PGN synthesis [69]. Polymyxins are notable antibiotics that target the OM by binding to LPS and create pores in the outer membrane leading to cell death [70].

There are multiple antibiotics that target various aspects of the cell envelope in each of the three main compartments (IM, periplasm, OM). Targeting the cell envelope is a proven method of antibiotic

development. It also has the benefit of being a differentiating factor between monoderm and diderm bacteria, possibly reducing the spectrum of activity.

Lipoproteins have a role in two essential cell envelope processes in OMP and LPS insertion which makes it an essential pathway. Alongside these essential processes, lipoproteins have a range of important functions in cellular processes (see below) making them, or their modification pathway, an attractive target for novel antibiotics. There are currently no clinically used in inhibitors of the LMP. Globomycin and Myxovirescin are well discussed inhibitors of Lsp [71, 72], alongside the recently described G0790 [73]. The Lol pathway has been shown to be inhibited by G0507 [74]and two other small molecule inhibitors [64]. Finally, a cyclic peptide inhibitor of Lgt has recently be described in the literature [75]. However, to date, none have been adapted for clinical use.

Lipoprotein biogenesis

As described above, the Gram-negative (proteobacterial) cell envelope contains a number of lipoproteins including those involved in essential cellular processes. Lipoproteins are characterized by fatty-acid acylation of an invariable cysteine residue which anchors the protein into the membrane.

We described the pathway by which lipoproteins are post-translationally modified in our 2021 review article:

Mode of action of lipoprotein modification enzymes - Novel antibacterial targets

Legood, S., G. Boneca, I., Buddelmeijer, N. (2021)

Molecular Microbiology 115(3): 356-365.

In this article we discuss the main stages of the classical lipoprotein modification pathway as found in proteobacteria. Namely, the translocation of a peptide across the inner membrane via the Sec or Tat machineries, modification with a diacylglyceryl moiety derived from the phospholipid phosphatidylglycerol (DAG-ylation) by Lgt, peptide cleavage by signal peptidase Lsp and N-acylation by Lnt. Modification is followed by the translocation across the periplasm and insertion into the outer membrane by the Lol machinery.

The article focused on the structure and mechanisms by which the enzymes function, their essentiality in proteobacteria and accessible domains, and the development of high-throughput *in vitro* assays to screen for inhibitors.

Considering the work presented in my thesis, I have included after this publication additional information about the lipoprotein modification pathway and in particular on Lgt, the first enzyme in the pathway.

Mode of action of lipoprotein modification enzymes - Novel antibacterial targets

Received: 30 July 2020 | Revised: 10 September 2020

DOI: 10.1111/mmi.14610

MICROREVIEW

WILEY

Mode of action of lipoprotein modification enzymes—Novel antibacterial targets

Simon Legood[1,2,3,4] | Ivo G. Boneca[1,2,3] | Nienke Buddelmeijer[1,2,3]

[1]Institut Pasteur, Unité Biologie et Génétique de la Paroi Bactérienne, Paris, France

[2]CNRS, UMR 2001 « Microbiologie intégrative et Moléculaire », Paris, France

[3]INSERM Groupe Avenir, Paris, France

[4]Université de Paris, Sorbonne Paris Cité, Paris, France

Correspondence
Nienke Buddelmeijer, Institut Pasteur, Unité Biologie et Génétique de la Paroi Bactérienne, 25-28 rue du docteur Roux, 75724 Paris cedex 15, France.
Email: nienke.buddelmeijer@pasteur.fr

Funding information
Investissement d'Avenir program, Laboratoire d'Excellence "Integrative Biology of Emerging Infectious Diseases, Grant/Award Number: ANR-10-LABX-62-IBEID; Institut Carnot Infectious Diseases, Grant/Award Number: 16 CARN 0023-01

Abstract

Lipoproteins are characterized by a fatty acid moiety at their amino-terminus through which they are anchored into membranes. They fulfill a variety of essential functions in bacterial cells, such as cell wall maintenance, virulence, efflux of toxic elements including antibiotics, and uptake of nutrients. The posttranslational modification process of lipoproteins involves the sequential action of integral membrane enzymes and phospholipids as acyl donors. In recent years, the structures of the lipoprotein modification enzymes have been solved by X-ray crystallography leading to a greater insight into their function and the molecular mechanism of the reactions. The catalytic domains of the enzymes are exposed to the periplasm or external milieu and are readily accessible to small molecules. Since the lipoprotein modification pathway is essential in proteobacteria, it is a potential target for the development of novel antibiotics. In this review, we discuss recent literature on the structural characterization of the enzymes, and the in vitro activity assays compatible with high-throughput screening for inhibitors, with perspectives on the development of new antimicrobial agents.

KEYWORDS

diacylglyceryl transferase, in vitro activity assays, inhibitors, lipoprotein, N-acyl transferase, phospholipid, signal peptidase, X-ray crystal structure

1 | INTRODUCTION

Volkmar Braun first discovered bacterial lipoproteins in 1973 through the identification of a fatty-acid modification of Lpp, or Braun's lipoprotein, in *E. coli* (Hantke and Braun, 1973). Through early biochemical and genetics studies and more recent structural analysis, the lipoprotein modification pathway is increasingly well understood. A general consensus exists regarding the well-studied tripartite stages of the lipoprotein modification pathway. Upon insertion into the cytoplasmic membrane, a diacylglyceryl group is added to the lipoprotein, the membrane-spanning signal peptide is cleaved and the protein stays membrane anchored by its diacylglyceryl moiety. Finally, N-acylation results in the formation of mature triacylated lipoprotein (Figure 1). In diderm bacteria, including proteobacteria and some high GC content Gram-positive bacteria, including *Streptomyces*, *Corynebacteria*, and *Mycobacteria*, lipoproteins are triacylated following this classical pathway, although in some instances Lnt and/or Lsp are not essential components for cell viability (discussed below). In monoderm bacteria it was long thought that only diacylated lipoproteins existed; however, recent studies illustrate that alternative lipid modifications occur in firmicutes and mollicutes, but not all enzymes catalyzing these reactions have been identified (Armbruster and Meredith, 2017; Asanuma et al., 2011;

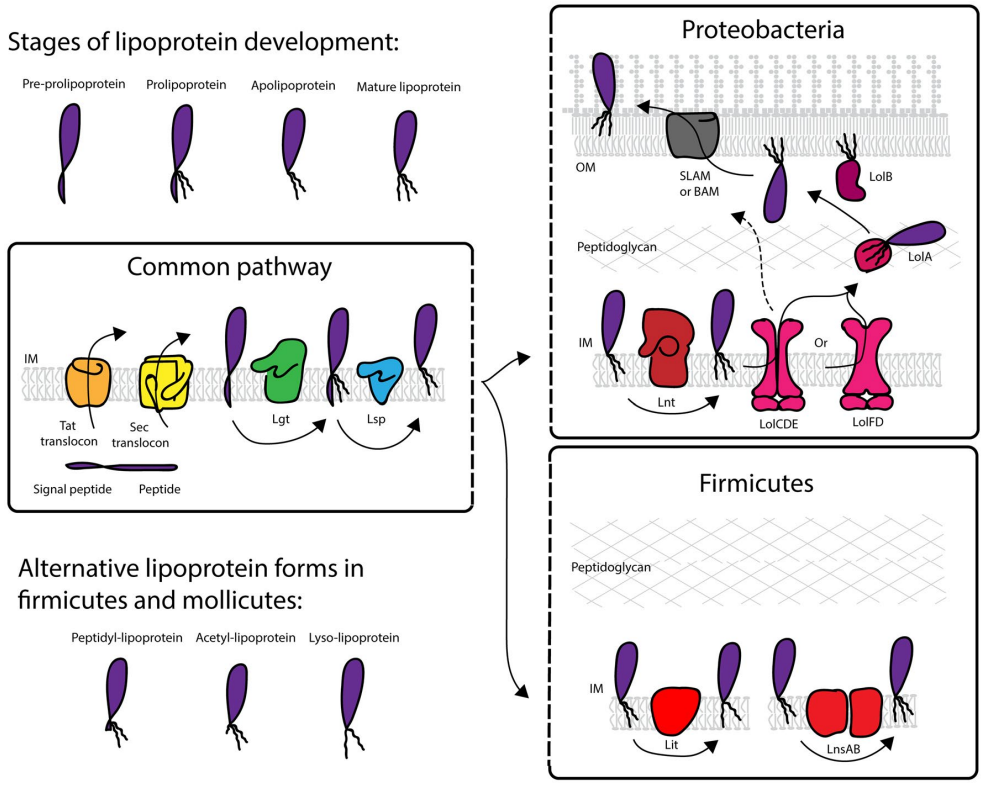


FIGURE 1 The lipoprotein biosynthesis pathway. Pre-prolipoprotein is translocated across the cytoplasmic membrane via the Sec or Tat translocons and the signal peptide is embedded in the membrane with the functional part exposed to the extra-cytoplasmic space (external to the cell in monoderm bacteria, the periplasm in diderm bacteria). The lipobox region of the signal peptide is recognized by Lgt that transfers diacylglyceryl from phosphatidylglycerol to an invariable cysteine in the lipobox forming prolipoprotein. The prolipoprotein is recognized by signal peptidase Lsp, which cleaves the signal peptide below the diacylated cysteine to form apolipoprotein. In proteobacteria, Lnt then N-acylates the apolipoprotein by transferring an acyl group from phosphatidylethanolamine to the α-amine group of the terminal cysteine to form a mature lipoprotein. The LolCDE (or LolFD) ABC-transporter transfers the lipoprotein to a periplasmic chaperone, LolA, which escorts the lipoprotein to the outer membrane where LolB inserts the triacylated protein into the membrane. In some monoderm bacteria, alternative forms of lipoproteins have been identified, including peptidyl-lipoprotein, acetyl-lipoprotein, and lyso-lipoprotein. In firmicutes, Lit forms lyso-lipoprotein from apolipoprotein and LnsA and LnsB are both involved in N-acylation of apolipoprotein resulting in triacylated lipoprotein

Kurokawa et al., 2009; Navarre et al., 1996) (Figure 1). An intra-molecular N-acyltransferase (Lit), which generates a lyso-form lipoprotein, is one such enzyme that has been characterized (Armbruster et al., 2020; Armbruster and Meredith, 2017). A recent study also identified two genes, *Lns*A and *Lns*B, in *Staphylococcus* species that are involved in N-acylation of lipoproteins (Gardiner et al., 2020). Lipoproteins are mainly located in the outer membrane and on the cell surface of proteobacteria (Wilson and Bernstein, 2016). The lipoprotein outer membrane localization (Lol) machinery is the canonical pathway for trafficking to the outer membrane, but recent studies suggest alternative Lol-independent mechanisms and other transport systems may exist in parallel.

The roles of lipoproteins in cellular processes are numerous, and include cell wall biogenesis, efflux of harmful substances and virulence. They also signal the innate immune system through recognition by Toll-like receptors where the lipid moiety is essential (Kovacs-Simon et al., 2011; Nguyen and Gotz, 2016). The essential nature of the pathway in proteobacteria is likely due to the essential function of some lipoproteins in outer membrane physiology, such as LptE in LPS translocation (Wu et al., 2006) or BamD in outer membrane protein assembly (Malinverni et al., 2006; Misra et al., 2015; Onufryk et al., 2005). In *Mycobacteria*, lipoprotein LpqW plays a key role in cell wall biogenesis and has been hypothesized as the reason for Lgt essentiality (Rainczuk et al., 2012; Tschumi et al., 2012).

New and exciting insights have been obtained in recent years on the molecular mechanism of lipoprotein modification enzymes and their structural arrangements in the membrane. The increase in antimicrobial resistance demands the identification of novel targets for the development of novel antibiotics. Due to its essential nature in proteobacteria, the accessibility of the catalytic domains of the

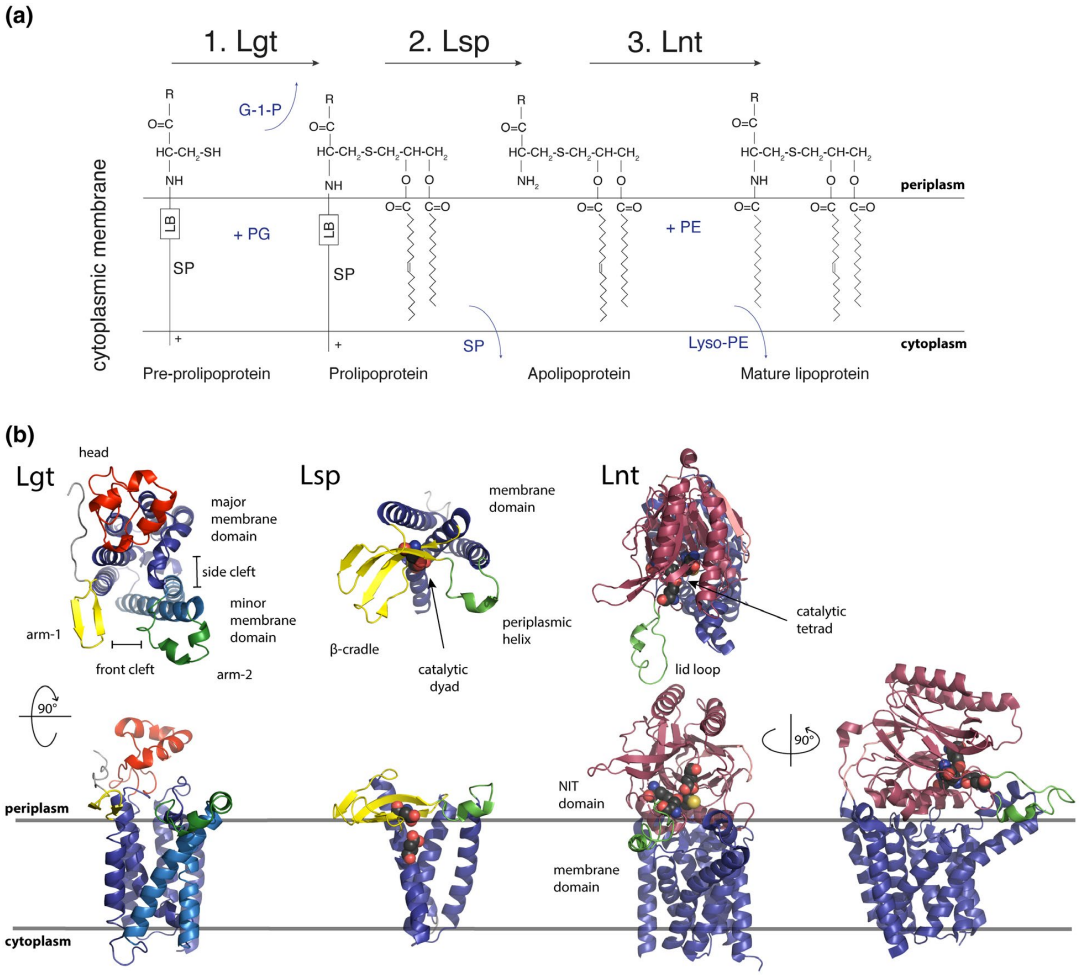


FIGURE 2 (a) Chemical modifications of the stages of lipoprotein biosynthesis. (b) X-ray crystal structure of Lgt, Lsp, and Lnt from proteobacteria (respective PDB files: 5AZC, 5DIR, 5N6H). Lgt is composed of a periplasm exposed head domain, two arm domains that rest on the periplasmic leaflet of the cytoplasmic membrane, and a minor and major body domain. Front and side clefts are formed between the two body domains where substrate is likely to enter (front cleft) and exit (side cleft). The arms possibly provide guidance and stabilizing functions for the incoming pre-prolipoprotein and the reaction is believed to occur within the central cavity. Lsp is composed of three domains, a β-cradle and periplasmic helix which both protrude from the membrane domain and rest on the periplasmic face of the cytoplasmic membrane between which the substrate binds. The two critical, catalytic residues which act to cleave the signal peptide are found in the upper half of the membrane domain (spheres). Lnt has two distinct domains: a NIT (nitrilase) domain where N-acylation occurs and a membrane domain. A flexible lid-loop (green) protrudes from the enzyme and may be correlated to binding and accessibility of substrate. The catalytic triad (spheres) is observed inside the NIT domain close to the lid-loop. G-1-P, glycerol-1-phosphate; LB, lipobox; PE, phosphatidylethanolamine; PG, phosphatidylglycerol; SP, signal peptide

enzymes, and the existence of high-throughput in vitro assays, the lipoprotein modification pathway is a promising target.

2 | ENTERING THE PATHWAY

Lipoproteins are synthesized in the cytoplasm as pre-prolipoproteins and contain an N-terminal signal sequence harboring a critical recognition sequence known as the lipobox. The signal sequence has a positively charged n-region, a hydrophobic h-region, and a lipobox containing c-region (Babu et al., 2006; Hayashi and Wu, 1990; von Heijne, 1989). The lipobox takes a standard form of $[LV]^{-3}$ $[ASTVI]^{-2}$ $[GAS]^{-1}$ $[C]^{+1}$, based on lipoprotein sequences from multiple organisms, where the invariant cysteine is the site of lipid modification that becomes the first amino acid in the mature lipoprotein (Babu et al., 2006). Variations in lipobox sequences have been reported but the invariant cysteine residue is always present (Valente et al., 2007).

The pre-prolipoprotein is translocated into the cytoplasmic membrane via the Sec translocon (Hayashi and Wu, 1985; Kosic et al., 1993; Watanabe et al., 1988) or Tat translocon (Shruthi et al., 2010a; 2010b; Thompson et al., 2010). The posttranslational targeting of secretory proteins by SecB and co-translational targeting of inner membrane proteins by signal recognition particle (SRP) converge at the Sec translocon (Tsirigotaki et al., 2017). Although the membrane translocation of lipoproteins has not been extensively studied, two examples of lipoproteins in *E. coli*, including Lpp, depend on SRP and YidC for targeting to the Sec machinery (Fröderberg et al., 2004). This may also be the pathway used in spirochetes since they lack SecB and there is no evidence for a functional Tat pathway (Zuckert, 2014). As a result of translocation, the Cys^{+1} in the lipobox is located at the membrane–periplasm interface of the outer leaflet of the cytoplasmic membrane, ready to be modified through fatty acid acylation, with the mature part of the protein located in the periplasm (Figure 1).

3 | BECOMING MATURE

3.1 | Diacylglyceryl transfer by Lgt

In the first modification step, phosphatidylglycerol: prolipoprotein diacylglyceryl transferase (Lgt) recognizes the lipobox Cys^{+1} of the bilayer embedded signal peptide, and transfers a diacylglyceryl group from phosphatidylglycerol to the pre-prolipoprotein, resulting in thioether-linked S-diacylglyceryl lipoprotein (Gan et al., 1993) (Figure 2a). Lgt possesses seven transmembrane helices (TH) with the N-terminus facing the periplasm and the C-terminus located in the cytoplasm (Daley et al., 2005; Mao et al., 2016; Pailler et al., 2012). The enzyme folds into five domains, a large transmembrane body domain with its seven THs, which in turn consists of a minor (TH2, TH3) and major (TH1, TH4-7) domain, a head domain extending into the periplasm, and two arms (arm-1, arm-2) in the periplasm facing away from the head domain parallel to the membrane (Figure 2b). Arm-1 is a short β-hairpin extending from TH1, and arm-2 consists of two α-helices extending from the minor body domain. This minor body domain is thought to be flexible and may have a role in guiding the pre-prolipoprotein substrate into the catalytic center (Mao et al., 2016). A front cleft is formed between the TH1 of the major domain and TH2 of the minor domain, and is a proposed entry site for the two Lgt substrates, pre-prolipoprotein and phosphatidylglycerol (Singh et al., 2019). A side cleft is formed between TH3 of the minor domain and TH7 of the major domain that functions as the exit portal of diacylglyceryl-modified product (Figure 2b).

A central cavity in the body domain, whose base is hydrophobic and contains a positively charged region above, with a large opening to the periplasmic side, houses the conserved H103-G-G-L sequence, the Lgt signature motif (G142–G154) and other essential residues as found in *E. coli* (Mao et al., 2016; Pailler et al., 2012). Within the cavity are two phosphatidylglycerol binding sites. At the first binding site, near the front cleft, arm-2 and Y26 interact with the phosphate group of the phospholipid. The second binding site is near essential residues R143 and R239 and is thought to be where diacylglyceryl transfer occurs. In the structure, diacylglycerol (DAG) is observed in a pocket formed by essential residues (Pailler et al., 2012; Sankaran et al., 1997), probably representing an intermediate state since DAG is not a substrate nor product of the Lgt reaction. Both alkyl groups pass through the side cleft (Mao et al., 2016). The following reaction mechanism is proposed based on the structural data. The Lgt signature motif binds the lipobox of pre-prolipoprotein coming in from the side cleft, such that the cysteine is in close proximity to the C3 ester group of phosphatidylglycerol. Upon lipoprotein binding, the thiol group of the cysteine is converted into a reactive thiyl radical via proton release to H103 that, in turn, attacks the ester bond in phosphatidylglycerol, transferring the diacylglyceryl group to the cysteine in the lipobox, releasing glycerol-1-phosphate (G1P) through a periplasmic exit.

Several models have been proposed for substrate entry and product exit: (a) the phospholipid substrate occupies the two binding sites simultaneously, and upon catalysis, phospholipid moves from site-1 to site-2 for a new round of catalysis and the product exits via the side cleft (Mao et al., 2016); or (b), binding of lipoprotein induces a conformational change that leads to entry of phospholipid in the catalytic site-2 (Mao et al., 2016). Alternatively, (c) phosphatidylglycerol and the pre-prolipoprotein both enter through the front cleft into the central cavity, where essential residue R239 acts as a gate that regulates the opening and closing of a loop in the major domain allowing products to leave via the side cleft (Singh et al., 2019).

In the early 1990's a first Lgt in vitro activity assay was reported based on a shift in mobility by high-resolution gel electrophoresis of a diacylglyceryl peptide, upon incubation with Lgt and phospholipid (Sankaran and Wu, 1994). Crude membrane fractions of bacteria with varying levels of Lgt were used as the enzyme source for the conversion of a synthetic peptide, composed of the first 24 residues of Braun's lipoprotein in the presence of radiolabeled membrane phospholipids (Gan et al., 1993; Sankaran and Wu, 1994). From these studies, the glycerol head group of phospholipid was shown to be specific for Lgt (Sankaran and Wu, 1994). A coupled enzymatic reaction described by (Sundaram et al., 2012) monitors Lgt activity through the formation of G1P, a by-product of the reaction directly correlated with enzyme activity. Dihydroxyacetone is formed from G1P using a combination of alkaline phosphatase and glycerol dehydrogenase. In a final step, resazurin is reduced to resorufin and fluorescence read-out monitored as a measure of Lgt activity. Both methods are based on the same *E. coli* strain to overproduce Lgt and the same synthetic peptide substrate, resulting in similar Km values for the peptide. In a recent report, peptide substrate LipoGFP, also containing the N-terminal sequence of Lpp fused to GFP, was used as Lgt substrate (Mao et al., 2016). This peptide was produced in *E. coli* as a glutathione-S-transferase (GST) fusion protein for purification purposes and after cleavage of GST used as substrate. Upon incubation with commercial phospholipids and purified enzyme, formation of diacylglyceryl-lipoGFP is followed by a shift in migration on SDS-PAGE and fluorescence detection.

The methods based on gel shift of diacylglyceryl peptides can be used in elaborate kinetic studies on Lgt, however, they are not compatible with high-throughput screening (HTS) required in the search for and development of novel antibiotics. Even though the resorufin fluorescence-based assay could be developed for multi-well plates, the necessity for two additional enzymes requires additional control steps and complicates the HTS set-up. Other challenges are the chemical nature of the reaction; acylated proteins, phospholipids, and integral membrane enzymes require nonclassical conditions for catalysis, as will be discussed below.

3.2 | Cleavage of the signal peptide by Lsp

Once diacylation of the lipobox cysteine by Lgt has occurred, Lsp cleaves the signal peptide liberating the α-amino group of the prolipoprotein (Figure 2a). The X-ray crystal structure of signal peptidase II (Lsp) from *P. aeruginosa* (Vogeley et al., 2016) and *S. aureus* (Olatunji et al., 2020) reveals two domains; a membrane domain consisting of the four transmembrane helices, with both the N and C terminus located in the cytoplasm (Munoa et al., 1991), and a periplasmic domain consisting of two sub domains—the β-cradle, resting on the outer leaflet of the inner membrane, and α-helix, with a single helical turn also resting on the membrane surface (Vogeley et al., 2016). Lsp belongs to the aspartate protease family (Tjalsma et al., 1999), where the catalytic aspartate residues reside at the membrane-periplasm interface in TH1 and TH4.

The incoming prolipoprotein likely enters between the β-cradle and the periplasmic helix, which form two arms extending away from the core of the enzyme (Vogeley et al., 2016). The scissile bond between the diacylglyceryl-modified cysteine and the amino acids at position-1 in the lipobox extends between the catalytic dyad (D124 and D143 in Lsp of *P. aeruginosa*) (Figure 2b) and is clamped by the β-cradle and the periplasmic helix (Olatunji et al., 2020; Vogeley et al., 2016). The catalytic site contains a water molecule in a deprotonated state. One aspartic acid residue acts as a base to attract hydrogen from the water molecule and creates a hydroxide that attacks the scissile peptide bond. This generates a tetrahedral intermediate. A second aspartic acid donates a proton to the amino terminus of the peptide, and the tetrahedral intermediate also donates a proton. This causes cleavage of the scissile bond and the substrates dissociate from the enzyme (Paetzel et al., 2002).

Lsp is the only enzyme in the lipoprotein modification pathway with known natural inhibitors. Globomycin is a cyclic peptide produced by *Streptomyces* (Inukai et al., 1978a; 1978b; Nakajima et al., 1978) that shares similarities to the signal peptide of lipoproteins (Inukai et al., 1978b). The second molecule, myxovirescin (also called TA), was isolated from *Myxococcus xanthus* (Rosenberg et al., 1973). The genome of *M. xanthus* encodes four Lsp genes (*lspA1* to *lspA4*) (Konovalova et al., 2010; Paitan et al., 1999; Xiao and Wall, 2014), two of which (*lspA3* and *lspA4*) are located in the myxovirescin biosynthetic gene cluster (Xiao and Wall, 2014). The mechanism of host protection is not fully understood but has been hypothesized due to either (over-)expression of *lspA3*, which conferred highest resistance when expressed in *E. coli* or regulation in antibiotic levels by LspA4 (Xiao and Wall, 2014). In the *S. aureus* Lsp structures, globomycin and myxovirescin share a 19-atom core structure bound in the central cavity of the enzyme, blocking the catalytic dyad (Olatunji et al., 2020), and is presumably where the signal peptide of prolipoprotein binds, whereas the macrocycles each occupy opposite sides of the catalytic site.

Proteolytic processing of prolipoprotein by Lsp in vitro was first shown in the early 1980's using a gel shift assay similar to those used in the study of Lgt (Tokunaga et al., 1982; Wu et al., 1983). Prior modification of substrate by Lgt is required for Lsp activity (Tokunaga et al., 1982; 1984). Recent work on the mode of action of globomycin and myxovirescin describe a similar coupled Lgt and Lsp reaction to obtain diacylglyceryl-modified substrate for Lsp (Olatunji et al., 2020; Vogeley et al., 2016). This study also highlights differences in enzymatic activity and inhibition by globomycin between Lsp enzymes of different bacterial species. Lsp of *P. aeruginosa* is more efficient in processing prolipopeptide than Lsp from *S. aureus* and has a lower inhibitory concentration for globomycin as measured by half maximal inhibitory concentrations (IC50 values) (Olatunji et al., 2020). Slight structural differences are observed between the Lsp enzymes in loop structures involved in keeping the antibiotic in place, and overall surface electrostatic differences between the two enzymes are also likely to play a role. Minimal inhibitory concentrations (MIC) of globomycin on bacterial cell cultures are much higher for *Pseudomonas* and *Staphylococcus* than for *E. coli* (Kiho et al., 2003; 2004). Specific small molecule inhibitors of Lsp were identified in a FRET assay based on processing of a synthetic diacylglyceryl-lipopeptide containing a fluorophore and quencher (Kitamura et al., 2018). Upon incubation with Lsp, processing of the peptide results in fluorescence of the fluorophore due to loss of the quencher. In an HTS, specific Lsp inhibitors were identified that could be optimized by medicinal chemistry to obtain IC50 values in the nanomolar range (Kitamura et al., 2018).

Lsp is essential for growth in proteobacteria and in *S. coelicolor* (Thompson et al., 2010) and probably also in *S. scabies* since suppressor mutants were readily obtained in attempts to delete *lsp* (Widdick et al., 2011). However, it is not essential in *Corynebacteria* (Dautin et al., 2020) and *Mycobacteria* but an *lsp* mutant in *M. tuberculosis* is attenuated for virulence (Rampini et al., 2008; Sander et al., 2004). The rationale for targeting lipoprotein biogenesis holds true.

3.3 | N-acyl transfer by Lnt

Lnt catalyzes a third and final step in the lipoprotein modification pathway, by N-acylation of the apolipoprotein formed by cleavage of the signal peptide by Lsp. The essential nature of Lnt is not completely conserved in proteobacteria. Recent studies demonstrate that *Francisella tularensis*, *Neisseria gonorrhoeae* (LoVullo et al., 2015), *Neisseria meningitidis* (da Silva et al., 2017), *Acinetobacter* spp (Gwin et al., 2018), and *Helicobacter pylori* (McClain et al., 2020) are viable

under laboratory conditions in the absence of Lnt. This phenomenon is possibly related to noncanonical Lol machinery in which LolF functions as LolCE in the translocation of OM lipoproteins, however, the basis and extent of this is not fully understood (see below).

Lnt is a member of the nitrilase superfamily catalyzing hydrolysis or condensation of carbon-nitrogen and nitrile bonds (Pace and Brenner, 2001). Within the enzyme, a catalytic triad E267, K335, C387 has been proposed for *E. coli* (Vidal-Ingigliardi et al., 2007). The enzyme exists in a thioester-acyl intermediate in vivo through acylation of the C387 sulfur group that is blocked for alkylation. Residues E267, K335, and E343 are involved in formation of this stable intermediate (Buddelmeijer and Young, 2010). The X-ray crystal structure of Lnt was reported by three research groups in quick succession (Lu et al., 2017; Noland et al., 2017; Wiktor et al., 2017) and has been reviewed in greater detail by (Cheng et al., 2018) (Figure 2b). Recently, Wiseman and Hogbom (2020) published a fourth similar structure. Due to the critical role of E343 and its fixed position in all structures, it has been proposed that the catalytic triad is in fact a tetrad (El Arnaout and Soulimane, 2019; Wiktor et al., 2017). Initial proton abstraction from the C387 sulfur by E267 generates a thiolate that in turn attacks the ester linkage between the *sn*-1 acyl of phosphatidylethanolamine, forming a tetrahedral intermediate stabilized by K335 and an oxyanion hole. When the tetrahedral intermediate collapses, proton abstraction from E267 releases the lyso-phospholipid by-product. When the apolipoprotein substrate enters the thioester acyl enzyme, the reaction passes through a second tetrahedral intermediate that forms when the α-amine of S-diacylglyceryl-cysteine in the apolipoprotein attacks the carboxyl carbon between C387 and the acyl chain. The mature lipoprotein is thereby formed and released. The reaction follows a ping-pong mechanism where lyso-phospholipid is released before binding of the second apolipoprotein substrate (Hillmann et al., 2011).

The characteristic catalytic domain, as seen in nitrilases, sits on top of the transmembrane domain composed of eight transmembrane helices (Figure 2b). Both termini are in the cytoplasm (Lu et al., 2017; Noland et al., 2017; Robichon et al., 2005; Wiktor et al., 2017; Wiseman and Hogbom, 2020). The nitrilase domain has a characteristic αββα fold and contains a domed cavity with an opening into the membrane domain. A phosphate-binding domain may be present which binds to and stabilizes the head group of the donor phospholipid (Noland et al., 2017). Extending from the catalytic domain is a lid loop (Lu et al., 2017) that is the most variable and flexible region between the multiple crystal structures, and contains several essential residues (Gelis-Jeanvoine et al., 2015; Lu et al., 2017; Vidal-Ingigliardi et al., 2007). It is observed resting on the membrane and also in two increasingly raised positions that may correlate with the proposed bound states of the substrates (Wiseman and Hogbom, 2020) echoing the flexibility also seen by molecular dynamics (Lu et al., 2017; Noland et al., 2017). The flexible nature of the lid loop may control entry of substrates into the active site (Lu et al., 2017; Wiseman and Hogbom, 2020). Wiseman and Hogbom (2020) propose that movement of this loop into its upward position creates a restricted access window allowing only apolipoprotein accommodation. TH3 and TH4 extend into the periplasm forming a portal for amphiphilic substrates (Wiktor et al., 2017) and various arms create an opening to the membrane playfully described as reflecting a hungry octopus (Wiktor et al., 2017; Wiseman and Hogbom, 2020). Noland et al. (2017) describe a gating phenylalanine and proposes a mechanism whereby a flexible loop, with F82 in the open position, allows phosphatidylethanolamine to bind the lipid binding groove and moves into the active site. F82 closes and positions the *sn*-1-acyl chain for nucleophilic attack by C387 generating acyl-Lnt. Then, in the open position, lyso-PE exits the enzyme allowing the entry of the fatty acid modified cysteine of apolipoprotein via the lipid channel. However, the observed gating by F82 was not correlated with the presence or absence of substrate (Wiseman and Hogbom, 2020) and is noncritical to activity (Noland et al., 2017).

Lnt activity was first demonstrated in detergent solubilized membrane vesicles with apolipoprotein substrates obtained from globomycin-treated cells (Gupta and Wu, 1991). The difference in temperature stability between Lsp and Lnt allowed for the accumulation of apolipoprotein substrate upon incubation at elevated temperatures. This study demonstrated the incorporation of palmitic acid from phospholipid through an amide bond in S-diacylglyceryl-cysteine. The initial determination of kinetic parameters of Lnt of *E. coli* was performed with an activity test based on purified Lnt, a synthetic biotinylated peptide (fibroblast-stimulating ligand 1 or FSL-1) and commercial phospholipids (Hillmann et al., 2011). The mobility shift of FSL-1 upon N-acylation by Lnt was monitored by high-resolution gel electrophoresis (Sankaran and Wu, 1994) and detection with streptavidin. Phosphatidylethanolamine was observed as the preferred acyl donor (Jackowski and Rock, 1986) with saturated fatty acids on *sn*-1 and unsaturated fatty acids on *sn*-2 (Hillmann et al., 2011). This test was recently developed into a fluorescence-based assay by using alkyne phospholipid as substrate and click-chemistry to render the N-acyl biotin peptide fluorescent, and could be detected in a sensitive manner on streptavidin-coated multi-well plates in a HTS compatible format (Nozeret et al., 2019; 2020).

4 | REACHING THE FINAL DESTINATION

In proteobacteria, the majority of lipoproteins are located in the outer membrane (OM), either in the inner leaflet of the membrane facing the periplasm, or exposed on the cell surface (Wilson and Bernstein, 2016). The nature of the +2 residue in the lipobox, and in some bacteria residues at +3 and +4, determine whether the lipoprotein is retained in the inner membrane or translocated to the OM (Narita and Tokuda, 2007; Tokuda and Matsuyama, 2004). A designated ABC-transporter, termed the Lol-machinery, is involved in translocation of lipoproteins to the OM. The Lol-machinery is generally composed of two integral membrane proteins LolC and LolE that together with ATP-ase LolD release lipoproteins from the cytoplasmic membrane to the periplasmic chaperone LolA, which transfers the protein to the OM receptor LolB (Okuda and Tokuda, 2011) (Figure 1). LolB is not strictly conserved, suggesting that other OM

receptors or alternative translocation pathways exists (Grabowicz, 2019; Liechti and Goldberg, 2012). Recent findings identified LolF as an alternative component of the ABC transporter that contains structural characteristics of both LolE and LolC and functions alongside LolD (LoVullo et al., 2015; McClain et al., 2020). Interestingly, in LolF containing bacteria, Lnt is not essential for viability, suggesting that LolF and LolD can release diacylated lipoproteins from the membrane (LoVullo et al., 2015). Lipoprotein trafficking to the OM can also occur through a LolAB-independent mechanism in certain mutant backgrounds (Grabowicz and Silhavy, 2017). Furthermore, in *Neisseria* a designated OM and surface transport machinery exists called SLAM (surface lipoprotein assembly modulator) (Hooda and Moraes, 2018) and in spirochetal species a Lol-independent proposed coupled "holding-flipping" machinery locates lipoproteins to the cell surface (Zuckert, 2014).

Three independent phenotypic screens identified inhibitors of *E. coli* growth that target the Lol machinery (McLeod et al., 2015; Nayar et al., 2015; Nickerson et al., 2018). The screens used bacteria with a permeabilized OM, either through a mutation or treatment with antimicrobial peptide to allow access of the molecules, or reduced ability to efflux toxic compounds through a mutation in an RND efflux pump. Pyridine imidazole compounds were shown to interfere with LolE and LolC (McLeod et al., 2015), and a pyrazole compound inhibits the LolCDE complex (Nayar et al., 2015). These results demonstrate the importance of OM lipoproteins in cell wall biogenesis and viability. Another inhibitor, pyrrolopyrimidinedione compound (G0507), targets LolCDE and stimulates LolD ATPase activity in vitro (Nickerson et al., 2018).

Inhibitory molecules of the Lol machinery were also identified in a chemical genomics approach. Overproduction of essential proteins were identified as suppressors of inhibition of growth in the presence of small molecules (Pathania et al., 2009). Molecule MAC13243 and its degradation products bind to the hydrophobic cavity of LolA, preventing interaction with lipoproteins (Barker et al., 2013; Pathania et al., 2009). The thiourea degradation product of MAC is an A22 analog that inhibits actin homolog MreB. A22 also acts on LolA (Barker et al., 2013). Molecular dynamic simulation experiments suggest that MAC13243 and lipoproteins occupy the LolA binding site simultaneously, that conformational changes in LolA upon lipoprotein binding are restricted, and that the interaction with lipoprotein is weakened (Boags et al., 2019). The MAC13243-LolA interaction leads to an increase in OM permeability (Muheim et al., 2017). The crucial role of the lipoprotein biosynthesis pathway is seen through these inhibitor studies of the downstream processes, and further affirms the potential antimicrobial benefits of targeting this pathway.

5 | TARGETING THE PATHWAY

The essential nature of the lipoprotein posttranslational modification pathway in proteobacteria makes it an intriguing novel target for antimicrobial therapy. Another advantage is the accessibility of the active sites from the periplasm as molecules need not traverse the cytoplasmic membrane. To date, globomycin and myxovirescin are the only inhibitors of the lipoprotein modification pathway, both targeting Lsp, but neither are in clinical use. Clinical trials for the treatment of gingivitis by myxovirescin did shown some promise (Manor et al., 1989). Stability, effectiveness and toxicity in host cells has proven to be an obstacle for antimicrobial peptides (Chen and Lu, 2020) as is probably the case for these compounds.

Until recently, assays developed to study the pathway involved radiolabeling and gel-shift analysis. These assays, albeit a very valuable tool, are low-throughput, and therefore, not suitable for HTS applications. The nature of the lipoprotein modification reactions is complex. The enzymes are integral membrane proteins, and the peptide substrates and phospholipids are also components of the lipid bilayer. It is therefore not straightforward to develop assays that are simple, homogeneous, soluble, and adaptable for the screening of inhibitors. A coupled fluorescence-based assay has been developed for Lgt that may be adapted for HTS (Sundaram et al., 2012) although no applications have been reported. Recently, an in vitro HTS Lsp assay was developed using FRET and was used to screen over 640,000 molecules for Lsp inhibition. This study yielded promising results (Kitamura et al., 2018). A fluorescence-based click-chemistry assay compatible for HTS has been developed for Lnt activity and is a promising tool for screening libraries of molecules (Nozeret et al., 2019; 2020). However, a drawback to target-based in vitro screening is the potential need to chemically alter inhibitors to enable passage of the OM and to access the periplasm of proteobacteria. The identification of targets of inhibitors found in phenotypic screens requires whole genome sequencing of resistant clones. Furthermore, the active compounds described so far are only able to prevent growth of bacterial cell cultures in the presence of a permeable outer membrane. An alternative approach is the use of structure-based drug design (Staker et al., 2015). Since the structures of the lipoprotein modification enzymes are known there is greater information available for this approach and has been reviewed in detail recently (El Arnaout and Soulimane, 2019). Ideally, these approaches should be used in parallel in the search for novel antibacterial agents.

6 | CONCLUSIONS

Tremendous progress has been made in recent years on the structural understanding of the lipoprotein modification enzymes. Some insights have been obtained in the molecular mechanism of the reactions, in particular on phospholipid substrate specificity for Lgt and Lnt, and inhibition of Lsp by globomycin and myxovirescin. The results with globomycin suggest that the same enzyme from different bacterial species may differ in substrate specificity and efficacy of the reaction. Furthermore, it is unknown how the enzymes bind the peptide substrates and how this affects conformational changes and catalysis of the reactions. In most studies on lipoprotein modification and targeting, Braun's lipoprotein of *E. coli* is used as a model protein but it is far from being conserved among bacteria. However, interesting similarities

are seen between the three enzymes, such as the arm domains or channels, which allow entry of substrate and phospholipids, and the flexibility of extended loops presumably permitting different substrates into close proximity of their active sites. To date, all studies have been conducted in isolation and there is little to no research into the functional interactions between the enzymes. The efficient nature of the system, and relative low abundance of the enzymes but high abundance of lipoproteins, hints toward a coordinated relationship to guarantee efficient lipidation of proteins as suggested in 1982 by Tokunaga (Tokunaga et al., 1982).

The development of tools to study this essential pathway has yielded the identification of inhibitors, which demonstrates progress in the race to develop or discover novel antibiotics. The combined, synergistic use of inhibitors targeting the lipoprotein modification and OM sorting pathways could be one method to increase efficacy of treatment and reduce frequency of resistance. Examples of successful combination therapy or a multi target approach have been reported (Tyers and Wright, 2019). Colistin, for example, is membrane-permeabilizing agent, which increases drug access to the cell that has been used in combination with other antibiotics. Alongside the identification of novel antibiotics, inhibitors are a useful tool in elucidating molecular mechanisms of proteins, and in the study of complex pathways. We believe the potential for inhibiting this pathway and the recent advances in our understanding make the lipoprotein modification pathway an exciting area for future study, and may play a key role in the fight against antimicrobial resistant pathogens.

ACKNOWLEDGMENTS

We thank Richard Wheeler and Karine Nozeret for critical reading the manuscript and all members of the BGPB Unit for their support. Simon Legood was supported by the Pasteur-Paris University (PPU) International PhD Program. Research received funding from the Institut Carnot Infectious Diseases Global Care (16 CARN 0023-01 Project iLiNT) and from the French Government's Investissement d'Avenir program, Laboratoire d'Excellence "Integrative Biology of Emerging Infectious Diseases" (grant n°ANR-10-LABX-62-IBEID).

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

ORCID


Nienke Buddelmeijer https://orcid.org/0000-0003-1116-3625

REFERENCES

Armbruster, K.M., Komazin, G. and Meredith, T.C. (2020) Bacterial lyso-form lipoproteins are synthesized via an intramolecular acyl chain migration. *Journal of Biological Chemistry*, *295*, 10195–10211.

Armbruster, K.M. and Meredith, T.C. (2017) Identification of the lyso-form N-Acyl intramolecular transferase in low-GC firmicutes. *Journal of Bacteriology*, *199*, e00099-17.

Asanuma, M., Kurokawa, K., Ichikawa, R., Ryu, K.H., Chae, J.H., Dohmae, N., et al. (2011) Structural evidence of alpha-aminoacylated lipoproteins of *Staphylococcus aureus*. *FEBS Journal*, *278*, 716–728.

Babu, M.M., Priya, M.L., Selvan, A.T., Madera, M., Gough, J., Aravind, L., et al. (2006) A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *Journal of Bacteriology*, *188*, 2761–2773.

Barker, C.A., Allison, S.E., Zlitni, S., Nguyen, N.D., Das, R., Melacini, G., et al. (2013) Degradation of MAC13243 and studies of the interaction of resulting thiourea compounds with the lipoprotein targeting chaperone LolA. *Bioorganic & Medicinal Chemistry Letters*, *23*, 2426–2431.

Boags, A., Samsudin, F. and Khalid, S. (2019) Details of hydrophobic entanglement between small molecules and Braun's lipoprotein within the cavity of the bacterial chaperone LolA. *Scientific Reports*, *9*, 3717.

Buddelmeijer, N. and Young, R. (2010) The essential *Escherichia coli* apolipoprotein N-acyltransferase (Lnt) exists as an extracytoplasmic thioester acyl-enzyme intermediate. *Biochemistry*, *49*, 341–346.

Chen, C.H. and Lu, T.K. (2020) Development and challenges of antimicrobial peptides for therapeutic applications. *Antibiotics*, *9*, 24.

Cheng, W., Doyle, D.A. and El Arnaout, T. (2018) The N-acyltransferase Lnt: structure-function insights from recent simultaneous studies. *International Journal of Biological Macromolecules*, *117*, 870–877.

da Silva, R.A.G., Churchward, C.P., Karlyshev, A.V., Eleftheriadou, O., Snabaitis, A.K., Longman, M.R., et al. (2017) The role of apolipoprotein N-acyl transferase, Lnt, in the lipidation of factor H binding protein of *Neisseria meningitidis* strain MC58 and its potential as a drug target. *British Journal of Pharmacology*, *174*, 2247–2260.

Daley, D.O., Rapp, M., Granseth, E., Melen, K., Drew, D. and von Heijne, G. (2005) Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science*, *308*, 1321–1323.

Dautin, N., Argentini, M., Mohiman, N., Labarre, C., Cornu, D., Sago, L., et al. (2020) Role of the unique, non-essential phosphatidylglycerol:prolipoprotein diacylglyceryl transferase (Lgt) in *Corynebacterium glutamicum*. *Microbiology*, *166*, 759–776.

El Arnaout, T. and Soulimane, T. (2019) Targeting lipoprotein biogenesis: considerations towards antimicrobials. *Trends in Biochemical Sciences*, *44*, 701–715.

Fröderberg, L., Houben, E.N., Baars, L., Luirink, J. and de Gier, J.W. (2004) Targeting and translocation of two lipoproteins in *Escherichia coli* via the SRP/Sec/YidC pathway. *Journal of Biological Chemistry*, *279*, 31026–31032.

Gan, K., Gupta, S.D., Sankaran, K., Schmid, M.B. and Wu, H.C. (1993) Isolation and characterization of a temperature-sensitive mutant of *Salmonella typhimurium* defective in prolipoprotein modification. *Journal of Biological Chemistry*, *268*, 16544–16550.

Gardiner, J.H., Komazin, G., Matsuo, M., Götz, F. and Meredith, T.C. (2020) Lipoprotein N-acylation in *Staphylococcus aureus* is catalyzed by a two-component acyl transferase system. *mBio*, *11*, e01619–e01620.

Gelis-Jeanvoine, S., Lory, S., Oberto, J. and Buddelmeijer, N. (2015) Residues located on membrane-embedded flexible loops are essential for the second step of the apolipoprotein N-acyltransferase reaction. *Molecular Microbiology*, *95*, 692–705.

Grabowicz, M. (2019) *Lipoproteins and Their Trafficking to the Outer Membrane*. Washington, DC: ASM Press.

Grabowicz, M. and Silhavy, T.J. (2017) Redefining the essential trafficking pathway for outer membrane lipoproteins. *Proceedings of the National Academy of Sciences*, *114*, 4769–4774.

Gupta, S.D. and Wu, H.C. (1991) Identification and subcellular localization of apolipoprotein N-acyltransferase in *Escherichia coli*. *FEMS Microbiology Letters*, *62*, 37–41.

Gwin, C.M., Prakash, N., Christian Belisario, J., Haider, L., Rosen, M.L., Martinez, L.R., et al. (2018) The apolipoprotein N-acyl transferase Lnt is dispensable for growth in Acinetobacter species. *Microbiology*, *164*, 1547–1556.

Hantke, K. and Braun, V. (1973) Covalent binding of lipid to protein. Diglyceride and amide-linked fatty acid at the N-terminal end of the murein-lipoprotein of the *Escherichia coli* outer membrane. *European Journal of Biochemistry*, *34*, 284–296.

Hayashi, S. and Wu, H.C. (1985) Accumulation of prolipoprotein in *Escherichia coli* mutants defective in protein secretion. *Journal of Bacteriology*, *161*, 949–954.

Hayashi, S. and Wu, H.C. (1990) Lipoproteins in bacteria. *Journal of Bioenergetics and Biomembranes*, *22*, 451–471.

Hillmann, F., Argentini, M. and Buddelmeijer, N. (2011) Kinetics and phospholipid specificity of apolipoprotein N-acyltransferase. *Journal of Biological Chemistry*, *286*, 27936–27946.

Hooda, Y. and Moraes, T.F. (2018) Translocation of lipoproteins to the surface of gram negative bacteria. *Current Opinion in Structural Biology*, *51*, 73–79.

Inukai, M., Enokita, R., Torikata, A., Nakahara, M., Iwado, S. and Arai, M. (1978a) Globomycin, a new peptide antibiotic with spheroplast-forming activity. I. Taxonomy of producing organisms and fermentation. *The Journal of Antibiotics*, *31*, 410–420.

Inukai, M., Takeuchi, M., Shimizu, K. and Arai, M. (1978b) Mechanism of action of globomycin. *The Journal of Antibiotics*, *31*, 1203–1205.

Jackowski, S. and Rock, C.O. (1986) Transfer of fatty acids from the 1-position of phosphatidylethanolamine to the major outer membrane lipoprotein of *Escherichia coli*. *Journal of Biological Chemistry*, *261*, 11328–11333.

Kiho, T., Nakayama, M., Yasuda, K., Miyakoshi, S., Inukai, M. and Kogen, H. (2003) Synthesis and antimicrobial activity of novel globomycin analogues. *Bioorganic & Medicinal Chemistry Letters*, *13*, 2315–2318.

Kiho, T., Nakayama, M., Yasuda, K., Miyakoshi, S., Inukai, M. and Kogen, H. (2004) Structure-activity relationships of globomycin analogues as antibiotics. *Bioorganic & Medicinal Chemistry*, *12*, 337–361.

Kitamura, S., Owensby, A., Wall, D. and Wolan, D.W. (2018) Lipoprotein signal peptidase inhibitors with antibiotic properties identified through design of a robust in vitro HT platform. *Cell Chemical Biology*, *25*, 301–308.e312.

Konovalova, A., Petters, T. and Sogaard-Andersen, L. (2010) Extracellular biology of *Myxococcus xanthus*. *FEMS Microbiology Reviews*, *34*, 89–106.

Kosic, N., Sugai, M., Fan, C.K. and Wu, H.C. (1993) Processing of lipid-modified prolipoprotein requires energy and sec gene products in vivo. *Journal of Bacteriology*, *175*, 6113–6117.

Kovacs-Simon, A., Titball, R.W. and Michell, S.L. (2011) Lipoproteins of bacterial pathogens. *Infection and Immunity*, *79*, 548–561.

Kurokawa, K., Lee, H., Roh, K.B., Asanuma, M., Kim, Y.S., Nakayama, H., et al. (2009) The triacylated ATP binding cluster transporter substrate-binding lipoprotein of *Staphylococcus aureus* functions as a native ligand for toll-like receptor 2. *Journal of Biological Chemistry*, *284*, 8406–8411.

Liechti, G. and Goldberg, J.B. (2012) Outer membrane biogenesis in *Escherichia coli*, *Neisseria meningitidis*, and *Helicobacter pylori*: paradigm deviations in H. pylori. *Frontiers in Cellular and Infection Microbiology*, *2*, 29.

LoVullo, E.D., Wright, L.F., Isabella, V., Huntley, J.F. and Pavelka, M.S. Jr (2015) Revisiting the gram-negative lipoprotein paradigm. *Journal of Bacteriology*, *197*, 1705–1715.

Lu, G., Xu, Y., Zhang, K., Xiong, Y., Li, H., Cui, L., et al. (2017) Crystal structure of *E. coli* apolipoprotein N-acyl transferase. *Nature Communications*, *8*, 15948.

Malinverni, J.C., Werner, J., Kim, S., Sklar, J.G., Kahne, D., Misra, R., et al. (2006) YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Molecular Microbiology*, *61*, 151–164.

Manor, A., Eli, I., Varon, M., Judes, H. and Rosenberg, E. (1989) Effect of adhesive antibiotic TA on plaque and gingivitis in man. *Journal of Clinical Periodontology*, *16*, 621–624.

Mao, G., Zhao, Y., Kang, X., Li, Z., Zhang, Y., Wang, X., et al. (2016) Crystal structure of *E. coli* lipoprotein diacylglyceryl transferase. *Nature Communications*, *7*, 10198.

McClain, M.S., Voss, B.J., and Cover, T.L. (2020) Lipoprotein processing and sorting in *Helicobacter pylori*. *mBio*, *11*, e00911-20.

McLeod, S.M., Fleming, P.R., MacCormack, K., McLaughlin, R.E., Whiteaker, J.D., Narita, S., et al. (2015) Small-molecule inhibitors of gram-negative lipoprotein trafficking discovered by phenotypic screening. *Journal of Bacteriology*, *197*, 1075–1082.

Misra, R., Stikeleather, R. and Gabriele, R. (2015) In vivo roles of BamA, BamB and BamD in the biogenesis of BamA, a core protein of the beta-barrel assembly machine of *Escherichia coli*. *Journal of Molecular Biology*, *427*, 1061–1074.

Muheim, C., Gotzke, H., Eriksson, A.U., Lindberg, S., Lauritsen, I., Norholm, M.H.H., et al. (2017) Increasing the permeability of *Escherichia coli* using MAC13243. *Scientific Reports*, *7*, 17629.

Munoa, F.J., Miller, K.W., Beers, R., Graham, M. and Wu, H.C. (1991) Membrane topology of *Escherichia coli* prolipoprotein signal peptidase (signal peptidase II). *Journal of Biological Chemistry*, *266*, 17667–17672.

Nakajima, M., Inukai, M., Haneishi, T., Terahara, A., Arai, M., Kinoshita, T., et al. (1978) Globomycin, a new peptide antibiotic with spheroplast-forming activity. III. Structural determination of globomycin. *Journal of Antibiotics*, *31*, 426–432.

Narita, S. and Tokuda, H. (2007) Amino acids at positions 3 and 4 determine the membrane specificity of *Pseudomonas aeruginosa* lipoproteins. *Journal of Biological Chemistry*, *282*, 13372–13378.

Navarre, W.W., Daefler, S. and Schneewind, O. (1996) Cell wall sorting of lipoproteins in *Staphylococcus aureus*. *Journal of Bacteriology*, *178*, 441–446.

Nayar, A.S., Dougherty, T.J., Ferguson, K.E., Granger, B.A., McWilliams, L., Stacey, C., et al. (2015) Novel antibacterial targets and compounds revealed by a high-throughput cell wall reporter assay. *Journal of Bacteriology*, *197*, 1726–1734.

Nguyen, M.T. and Gotz, F. (2016) Lipoproteins of gram-positive bacteria: key players in the immune response and virulence. *Microbiology and Molecular Biology Reviews*, *80*, 891–903.

Nickerson, N.N., Jao, C.C., Xu, Y., Quinn, J., Skippington, E., Alexander, M.K., et al. (2018) A novel inhibitor of the LolCDE ABC transporter essential for lipoprotein trafficking in gram-negative bacteria. *Antimicrobial Agents and Chemotherapy*, *62*(4), e02151-17.

Noland, C.L., Kattke, M.D., Diao, J., Gloor, S.L., Pantua, H., Reichelt, M., et al. (2017) Structural insights into lipoprotein N-acylation by *Escherichia coli* apolipoprotein N-acyltransferase. *Proceedings of the National Academy of Sciences*, *114*, E6044–E6053.

Nozeret, K., Boucharlat, A., Agou, F. and Buddelmeijer, N. (2019) A sensitive fluorescence-based assay to monitor enzymatic activity of the essential integral membrane protein Apolipoprotein N-acyltransferase (Lnt). *Scientific Reports*, *9*, 15978.

Nozeret, K., Pernin, A. and Buddelmeijer, N. (2020) Click-chemistry based fluorometric assay for apolipoprotein N-acyltransferase from enzyme characterization to high-throughput screening. *Journal of Visualized Experiments*, *159*, e61146.

Okuda, S. and Tokuda, H. (2011) Lipoprotein sorting in bacteria. *Annual Review of Microbiology*, *65*, 239–259.

Olatunji, S., Yu, X., Bailey, J., Huang, C.Y., Zapotoczna, M., Bowen, K., et al. (2020) Structures of lipoprotein signal peptidase II from *Staphylococcus aureus* complexed with antibiotics globomycin and myxovirescin. *Nature Communications*, *11*, 140.

Onufryk, C., Crouch, M.L., Fang, F.C. and Gross, C.A. (2005) Characterization of six lipoproteins in the sigmaE regulon. *Journal of Bacteriology*, *187*, 4552–4561.

Pace, H.C. and Brenner, C. (2001) The nitrilase superfamily: classification, structure and function. *Genome Biology*, *2*, 1–9.

Paetzel, M., Karla, A., Strynadka, N.C. and Dalbey, R.E. (2002) Signal peptidases. *Chemical Reviews*, *102*, 4549–4580.

Pailler, J., Aucher, W., Pires, M. and Buddelmeijer, N. (2012) Phosphatidylglycerol:prolipoprotein diacylglyceryl transferase (Lgt) of *Escherichia coli* has seven transmembrane segments, and its essential residues are embedded in the membrane. *Journal of Bacteriology*, *194*, 2142–2151.

Paitan, Y., Orr, E., Ron, E.Z. and Rosenberg, E. (1999) A nonessential signal peptidase II (Lsp) of *Myxococcus xanthus* might be involved in biosynthesis of the polyketide antibiotic TA. *Journal of Bacteriology*, *181*, 5644–5651.

Pathania, R., Zlitni, S., Barker, C., Das, R., Gerritsma, D.A., Lebert, J., et al. (2009) Chemical genomics in *Escherichia coli* identifies an inhibitor of bacterial lipoprotein targeting. *Nature Chemical Biology*, *5*, 849–856.

Rainczuk, A.K., Yamaryo-Botte, Y., Brammananth, R., Stinear, T.P., Seemann, T., Coppel, R.L., et al. (2012) The lipoprotein LpqW is essential for the mannosylation of periplasmic glycolipids in *Corynebacteria*. *Journal of Biological Chemistry*, *287*, 42726–42738.

Rampini, S.K., Selchow, P., Keller, C., Ehlers, S., Bottger, E.C. and Sander, P. (2008) LspA inactivation in *Mycobacterium tuberculosis* results in attenuation without affecting phagosome maturation arrest. *Microbiology*, *154*, 2991–3001.

Robichon, C., Vidal-Ingigliardi, D. and Pugsley, A.P. (2005) Depletion of apolipoprotein N-acyltransferase causes mislocalization of outer membrane lipoproteins in *Escherichia coli*. *Journal of Biological Chemistry*, *280*, 974–983.

Rosenberg, E., Vaks, B. and Zuckerberg, A. (1973) Bactericidal action of an antibiotic produced by *Myxococcus xanthus*. *Antimicrobial Agents and Chemotherapy*, *4*, 507–513.

Sander, P., Rezwan, M., Walker, B., Rampini, S.K., Kroppenstedt, R.M., Ehlers, S., et al. (2004) Lipoprotein processing is required for virulence of *Mycobacterium tuberculosis*. *Molecular Microbiology*, *52*, 1543–1552.

Sankaran, K., Gan, K., Rash, B., Qi, H.Y., Wu, H.C. and Rick, P.D. (1997) Roles of histidine-103 and tyrosine-235 in the function of the prolipoprotein diacylglyceryl transferase of *Escherichia coli*. *Journal of Bacteriology*, *179*, 2944–2948.

Sankaran, K. and Wu, H.C. (1994) Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *Journal of Biological Chemistry*, *269*, 19701–19706.

Shruthi, H., Anand, P., Murugan, V. and Sankaran, K. (2010a) Twin arginine translocase pathway and fast-folding lipoprotein biosynthesis in *E. coli*: interesting implications and applications. *Molecular BioSystems*, *6*, 999–1007.

Shruthi, H., Babu, M.M. and Sankaran, K. (2010b) TAT-pathway-dependent lipoproteins as a niche-based adaptation in prokaryotes. *Journal of Molecular Evolution*, *70*, 359–370.

Singh, W., Bilal, M., McClory, J., Dourado, D., Quinn, D., Moody, T.S., et al. (2019) Mechanism of phosphatidylglycerol activation catalyzed by prolipoprotein diacylglyceryl transferase. *The Journal of Physical Chemistry B*, *123*, 7092–7102.

Staker, B.L., Buchko, G.W. and Myler, P.J. (2015) Recent contributions of structure-based drug design to the development of antibacterial compounds. *Current Opinion in Microbiology*, *27*, 133–138.

Sundaram, S., Banerjee, S. and Sankaran, K. (2012) The first nonradioactive fluorescence assay for phosphatidylglycerol:prolipoprotein diacylglyceryl transferase that initiates bacterial lipoprotein biosynthesis. *Analytical Biochemistry*, *423*, 163–170.

Thompson, B.J., Widdick, D.A., Hicks, M.G., Chandra, G., Sutcliffe, I.C., Palmer, T., et al. (2010) Investigating lipoprotein biogenesis and function in the model Gram-positive bacterium *Streptomyces coelicolor*. *Molecular Microbiology*, *77*, 943–957.

Tjalsma, H., Zanen, G., Venema, G., Bron, S. and van Dijl, J.M. (1999) The potential active site of the lipoprotein-specific (type II) signal peptidase of *Bacillus subtilis*. *Journal of Biological Chemistry*, *274*, 28191–28197.

Tokuda, H. and Matsuyama, S. (2004) Sorting of lipoproteins to the outer membrane in *E. coli*. *Biochimica et Biophysica Acta*, *1694*, 1–9.

Tokunaga, M., Loranger, J.M. and Wu, H.C. (1984) Prolipoprotein modification and processing enzymes in *Escherichia coli*. *Journal of Biological Chemistry*, *259*, 3825–3830.

Tokunaga, M., Tokunaga, H. and Wu, H.C. (1982) Post-translational modification and processing of *Escherichia coli* prolipoprotein *in vitro*. *Proceedings of the National Academy of Sciences*, *79*, 2255–2259.

Tschumi, A., Grau, T., Albrecht, D., Rezwan, M., Antelmann, H. and Sander, P. (2012) Functional analyses of mycobacterial lipoprotein diacylglyceryl transferase and comparative secretome analysis of a mycobacterial lgt mutant. *Journal of Bacteriology*, *194*, 3938–3949.

Tsirigotaki, A., De Geyter, J., Sostaric, N., Economou, A. and Karamanou, S. (2017) Protein export through the bacterial Sec pathway. *Nature Reviews Microbiology*, *15*, 21–36.

Tyers, M. and Wright, G.D. (2019) Drug combinations: a strategy to extend the life of antibiotics in the 21st century. *Nature Reviews Microbiology*, *17*, 141–155.

Valente, F.M., Pereira, P.M., Venceslau, S.S., Regalla, M., Coelho, A.V. and Pereira, I.A. (2007) The [NiFeSe] hydrogenase from *Desulfovibrio vulgaris* Hildenborough is a bacterial lipoprotein lacking a typical lipoprotein signal peptide. *FEBS Letters*, *581*, 3341–3344.

Vidal-Ingigliardi, D., Lewenza, S. and Buddelmeijer, N. (2007) Identification of essential residues in apolipoprotein N-acyl transferase, a member of the CN hydrolase family. *Journal of Bacteriology*, *189*, 4456–4464.

Vogeley, L., El Arnaout, T., Bailey, J., Stansfeld, P.J., Boland, C. and Caffrey, M. (2016) Structural basis of lipoprotein signal peptidase II action and inhibition by the antibiotic globomycin. *Science*, *351*, 876–880.

von Heijne, G. (1989) Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. *Nature*, *341*, 456–458.

Watanabe, T., Hayashi, S. and Wu, H.C. (1988) Synthesis and export of the outer membrane lipoprotein in *Escherichia coli* mutants defective in generalized protein export. *Journal of Bacteriology*, *170*, 4001–4007.

Widdick, D.A., Hicks, M.G., Thompson, B.J., Tschumi, A., Chandra, G., Sutcliffe, I.C., et al. (2011) Dissecting the complete lipoprotein biogenesis pathway in *Streptomyces scabies*. *Molecular microbiology*, *80*, 1395–1412.

Wiktor, M., Weichert, D., Howe, N., Huang, C.Y., Olieric, V., Boland, C., et al. (2017) Structural insights into the mechanism of the membrane integral N-acyltransferase step in bacterial lipoprotein synthesis. *Nature Communications*, *8*, 15952.

Wilson, M.M. and Bernstein, H.D. (2016) Surface-exposed lipoproteins: an emerging secretion phenomenon in gram-negative bacteria. *Trends in Microbiology*, *24*, 198–208.

Wiseman, B. and Hogbom, M. (2020) Conformational changes in apolipoprotein N-acyltransferase (Lnt). *Scientific Reports*, *10*, 639.

Wu, H.C., Tokunaga, M., Tokunaga, H., Hayashi, S. and Giam, C.Z. (1983) Posttranslational modification and processing of membrane lipoproteins in bacteria. *Journal of Cellular Biochemistry*, *22*, 161–171.

Wu, T., McCandlish, A.C., Gronenberg, L.S., Chng, S., Silhavy, T.J. and Kahne, D. (2006) Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of *Escherichia coli*. *Proceedings of the National Academy of Sciences*, *103*, 11754–11759.

Xiao, Y. and Wall, D. (2014) Genetic redundancy, proximity, and functionality of lspA, the target of antibiotic TA, in the *Myxococcus xanthus* producer strain. *Journal of Bacteriology*, *196*, 1174–1183.

Zuckert, W.R. (2014) Secretion of bacterial lipoproteins: through the cytoplasmic membrane, the periplasm and beyond. *Biochimica et Biophysica Acta*, *1843*, 1509–1516.

How to cite this article: Legood S, Boneca IG, Buddelmeijer N. Mode of action of lipoprotein modification enzymes—Novel antibacterial targets. *Mol Microbiol* 2021;115:356–365. https://doi.org/10.1111/mmi.14610

Diacylglyceryl transfer by Lgt

Although our understanding of Lgt is limited there have been a few advances in recent years. As the focus of this thesis, a more complete introduction to Lgt will be presented here. It aims to demonstrate the current understanding in the field and to highlight areas where further research is required.

Protein sequence conservation

Early analysis of the conservation of the Lgt amino-acid sequence compared *E. coli*, *S. typhimurium*, *H. influenzae* and *S. aureus* and found the sequences to be closely aligned [76, 77]. The first mention of a conserved region, namely the H_{103}GGL motif was identified, although initially described as an extended motif of H_{103}GGLIG. This motif was proposed to have catalytic activity [77]. However, H_{103} was soon noted to be less well conserved than originally thought with glutamine, tyrosine and tryptophan all being observed at this position in different species (Table 3). Although there is a correlation between Lgt proteins of Gram-positive and Gram-negative bacteria having either tyrosine or histidine, respectively, it is not entirely conserved in this manner. A phylogenetic tree of the Lgt amino-acid sequence observed two roots, Gram-positive and Gram-negative. Banerjee *et al.* (2013) state but do not show that inclusion of *M. smegmatis* in the tree places it between Gram-positive and Gram-negative bacteria suggesting it is close to where the shift from histidine to tyrosine took place [78].

Singh *et al.* (2019) also state that H_{103} is not widely conserved beyond proteobacteria and display a Hidden Markov Model (HMM) motif from 4,860 Lgt sequences where phenylalanine can also be observed at this position [79]. However, no breakdown of the sequences is presented, and no analysis discussed. The broadest ranging alignment only contained 8 strains: *E. coli*, *T. maritima*, *B. cereus (containing 2 lgt genes)*, *P. chromatophore (amoeba)*, *L. lactis*, *S. coelicolor* and *C. perfringens*, but there is no comment on the data. However, it is clear that the HGGL motif has variations between species (Table 3).

The most extensive sequence conservation study to be well described of Lgt compared 446 annotated protein sequences from proteobacteria, firmicutes and actinobacteria. A number of highly conserved residues were noted across all the phylogenetic groups and between groups (Figure 6) [80]. The majority of conserved residues were found within the predicted transmembrane domains. H_{103} was not noted as fully conserved in either group.

Table 3. Conservation of residue 103 as described in the literature.

Residue at site 103	Organism	Detail	Reference
Histidine	*Escherichia coli*	Gram-negative	[76, 77, 81]
	Salmonella typhimurium	Gram-negative	[76, 77]
	Haemophilus influenzae	Gram-negative	[77]
	Staphylococcus aureus	Gram-positive	[77, 78]
	Thermotoga maritima	Gram-negative	[81]
	Bacillus cereus, Bacillus subtilis	Gram-positive	[78]
	Paulinella chromatophora	Amoeba	[81]
	Clostridium sp.	Gram-positive	[78]
Glutamine	*Mycoplasma gentalium*	Mycoplasma	[82]
Tyrosine	*Streptococcus sp*	Gram-positive	[78]
	Enterococcus sp	Gram-positive	[78]
	Lactobacillus sp	Gram-positive	[78]
	Pediococcus sp	Gram-positive	[78]
	Carnobacterium sp	Gram-positive	[78]
	Aerococcus sp	Gram-positive	[78]
	Mycobacterium smegmatis	Actinobacteria	[78]
	Lactococcus lactis	Gram-positive	[78, 81]
	Clostridium perfringens	Gram-positive	[81]
Tryptophan	*Streptomyces coelicolor*	Gram-positive	[81]

Another motif, termed the Lgt signature motif, has been shown to be well conserved [79, 81]. This motif is housed within the catalytic core of the protein (described below) (Figure 8, 21).

Lgt is generally considered to be present as a single copy gene as this is the case for the majority of strains mentioned thus far. However, it has been reported that some species contain more than one copy of *lgt*. *M. smegmatis* has two *lgt* homologs of which only one is active and the other contains sequence variations at key sites ($Y_{26}H$ and $N_{146}C$) [83]. *B. cereus* also has two genes which were aligned as part of the work by Mao *et al.* (2016). Interestingly, two copies of *lgt* are reported in *S. coelicolor* and *S. clavuligerus* [84] but only one in the closely related *S. scabies* [85] raising questions about possible differentiating roles of different Lgts within a single species.

Although some analysis of the sequence conservation of Lgt has been completed as described here, some gaps in our knowledge still remain. A more detailed analysis of sequence conservation may assist in identifying areas of functional interest. A greater look at the conservation of H_{103} and whether there are more conserved motifs in the sequence may also help elucidate information about species specificity or substrate recognition. The specific role of different Lgts within the same species has also been overlooked and merits further study.

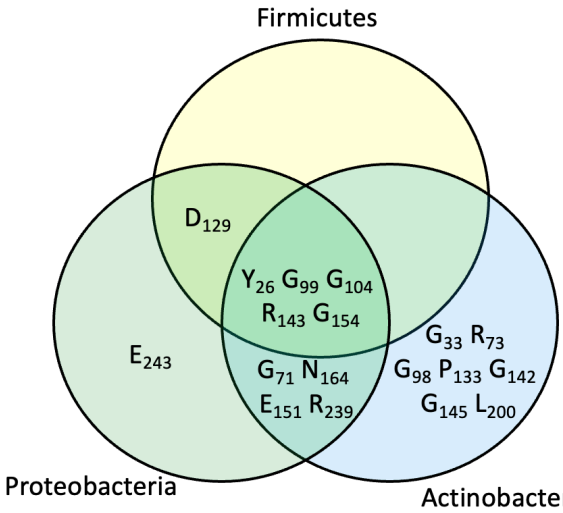


Figure 6. Conservation of residues between three phyla of bacteria. *Adapted from Pailler et al., 2012 [80].*

Genetic context and regulation of expression

The earliest studies to take place in the region of the then undescribed *lgt* gene found that *thyA* (encoding thymidylate synthase) could not be deleted due to a gene in the region having an essential function [86]. A transcription stop sequence was discovered within *thyA* which led to the hypothesis that there was a transcriptional-linked upstream gene and they observed that disruption was indeed only possible after the stop sequence of this upstream gene [87]. This upstream gene was later identified as *lgt* (previously named unidentified membrane protein A *ump*A) [88]. A temperature sensitive *S. typhimurium* strain was used to determine that *lgt* is similarly situated upstream of *thyA* as in *E. coli* [89]. Early studies looked at the genetic context of *lgt* but due to the available techniques and resources the results were limited. The -10 site was found but not the -35 site. *ptsP* was described as upstream of *lgt* and the intergenic region between *ptsP* and *lgt* was predicted to have an extensive secondary structure but no functional significance could be deduced. *ptsP* is not transcriptional linked to *lgt* unlike *thyA.* When *lgt* was partially deleted but the promotor sequence remained intact in a strain with a complementing *lgt*, ThyA production was dramatically increased suggesting *lgt* plays a role in reducing excess *thyA* expression [76].

The *lgt* gene from various other species were later described. In the Gram-positive *S. aureus* and *E. faecalis*, *lgt* is flanked upstream by *hprK* [90, 91]. Differences are found downstream with an *O*-acetyltransferase present in *S. aureus* [91] and *gpsA* and *galU* in *E. faecalis* although no comment on co-transcription is made [90]. In *S. aureus* it is thought to form a three-gene operon. It has been proposed that HprK may dephosphorylate PG and provide the DAG for the reaction [91]. However, *in vitro* activity assays show PG as the primary substrate for *E. coli* Lgt [81]. Lgt is not essential in Gram-positive bacteria. In a strain of *E. faecalis* containing a stop codon in *lgt*, no changes in the transcription of nine selected lipoproteins were observed, suggesting that lack of Lgt does not have downstream effects on the regulation of lipoproteins [90].

It is still unknown whether these genetic arrangements are true across a broader range of organisms or across pathogenic bacteria. Understanding the synteny of *lgt* may help elucidate understanding of function or regulation of the enzyme.

Peptide substrates

It is unclear how Lgt recognizes and binds the substrate lipobox and if it is indeed the lipobox that is recognized. The signal sequence of lipoproteins was found to have a majority hydrophobic peptide with a positively charged 'n-region' of 5-7 residues at the N-terminus and hydrophobic 'h-region' before the well characterized lipobox [92, 93]. The lipobox, as discussed in our review article is composed of four residues: $(LV)^{-3}(ASTVI)^{-2}(GAS)^{-1}(C)^{+1}$. A tailored lipobox for Gram-positives was proposed from data acquired from *L. monocytogenes* lipoproteins; $(L)^{-3}(SA)^{-2}(AG)^{-1}(C)^{+1}$ [94] which encompasses the same characteristics as those proposed previously. Studies have shown that replacement of the conserved cysteine led to the accumulation of unmodified Lpp [75, 95].

Beyond the lipobox, a $G_{14}D$ mutation in the signal sequence of the lipoprotein Lpp was characterized and Lpp was found to have a non-modified cysteine and to have retained the signal peptide. Lpp is covalently bound to PGN under normal conditions but a reduction in the PGN bound form of Lpp was observed in the mutant. Lpp was still located in the OM but was significantly increased in the IM [96-98]. Lpp ΔG14 showed temperature sensitivity when expressed on a plasmid in a Δ*lpp* background by failing to grow at non-permissive temperatures and contrary to Lpp G14D, accumulated glyceryl modified Lpp. It was also shown to be a poorer substrate for Lsp. It is noted that the mutant is still functional and mature Lpp was detected, albeit at a slower rate [99]. G14D mutations have also been observed as a resistance mechanism to Lsp inhibitor globomycin and G0507, a LolCDE inhibitor [74, 100]. Although an interesting avenue to explore, the role of G14 has not been tested further *in vitro* to assess whether the mutation effects Lgt activity directly of it the lack of modification is due to other factors.

For use in *in vitro* assays of Lgt, the first 24 amino acids of Lpp were made synthetically and were confirmed as a substrate for the reaction [101] suggesting that the mature region of the protein is not important for the Lgt reaction. A less hydrophobic peptide was also found to be a substrate *in vitro* [93]. However, comparisons between mature proteins and truncated forms have not been completed. We cannot therefore say whether aspects of the mature form affect or not the reaction kinetics of Lgt.

Lipid substrates

A central cavity within the enzyme containing a hydrophobic base and a positively charged region above has been speculated as being a binding site for negatively charged phospholipids [77] (Figure 8). Phosphatidylglycerol (PG) has been shown to be the primary phospholipid substrate for Lgt with specificity for the negatively charged head group the likely reason [81, 101]. Phosphatidic acid (PA), phosphatidylserine (PS) and cytidine diphosphate diacylglycerol (CDP DG) are moderate/weak substrates, whereas phosphatidylethanolamoine (PE) and cardiolipin (CL) are not substrates [101] (Figure 5, Table 4). The lipid composition has been described briefly and palmitoyl-oleyl-PG (POPG) is a better substrate than dipalmitoyl-PG (DPPG) but this assertion is based on visual interpretation of a shift in mobility of the peptide substrate on SDS-PAGE that appears to show little or no difference [81] (Table 4). Although PG is the main substrate, it is has been shown that cells lacking PG synthesis machinery and therefore PG itself, are still viable but only in the absence of Lpp [102]. As Lgt is considered essential and PG its primary phospholipid substrate it raises questions of Lgts ability to utilize other phospholipids. However, as viability is only possible in the absence of the highly abundant Lpp, this suggests alternative phospholipid substrates are not wholly sufficient to compensate for PG when this most abundant lipoprotein is present. It does however suggest enough lipoproteins may be modified to enable growth in the absence of PG. This reinforces the crucial role of PG as the lipid donor for Lgt diacyglyceryl transfer.

Table 4. Phospholipid substrates of Lgt

Phospholipid	Substrate	Reference
Phosphatidylglycerol (PG)	best	[81, 101]
POPG	best	[81]
DPPG	better	[81]
Phosphatidylethanolamine (PE)	Not a substrate	[81, 101]
Cardiolipin (CL)	Not a substrate	[81, 101]
Phosphatidic acid (PA)	Moderate	[81, 101]
DPPA	Moderate	[81]
Phosphatidylserine (PS)	Moderate	[81]
DPPS	Moderate	[81]

Although the head groups of phospholipids have been analysed, little work has looked at the role of lipid length and degree of saturation on substrate specificity in Gram-negative bacteria. DPPG was shown to stabilize the enzyme more than POPG, in thermodynamic experiments, lyso-PG also

stabilized the protein and the authors concluded that phospholipids with saturated acyl chains are better at stabilizing Lgt [81].

Some mass spectrometry analysis of lipoproteins in *Mycobacterium bovis* revealed the acylation states of sn-1 as C16, of sn-2 as C16, C18 or C19 and C16/C19 as the N-acylation after complete maturation by Lnt [103, 104]. Tuberculostearic acid (C19/0) is observed as ester-linked to the lipoprotein. Phospholipids in *Mycobacteria* are usually C16:0, C16:1, C18:1 and C19:0 [105]. Analysis of head-group specificity has not been conducted in *Mycobacterium spp*.

Essentiality of Lgt

Whole genome studies

Three global genetic screens of *E. coli* using different methods have concluded that *lgt* is an essential gene under laboratory conditions. The first is the single gene knock-out library, known as the Keio collection, which demonstrated *lgt* essentiality through an inability to replace the gene with a resistance cassette [23]. A random transposon library which inserts a transposon into random locations in a genome and is later sequenced to allow determination of insertion site also showed through a lack of viable insertions in the *lgt* gene that it was essential [24]. Finally, a CRISPRi gene silencing library also showed that *lgt* could not be silenced and therefore was likely essential [25]. No successful deletion has been done in Gram-negative bacteria.

Species specific essentiality

In firmicutes, Lgt has successful been altered or removed from *E. faecalis* [90] and multiple *Streptococcus* species including *S. mutans, S. agalactiae, S. uberis* and *S. pneumoniae [106-108]. S. mutans* Δ*lgt* delayed growth compared to the wild type [106]. In *S. agalactiae*, deletion of *lgt* had no effect on growth in complex media but showed reduced growth in minimal media [108]. Whereas *lgt* deletion had no or little effect on growth in rich media for *S. pneumoniae* but a reduced survival in blood [109, 110]. In *L. monocytogenes*, removal of *lgt* had no effect on cell or colony morphology or growth in rich media but a slight attenuation of growth was seen in minimal media [94]. A similar effect on growth was observed in a *S. aureus lgt* mutant [91]. Lgt in *Streptomyces scabies* is non-essential, but deletion did effect colony morphology which was only partially restored when *lgt* was complemented *in-trans* [85]. These data suggest Lgt is not essential in Gram-positive bacteria.

In actinobacteria, deletion of *lgt* from *Corynebacterium glutamicum* revealed no changes to colony morphology, cell morphology or susceptibility to antibiotics. Albeit, there was a very slight growth delay in the absence of *lgt* [111]. In a closely related organism, Lgt was also deleted from *M. smegmatis* but more severe effects on growth were observed alongside changes in colony morphology. Interestingly, Lgt could not be deleted from *M. tuberculosis* [83].

To date, Lgt has only been shown essential in Gram-negative bacteria and *M. tuberculosis*. Although growth defects have been observed in Gram-positive bacteria that lack Lgt, particularly under minimal nutrient conditions, Lgt does seem to be dispensable under laboratory conditions.

Essential residues for Lgt activity and function

The majority of experimental studies assessing key residues in Lgt were performed in *E. coli* or *S. enterica* Typhimurium due to the essential nature of the protein in Gram-negative bacteria under laboratory conditions. A table highlighting functional and non-functional mutations can be found in Table 5. Initially, Lgt habouring mutations $W_{25}R$, $G_{104}S$, $L_{139}F$ and $D_{249}N$ were shown to be defective in lipoprotein modification in a temperature sensitive mutant [77] (Figure 7). Diethyl pyrocarbonate (DEPC), which covalently modifies histidine, lysine, cysteine, and tyrosine residues, was found to cause a reduction in Lgt activity suggesting DEPC-modified histidine, tyrosine or lysine residues were key for activity as Lgt does not contain cysteine [77]. H_{103} and Y_{235} were found to be essential for complementing the temperature sensitive mutant and in an *in vitro* activity assay. The authors concluded that H_{103} was likely the key catalytic residue [82]. Further site-directed mutagenesis by alanine substitution of highly conserved residues was conducted to assess their role in Lgt function. Models used to study Lgt then moved away from the temperature sensitive mutant to a more malleable *E. coli* depletion strain (Figure 7). This strain (PAP9403, Δlgt^{P}) contains an *lgt* deletion on the chromosome made possible by the introduction of a primary vector (pBAD18) that contains an arabinose inducible wild-type *lgt* gene [80]. To analyse complementation ability of mutated *lgt*, a secondary vector (pAM238) containing mutated *lgt* variants, differentially inducible by IPTG, was added. Although Lgt protein production from the pAM238 expression vector varied for different mutants, conserved $R_{143}A$, $E_{151}A$ and $E_{243}A$ could partially restore growth. Growth was observed for $G_{98}A$, $G_{104}A$, $D_{129}A$ and interestingly, $H_{103}Q$, suggesting these residues are not essential for activity. $Y_{26}A$, $N_{146}A$ and $G_{154}A$ were not functional (Table 5), although the latter two showed low levels of production from the pAM238 vector. It should be noted that wild type Lgt showed low level production also suggesting high concentrations of Lgt are not required for viability [80]. The largest analysis of conserved residues was conducted by Mao *et al.* (2016) who completed site directed mutagenesis on dozens of residues (Table 5). Although comprehensive, protein production is not shown for all the mutants and therefore some doubt is raised as to whether mutated genes are being expressed or if the proteins are unstable. Once again, H_{103} was shown to be essential, adding greater interest to this residue. Y_{235} was essential in accordance to previous reports.

Table 5. Functional and non-functional mutations of E. coli Lgt.

	Mutation			Mutation	
Residue	**Functional**	**Non-functional**	**Residue**	**Functional**	**Non-functional**
W25		R[3]	**W153**	A[4]	
Y26	F[2]	A[1,2], Q[4]	**G154**		A[1,4]
G27		L, W, Q [4]	**R155**	A, Q [4]	
Y30	F[4]	A[4]	**D157**	A, N [4]	
E56	Q[4]		**E172**	Q[4]	
G64	C[4]		**H196**	A[4], Q[1,4]	
Y80		F, Q [4]	**P197**		A[4]
D88	N[4]		**Y201**	F[4]	
D97	N[4]		**E202**	A, Q, L [4]	
G98	A[1], P[4]		**Y235**	S[2]	F[2,4], S[4], T[2]
G99	P[4]		**R239**	A[1]	A[1,4], K, H, Q [4]
M100		W[4]	**E243**	A[1]	A[1,4]
S101		D[4]	**D249**		N[3]
F102		A, W[4]	**R246**		A, K, H ,Q, E [4]
H103	Q[1]	Q[2,4], N[2,4], D[2], R[4], Y[4]	**P248**	A[4]	
G104	A[1]	S[4]	**A250**		L[4]
L106		W[4]	**F252**	L[4]	
I110	W[4]		**T253**	L[4]	
D129	A[1,4], N[4]		**Q258**	A[4]	
G138	A[4]		**Y259**	A, E, F[4]	
L139		F[3]	**S261**	A[4]	
G142		A, I [4]	**M262**		Q, Y [4]
R143	A[1]	A[14]	**G263**		A, V, L [4]
G145		V, I [4]	**Q264**	A[4]	
N146		A[1,4]	**P269**		A[4]
E151	A[1]	A[1,4], Q[4]			

[1][80], [2][82], [3][77], [4][81].

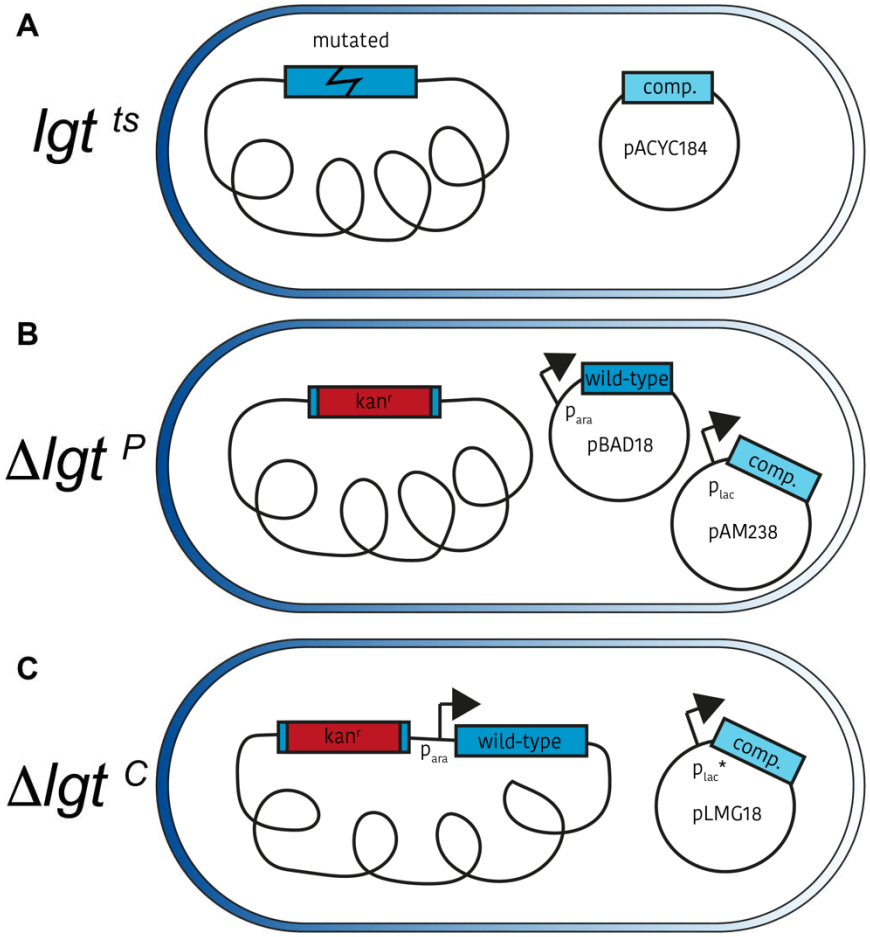


Figure 7. Models of strains to study Lgt function. *A) Temperature sensitive strain [77] with a mutated lgt gene and lgt restored on a complementing pACYC184 plasmid . B) PAP9403, lgt is replaced on the chromosome by a kanamycin resistance cassette (kan^r) and wild type lgt is provided on a pBAD18 plasmid induced by L-arabinose and repressed by D-glucose, a secondary vector containing complementing lgt is present on a pAM238 plasmid, under the control of a P_{lac} promoter. [80]. C) lgt is removed from the chromosome in the same manner as B) but wild type lgt is restored on the chromosome with the same promoter system as B by insertion at the lambda phage integration site (λatt)[75], a complementing lgt gene is present on a pLMG18 plasmid under the control of a P_{lac} promoter.*

Conditional lethality depletion studies

Some work has been carried out which assesses the effects of reducing the levels of Lgt in depletion strains. Initially, work was carried out on a temperature sensitive *lgt S. typhimurium (lgt^{ts})* mutant which deemed Lgt essential as growth did not occur at the non-permissive temperature. A morphology of rounded, oval shaped cells and cell lysis was described (but not shown) at non-permissive temperatures [89]. Another strain was developed whereby *lgt* is deleted from the chromosome and restored on a plasmid under the control of a L-arabinose inducible promoter in *E. coli* K12 BW25113, a wild-type strain genetically altered to not metabolize arabinose (Δ*lgt*^{P}, Figure 7). Through this work it was confirmed that Lgt was essential for growth and that reducing levels of Lgt led to rounding of the cell poles in early exponential phase and DNA leakage occurred from the cell pole [80]. A chromosomal strain was more recently developed where endogenous *lgt* was replaced by a kanamycin resistance cassette and a complementing single copy *lgt* was inserted onto the chromosome of *E. coli* under the same promotor control as described above [75] (Δ*lgt*^{C}, Figure 7). This was produced in *E. coli* K12 MG1655, a laboratory wild-type strain which maintains its ability to metabolize arabinose and CFT073, a clinical isolate causing urinary tract infections. The authors describe that *lgt* is required for growth and this requirement is still needed when downstream enzymes Lsp, Lnt and LolCDE are overexpressed suggesting the pathway is not functional without Lgt. They found that around 25% depletion was sufficient to induce loss in cell viability and more modest levels of depletion increased susceptibility to serum killing and increased membrane permeability. They also describe an increase in cell area upon Lgt depletion and the Minimal Inhibitory Concentration (MIC) for multiple antibiotics is decreased.

The role of Lpp in Lgt essentiality

Lpp was the first lipoprotein discovered [112-114]. Interestingly, in the temperature sensitive *lgt* strain, growth was restored when *lpp* was deleted suggesting *lgt* was only essential in the presence of Lpp and therefore Lpp was a major factor in the essentiality of Lgt [89]. In the chromosomal depletion strain, Δ*lgt*^{C}, Diao *et al.* (2021) state that *lpp* deletion does not increase survival even at low concentrations of Lgt, contrary to the findings by Gan *et al*. (1993), but in fact *lpp* deletion has deleterious effects on cell viability when *lgt* is depleted. A result that was surprising, given that Lpp deletion is sufficient to rescue growth of Lsp, Lnt and Lol depletion or inhibition [64, 65, 67, 72].

Inhibition of Lgt

Diao *et al.,* (2021) reported an inhibitor of Lgt (G2824) discovered through the generation of a macrocyclic peptide library. The authors demonstrate that their inhibitor was sufficient to block Lgt activity in an *in vitro* activity assay (described below) and showed a lack of modification of the peptide

substrate, when analysed by Mass Spectrometry, in the presence of the inhibitor. G2824 altered Lpp processing and led to a modest increase in cell area. G2824 was more effective in the absence of Lpp, contrary to observations that Lpp deletion reduces the efficacy of inhibitors to Lsp and LolC or LolE [67, 72]. This result is indeed surprising and they note that there is reduced accumulation of peptidoglycan-bound lipoprotein precursors when Lgt is inhibited, compared to inhibition of Lsp which sees accumulation of the bound-form of Lpp and speculate this may be why these unanticipated effects are observed. Although not discussed or explored, it may be that the unmodified lipoprotein with its signal peptide still attached may not be able to form the trimeric form possibly preventing covalent attachment to PGN by L-D-transpeptidases.

It is also possible that although Lgt may be inhibited by G2624, there are also other targets within the cells causing multiple effects not drawn out by the current research and therefore some doubts are raised as to whether Lgt is the sole target. It should be noted that *E. coli* CFT073, the primary strain used in this study, is not directly comparable to *E. coli* MG1655 due to the presence of a pathogenicity island [115]. It may also be possible that stress responses are triggered in the presence of the inhibitor, helping to overcome some phenotypic effects.

This study also looked at the effects of reduced Lgt on the MICs of a small range of antibiotics. In each instance, reduced *lgt* expression led to a decrease in MIC. This suggests that Lgt could be a good target when combined with other antibiotics as reduced (not removed) Lgt is sufficient to make the cell susceptible to other therapies. However, the reasons why reducing Lgt would lead to increased susceptibility is yet to be discussed.

Functional conservation

A few studies have explored the functional conservation of Lgt. Initially it was shown that *S. aureus lgt* complements a temperature sensitive *E. coli* mutant [77]. In the same year a temperature sensitive *Salmonella* strain was complemented by *lgt* from *E. coli* and vice versa [76]. More recently it has been shown that *lgt* from *P. aeruginosa* and *A. baumannii* can complement growth of an *E. coli* mutant generated in CFT073 [75].

In Actinobacteria, *E. coli lgt* could complement a *C. glutamicum* strain demonstrated by the association of the lipoprotein MusE to the membrane. As mentioned above, *M. smegmatis* has two alleles of Lgt. *M. smegmatis* MSMEG_3222 (with and without its extended C-terminus) could complement *C. glutamicum* but MSMEG_5408 could not, likely due to the fact the MSMEG_5408 has multiple

mutations in key residues [83]. Interestingly, *C. glutamicum lgt* could not complement the *E. coli* depletion strain PAP9403 [111]. Lgt from *M. tuberculosis* was able to complement an *M. smegmatis lgt* deletion strain [83].

In firmicutes, the two *lgt* genes from *S. coelocolor* complemented an *lgt* mutant of *S. scabies* and saw partial restoration of colony morphology [85].

However, much more work is needed to explore the functional conservation of Lgt in different backgrounds and elucidate information regarding species specificity. A systematic approach with well defined models would enable better comparisons between Lgt from different organisms.

Structure of Lgt

The cellular location of Lgt has been debated in the literature. Solubilisation in salt-containing buffer and activity of Lgt in water soluble fractions has led one research group to hypothesise that Lgt is merely associated with the cytoplasmic leaflet of the inner membrane [116, 117]. However, *in silico* analysis by eight prediction tools predicted multiple transmembrane domains (TM). This work was followed up experimentally by the inability to solubilize the protein in high salt concentrations but solubilization in the presence of detergent n-octyl b-D-glucoside (OG). It was also shown that Lgt was located in the inner membrane by sucrose floatation gradients [80]. Furthermore, this study showed that Lgt had seven TM domains by using multiple *lac*Z and *pho*A translational fusions. Alkaline phosphatase (PhoA) is only active in the periplasm and β-galactosidase (LacZ) in the cytoplasm. The experiments were completed by Substituted Cysteine Accessibility Method (SCAM) analysis that is based on the accessibility of cysteine residues to alkylating reagents. Finally, an X-ray crystal structure was solved and the authors concur that Lgt is an integral membrane protein [81].

Two structures of Lgt with lipid substrates were solved, form-1 and form-2 (Figure 8, Appendix III). Lgt has two 'arm' loops that protrude from the body of the enzyme, along the surface of the phospholipid membrane. Seven transmembrane (TM) helices create the membrane domain which itself appears to have two domains, a large and small TM domain. Inside this TM domain, a central cavity is formed. Above the membrane is a large head group, protruding into the periplasmic space. The notable difference between the two structures (form-1 and form-2) was a change in the loop between TM6-7 (L6-7), increasing the size of the central cavity. Computational modeling of Lgt revealed that L6-7 has structural flexibility and when two PG molecules are modeled inside the enzyme, the loop moves to a more open conformation [79]. The authors propose a role of L6-7 in moving to create space for the PG head group (Figure 9). The H_{103}GGL motif is just below the periplasm-membrane interface after the arm-2 loop and this region also showed some flexibility [79] (Figure 21). The Lgt signature motif is embedded in the central cavity, at the periplasmic half of the membrane.

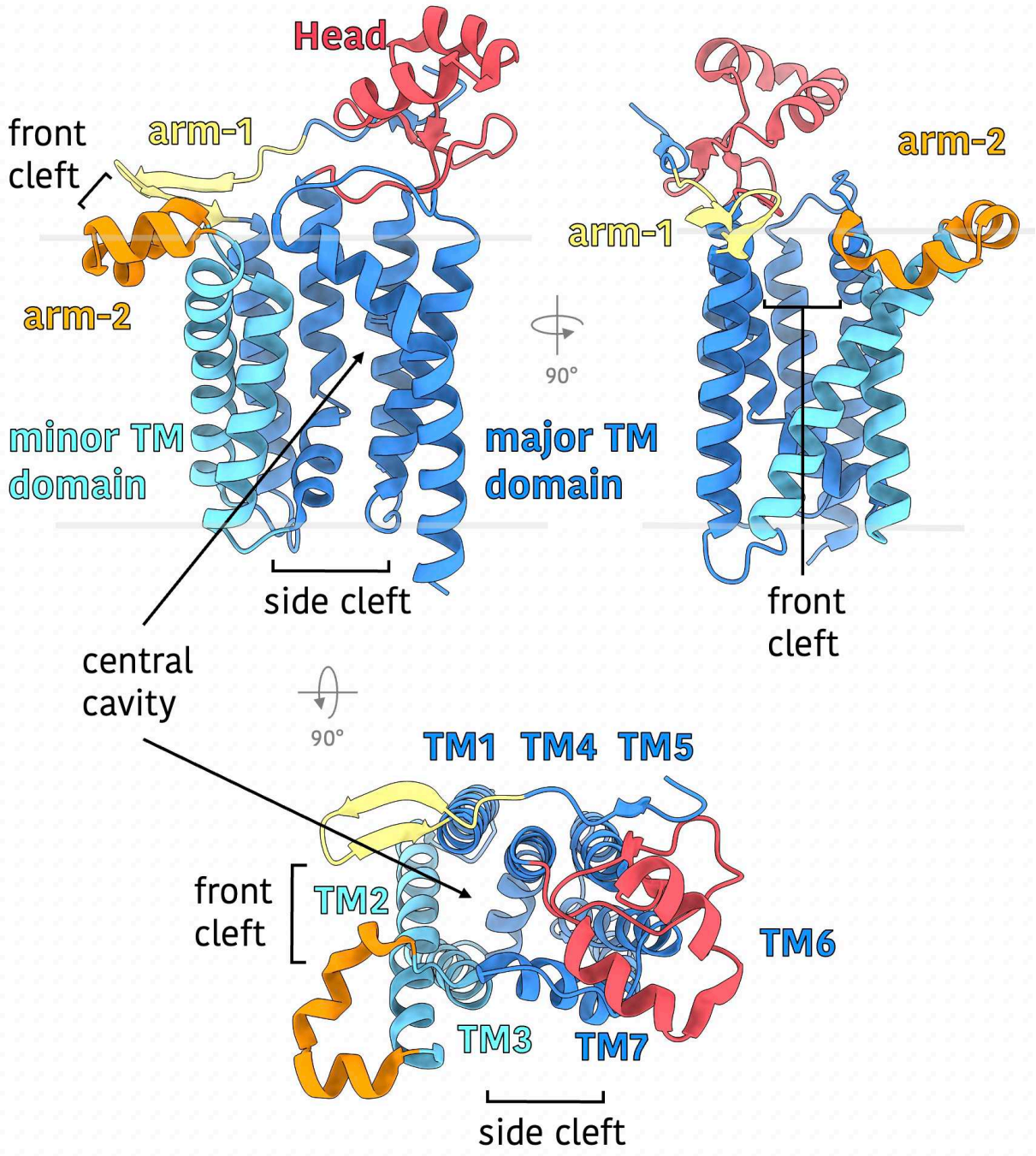


Figure 8. Structure of Lgt. *X-ray crystal structure of Lgt key regions of the protein. Structure from Mao et al. (2016). Grey lines indicated upper and lower bounds of the phospholipid bilayer.*

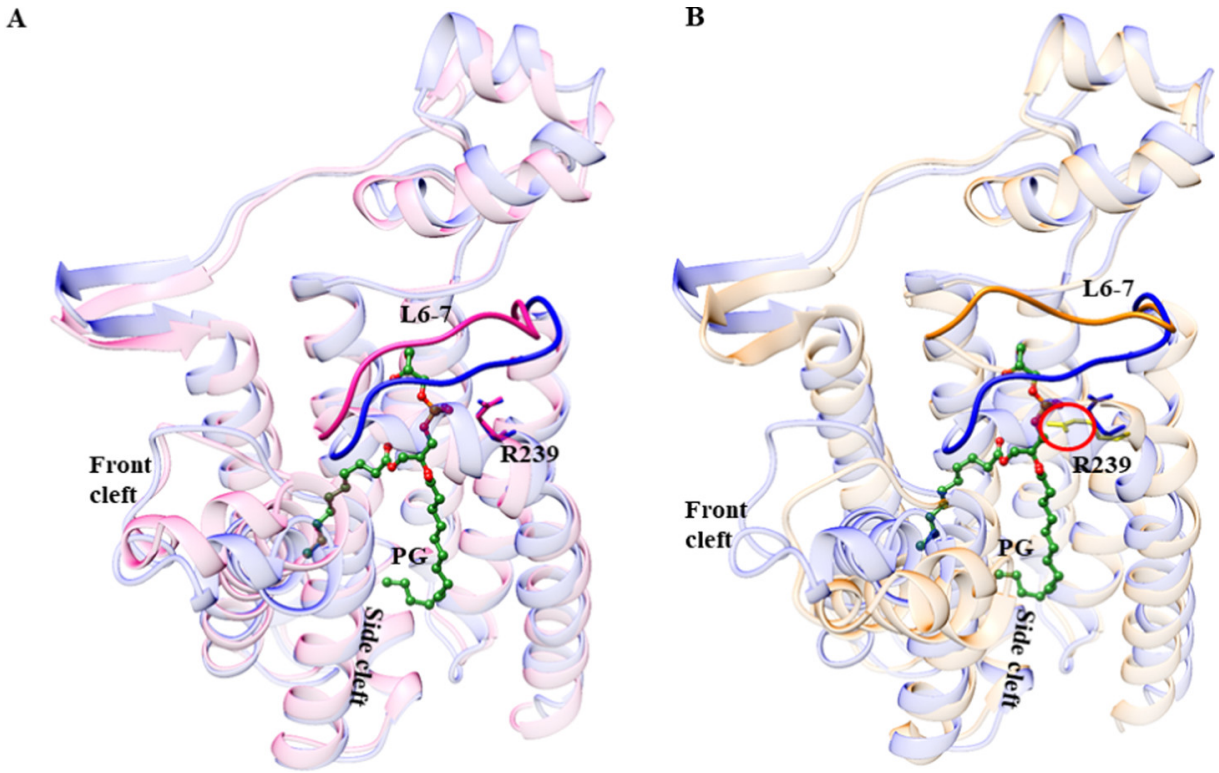


Figure 9. Flexibility of L6-7 and arm-2. *A) movement of L6-7 and R239 in the presence of PG in the central cavity. B) movement of L6-7 and R239 in the presence of PG and cysteine containing lipobox. Pink = solved structure of Lgt, blue = with PG, orange = with PG and lipobox. The red circle highlights the movement of R239 in the presence of the lipobox substrate. Image from Singh et al. (2019)*

Catalysis

Initially it was proposed that PG substrates move through the enzyme from the front cleft into site-1 before sliding into the catalytic cavity of the enzyme (site-2) (Figire 8) [81]. Molecular modeling and simulations have since shed some light on a possible mechanism [79].

It was observed in models of Lgt with two PG molecules bound that in site-1 PG forms H-bonds with N_{149} and G_{151} and when L6-7 is in an open position the PG moves towards the central cavity. R_{239} in this site turns towards the head-group of PG stabilizing it in site-2 (Figure 9). S_{198} turns towards PG and G_{202} and R_{246} further stabilize the head-group via H-bonds and R_{143} and R_{239} are reorientated, preparing the active site for catalysis. Adding a peptide into the model caused G_{243} to turn away from R_{239} and R_{238} turning toward the phosphate group of PG. The authors propose R_{239} as a gate that is open in the absence of a lipobox peptide but when the peptide binds, the gate closes and L6-7 opens, allowing the reaction product out of the side cleft and new PG substrate to enter through the front cleft (Figure 9).

In the catalytic cavity the authors simulate a reaction with residues they determined as key. H_{103}, R_{143}, R_{146}, Y_{235}, R_{239} along with a cysteine from the lipobox where chosen. They observe in their simulations that H_{103} forms an H-bond with the thiol group of the cysteine and the C3-O ester bond forms an H-bond with R_{143}. H_{103} acts as a catalytic base, abstracting a proton from the cysteine which then undergoes nucleophilic attack on the C3-O carbon of PG, forming diacylglyceryl-cysteine and releasing glycerol-1-phosphate (G1P) (Figure 10).

An issue raised is that H_{103} is not entirely conserved (Table 3) and therefore adds doubt to the possible mechanism of the reaction. The authors do not perform the modelled reaction with other residues common at residue 103, such as tyrosine. The peptide used in the modeling is also shorter than a standard signal peptide and the additional residues of a real peptide may induce further conformational changes. Therefore, although this model may be correct, there are still unanswered questions as to whether we fully understand the catalytic activity of Lgt.

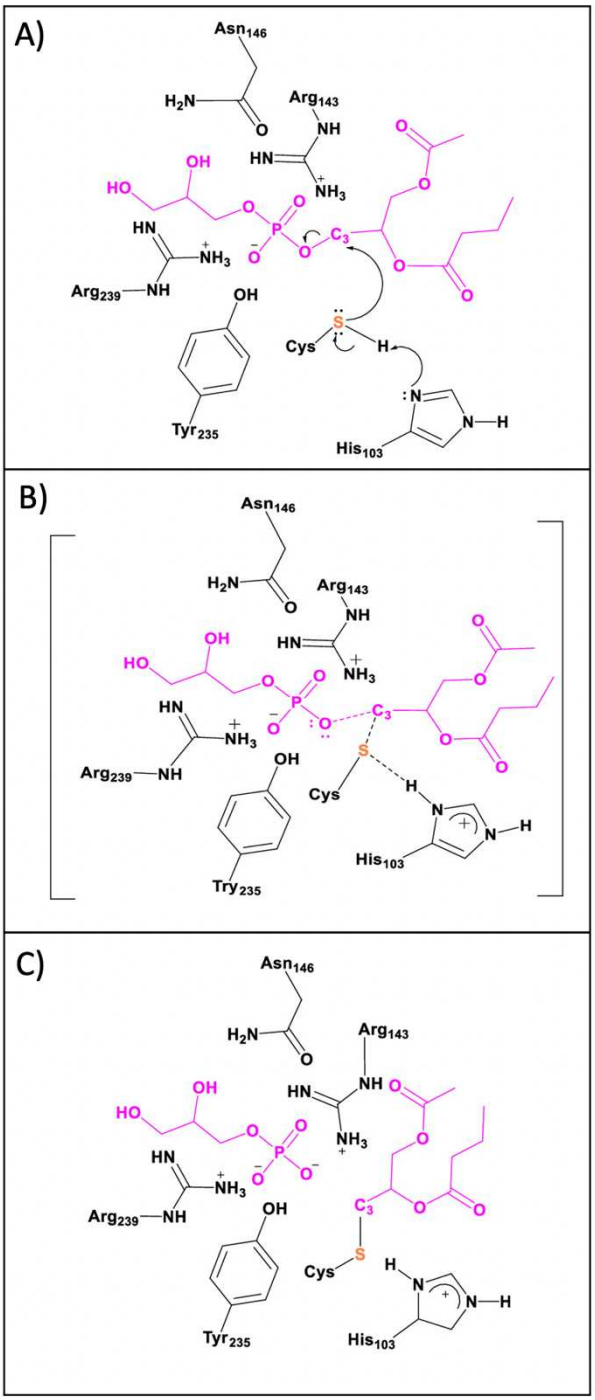


Figure 10. Predicted reaction catalysed by Lgt. *Side chains from Lgt are depicted in black, PG is shown in pink and the sulfhydryl group of the lipobox cysteine is shown in orange. A) Pre-reaction positioning of residues. B) representative transition state. C) completed reaction. From Singh et al. (2019)*

In vitro activity of Lgt

Gan *et al.* (1995) observe that although the predicted molar mass of Lgt is 33 kDa, they found it to be 27-30 kDa based on mobility on SDS-PAGE through plasmid-encoded expression, and Williams *et al.* (1989) observed a 25 kDa band corresponding to Lgt. It was suggested that the acrylamide concentration and high number of non-polar residues (61%) in Lgt may be the cause of this observed discrepancy in size. Differences in migration for membrane proteins are common and is noted in the literature [118]. Williams *et al.* (1989) also determined that the protein was present in the membrane fraction of the cell as opposed the soluble fraction. Several assays have been used to study Lgt activity and are summarized in Figure 11.

Radiolabeled Lpp - gel-shift assay

The first *in vitro* activity assay of Lgt used a lipoprotein peptide of Lpp, synthesized by *in vitro* protein synthesis. This process incorporated radiolabeled methionine, an amino acid present three times in Lpp. Inverted membrane vesicles (IMVs) were prepared from *E. coli* wild type cell cultures and these IMVs were incubated with the peptide for one hour. The reaction was stopped by the addition of an SDS-based buffer and boiling for five minutes. The samples were loaded onto Tricine-SDS gel and separated by electrophoresis. The bands were revealed by autoradiography. Due to a change in mass between the lipidated and non-lipidated peptides, a shift in migrations indicated Lgt activity [89] (Figure 11A).

Radiolabeled phospholipids – peptide assay

Another *in vitro* activity assay used to study Lgt kinetics incorporated radiolabeled phospholipids, as opposed to a radiolabeled peptide. Either [9,10-^{3}H] palmitate-labeled phospholipid [77, 82] or [2-^{3}H] glycerol labeled phospholipid [101] were used. These phospholipids were incubated with either IMVs, membranes solubilized in detergent or water-soluble fractions derived from IMV preparations, and an Lpp based peptide. The reaction was quenched and the peptide precipitated by ammonium sulfate and acetone, and dissolved in SDS. The collected peptide was analysed by scintillation counting and increases in counts relative to negative control indicate incorporation of diacylglceryl onto the peptide and therefore Lgt activity (Figure 11D).

Radiolabeled phospholipids – paper-shift assay

Another application of radiolabeled phospholipids used a similar method as described above whereby [9,10-^{3}H] palmitate-labeled phospholipid was incubated with an Lpp derived peptide. The reaction products were then separated by paper electrophoresis based on changes in charge of modified and unmodified peptide. The radioactivity was measured by scintillation counting [93] (Figure 11C).

Fluorescence – gel-shift assay

A peptide substrate Lipo-GFP, containing the N-terminal sequence of Lpp fused to GFP, was used as Lgt substrate [81] in an alternative gel-shift assay. This peptide was produced in *E. coli* as a glutathione-S-transferase (GST) fusion protein for purification purposes and after cleavage of GST used as substrate in the Lgt reaction. Upon incubation with commercial phospholipids and detergent purified enzyme, formation of diacylglyceryl-lipoGFP is followed by a shift in migration on SDS-PAGE and fluorescence detection (Figure 11B). However, lipoGFP is a large protein and the addition of diacylglyceryl leads to a relatively small change in mass, and therefore a small shift in migration.

Coupled-enzymatic reactions

A coupled enzymatic reaction [117] monitors Lgt activity through the formation of glycerol-1-phosphate (G1P), a by-product of the reaction directly correlated with enzyme activity. G1P is cleaved by alkaline phosphatase to produce glycerol and phosphate of which the glycerol is then dehydrogenated in the presence of NAD^+ and glycerol dehydrogenase (GDH) to produce dihydroxyacetone (DHA) and $NADH^+$ and H^+. In a final step, with $NADH^+$ now present, diapharose catalyses the reduction of resazurin to resorufin, producing a measurable fluorescent signal proportional to Lgt activity (Figure 11Ei). This method was used with IMVs or water eluted protein.

However, to improve this assay it was later divided into a two-step reaction [78]. The reaction was halted after the initial Lgt reaction and restarted in the presence of the coupled enzymes. This method was used to study Lgt activity in IMVs of enteropathogenic *E. coli* (EPEC), *Lactobacillus sp., Lactococcus lactis, Samonella Typhi, Shigella flexneri* and *S. aureus*. Although they all exhibited Lgt activity, *L. lactis* had considerably higher activity. This was true when Lgt was purified in a single-step by cation exchange chromatography where purified Lgt from *L. lactis* had considerably higher activity than that of *E. coli* Lgt. The authors observe that *L. lactis* Lgt required detergents to be solubilized contrary to their previous (and subsequent) reports that *E. coli* Lgt was peripherally associated with the membrane. They do not use antibody detection or peptide tags to visualize and verify the protein purification but only Coomassie Brilliant Blue (CBB) stained gels and detect a 31 kDa band as evidence for Lgt presence, but we have seen that Lgt migrates corresponding to a variety of different masses depending on the electrophoresis conditions. After purification Banerjee *et al.* (2013) observe a single protein band and this fraction contains Lgt activity as measured by the described coupled reaction suggesting that this method of purification is viable.

A similar coupled reaction was used by Diao *et al.* (2021) to assess the effect of possible inhibitors whereby they measure glycerol-3-phosphate (G3P) release using a G3P-PG as substrate. G3P dehydrogenase (G3PDH) catalyses the production of dihydroxyacetone phosphate (DHAP) and NADH,

NADH reacts with proluciferin in a reaction catalysed by proluciferin reductase (PR) to produce luciferin. Luciferin in turn is converted into a fluorescent signal by luciferase (Figure 11Eii). This assay uses detergent purified Lgt.

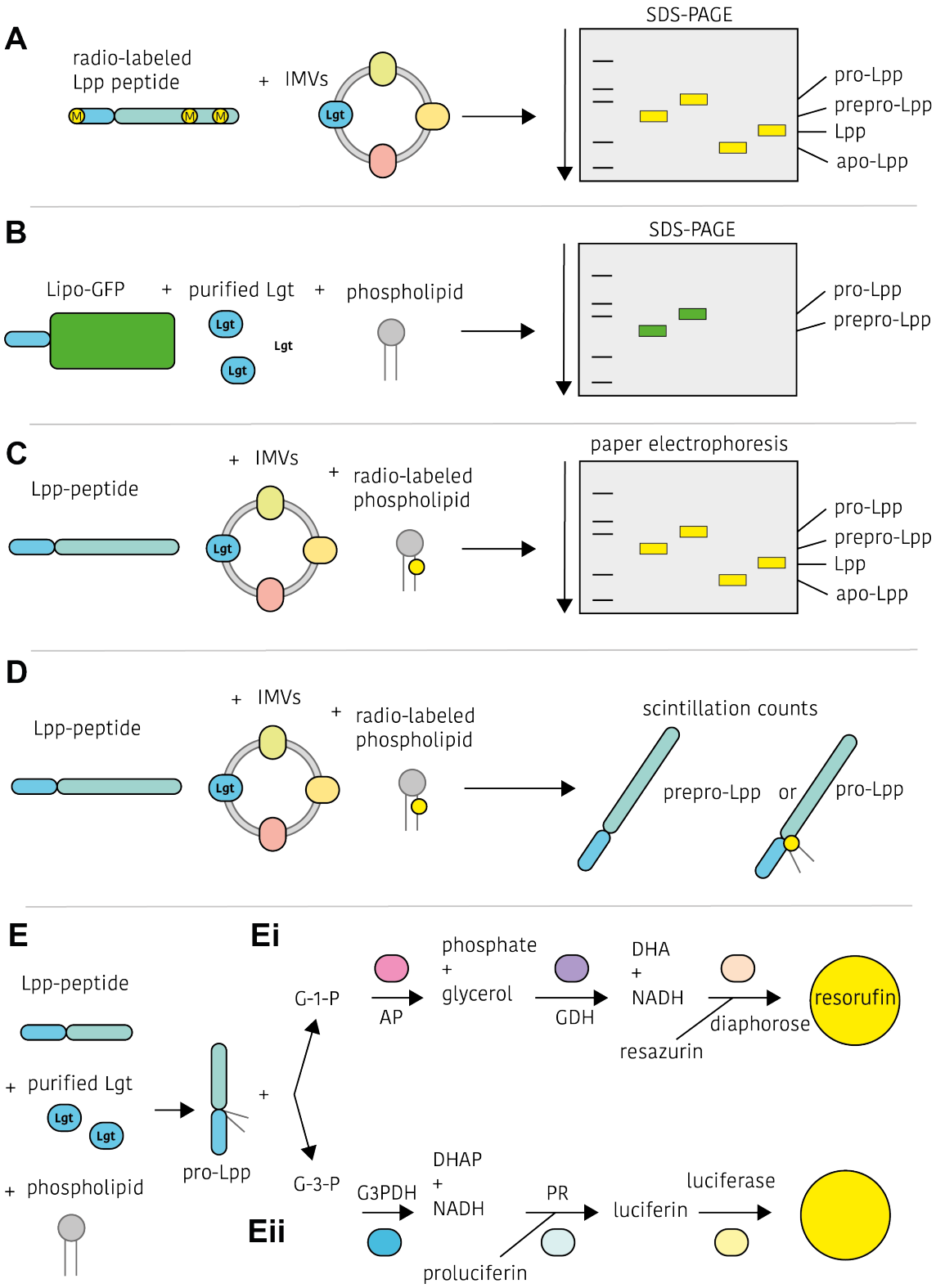


Figure 11. In vitro activity assays of Lgt. *Detailed descriptions can be found in the text.*

Pros and cons of different in vitro activity assays

Each assay has its draw backs and advantages. The coupled enzymatic reaction is complex and the addition of more enzymes to the reaction could interfere with the Lgt reaction. However, it does produce a quantitative read-out and does not require SDS-PAGE. The other assays rely on migration shifts on SDS-PAGE which is slow and difficult to quantify or on radio-labeled phospholipids and peptides which may constitute unnecessary risk. An ideal assay would be adaptable for high throughput, not have the risks associated with radioactivity, could be performed with unmodified substrates and would have a quantitative output.

Table 6. Comparison of Lgt kinetics.

Assay	Lgt Source	Details	Peptide		PG		Other	Ref
			Km (mM)	Vmax (pmol/ min/mg)	Km (mM)	Vmax (pmol/ min/mg)	Vmax (nmol/ min/mg)	
Radioactive phospholipid - Peptide assay	Solubilised membranes	WT	30	25	30	16 nmol	-	[77, 101]
		G104S	-	-	100	7 nmol	-	
	IMVs	WT	30	-	10	-	26	[82]
		H196Q	28	-	7	-	13	
		Y235S	67	-	30	-	18	
		WT	13	12	59	15	-	[117]
	Water-soluble	WT	10	13	52	16	-	
Single-step coupled reaction		WT	24	13	68	8	-	
	IMVs	WT	32	12	69	7	-	
Two-step coupled reaction	Single-step purification	*L. lactis* Lgt	20	-	100	-	-	[78]
Radioactive phospholipid -Paper-shift	IMVs	Soluble peptide	6	-	10	-	64 pmol	[93]
	Water-soluble		6	-	10	-	320 pmol	

However, these assays have allowed us to gain some insight into the kinetic parameters of Lgt that are highlighted in Table 6. Most of these data were obtained with enzymes in a membrane environment

and not with purified protein. The papers with detergent purified enzymes by Mao *et al.* (2016) and Diao *et al.* (2021) do not describe kinetic studies.

Lgt and virulence

The study of the role of Lgt in virulence has generally been restricted to firmicutes due to the non-essential nature of Lgt in these organisms.

At a host-cell interaction level we can observe differential effects of Lgt removal from various organisms. *lgt* mutants of *L. monocytogenes* and *S. aureus* retained their ability to gain entry in host cells [91, 94] and *L. monocytogenes* had reduced intracellular survival once inside [94]. In another study of *L. monocytogenes*, *lgt* mutants had reduced ability to invade and survive intracellularly and failed to activate the TLR2 dependent NF-κB response [119] a key signaling pathway for host immune responses. An *lgt* mutant of *E. faecalis* had greater survival in the presence of oxidative stress than a wild type *E. faecalis* [90] but the opposite was true for *S. pneumoniae* [109]. Deletion of *lgt* from *S. pneumoniae* has major impacts on the surface and exoproteome which in turn may have wide ranging effects on virulence[120]. Lipoproteins have been shown as the important component in TLR2 activation in *S. agalactiae* and interestingly, signal peptidase I was able to cleave the peptide of lipoproteins in the absence of *lgt* and *lsp* [108]. Okugawa *et al.* (2012) created a *lgt* deletion in *B. anthracis*. TLR2 response to heat-killed bacteria was reduced along with inefficient sporulation. Ultimately, there was a reduction in virulence in spores but not in vegetative cells of the Δ*lgt* strain. These mixed data suggest overall that Lgt deletion reduces the virulence in cell-culture models.

A number of studies have been carried out assessing the effect of *lgt* deletion in animal and insect infection models. Reffuveille *et al.* (2012) found that an *E. faecalis lgt* mutant killed the host more slowly in a Galleria (wax moth larvae) model. An *lgt* deletion strain of *S. pneumoniae* was shown to have attenuated virulence in a mouse respiratory tract infection model [109, 110] and *S. equi*, a pathogen of horses was shown to be attenuated by the deletion of *lgt* in an intranasal mouse infection model but interestingly not in a pony model. In the pony model, although not significant, there was a decrease in disease progression by the deletion strain [121]. *lgt* mutants in *S. agalactiae* were more lethal in neonatal sepsis mouse model [108]. In *L. monocytogenes*, *lgt* deletion attenuated infection in a mouse model [119]. In *E. coli* a depletion strain of Lgt was less adapted to survival in a mouse model [75]. A *S. aureus* Δ*lgt* strain caused a less severe septic polyarthritis in a TLR2 double knock-out mouse model compared to a wild type *S. aureus* but this effect was not seen in wild type mice [122]. Therefore, in general, Lgt mutants have attenuated virulence in animal models.

Lgt deletion in *S. scabies* had a minor effect on virulence in a potato model but this strain displayed a growth defect phenotype and therefore its reduced viability may be the cause of reduced disease progression [85].

Although it seems clear that removal or depletion of Lgt reduces virulence in animal and plant models the direct cause of this have not been explained.

Additional information regarding Lsp and Lnt

As our review article focused primarily on the classical lipoprotein modification pathway with a focus on Gram-negative bacteria some differences found in other organisms will be highlighted here.

Lsp

In *L. monocytogenes, C. glutamicum*, and *S. uberis* Lgt activity was not required for Lsp activity [94, 107, 111] which is contrary to proteobacteria where diacylation by Lgt is a pre-requisite for Lsp activity [123]. Although the pathway is not considered essential in Gram-positive bacteria, Lsp was shown to be essential for growth in the Actinobacteria *S. coelicolor* [84] raising the question as to why Lgt could be readily removed from multiple other Gram-positive bacteria but not high G-C Gram-positive bacteria.

A peptidyl-form of lipoprotein where an alanine and serine/glycine are found present in front of the normally terminal cysteine has been described in *M. fermentans* (Figure 12). As a conventional diacylated form of these lipoproteins is also present it suggests unusual cleavage activity [124]. It remains unknown which peptidase cleaves these proteins and what the possible physiological benefits may be.

Lsp has known inhibitors, chiefly the cyclic-peptides globomycin and myxovirescin [38], but these are not clinically available.

Lnt

Initially, Lnt was thought to be essential in all Gram-negative bacteria. However, increasingly, studies have found that it may be dispensable in a number of organisms such as *F. tularensis*, *N. gonorrhoeae*, *N. meningitides*, *Acinetobater. spp.* and *H. pylori* [125-128]. There are also cases where α-proteobacteria (*Wolbachia sp.*) and γ-proteobacteria (*Buchnera sp.*) lack an Lnt homolog [129, 130].

Lnt is present in *Actinobacteria* such as *C. glutamicum* [131] and *Mycobacterium sp.* [104, 132] but was shown not to be essential in *M. bovis* [104].

The largest difference between the lipoprotein modification pathway in Gram-positive and Gram-negative bacteria is that the former do not contain an Lnt homolog and it was therefore thought for a long time to be synonymous with these organisms lacking triacylated proteins [91, 133, 134]. However, more recent advances have demonstrated triacylated forms in firmicutes [135] such as *S. aureus* [136, 137], *B. subtlis* [138], *Mycoplasma* [139] and *A. maidlawii* [140]. LnsA and LnsB have been recently discovered to have N-acylation activity on *S. aureus* lipoproteins [141].

Alternative forms of modification have also been described. A lyso-form was discovered with an acyl moiety on the amino group and on the typical sn-1 position (Figure 12) [135, 142]. This transfer of the sn-2 acyl chain to the α-amino group is catalysed by Lit [1, 143]. An N-acetyl form that has an acetyl group attached to the α-amino group [135] has also been recently been discovered (Figure 12). Switching to and from diacyl and triacyl forms my be induced by environmental factors as it has been shown that ratios of the two forms fluctuate under different pH, salt, temperature, and growth phase conditions in *S. aureus* [124]. There are instances, for example, cytochrome C of *Blastochloris viridis,* which contains an Lnt homolog but the primary form of the lipoprotein is a diacylated form [144].

So far no inhibitors of Lnt have been described.

This adds to the increasingly complex nature of lipoprotein modification and demonstrates that a great deal of further knowledge is required to help our understanding of the modification process and physiological role of each form of lipoprotein.

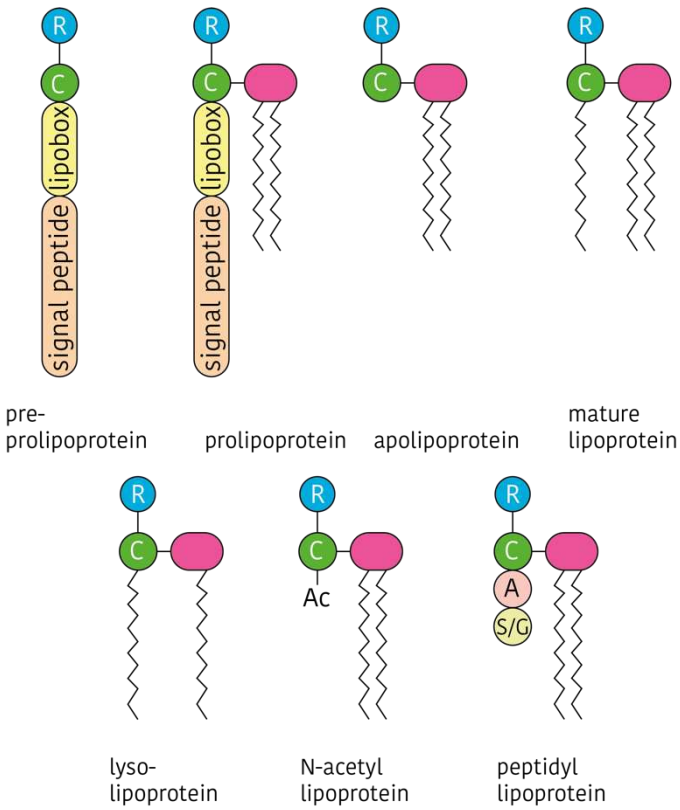


Figure 12. Forms of lipoproteins. *R = mature protein, Ac = acetyl*

Bacterial lipoproteins

The full lipoproteome of *E. coli* is yet to be characterized experimentally by mass spectrometry but a combination of experimental and computational approaches have sought to determine it. A recent study looked at the dynamics of lipoproteins in membranes via computational models and use 114 predicted lipoproteins, of which 30 have structures validated by X-ray crystallography [58]. Albeit, of the 30 structures only a handful have the lipid moiety; CusC [145], ActC [146], and LpoB [44]. The most notable evidence of proteins being fatty acid acylated is the work by Matsuyama *et al.* (2007) who discuss lipoproteins they elucidated by labeling with ^{3}H-palmitate and globomycin inhibition. From early predictions of lipoproteins, they confirm 90 proteins are lipidated in *E. coli* (Table 7) [147]. However, they did not demonstrate a non-biased proteome-wide screen but focused on predicted lipoproteins, limiting their analysis. Another example of experimental analysis was conducted by Rangan *et al.* (2010) who use clickable alkyne fatty acids to identify lipidated proteins. After click-chemistry reaction with azido-diazo-biotin the biotinylated proteins were enriched on streptavidin beads and identified by proteomics analysis [148]. The advantage of this study is the lack of bias toward prediction models. They find 44 high confidence lipoproteins and 43 medium confidence lipoproteins. However, due to the nature of the experiment, some lipid associated proteins are also extracted (Table 7).

The most commonly used prediction models are DOLOP [92] which predicts 86 lipoproteins in *E. coli* and Prosite which predicts 150 lipoproteins in *E. coli* K12 (Figure 13, Table 7). Both models are based on conserved lipobox sequences.

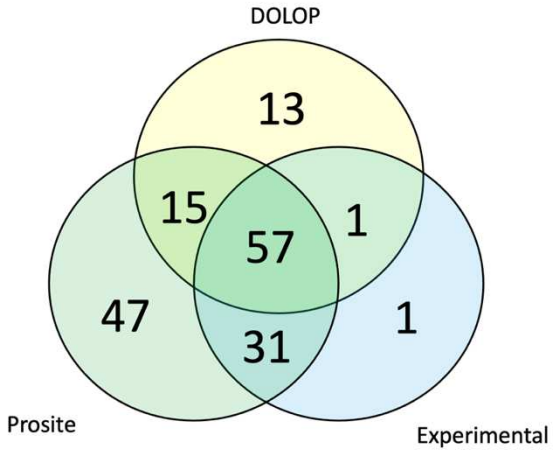


Figure 13. Comparison of lipoprotein prediction tools and experimental data. *DOLOP prediction tool [92], Prosite reference* (PRU00303) *and experimental lipoproteins determined by* [147]

Table 7. Possible lipoproteins of E. coli.

Protein	[148]	[92]	[147]	*	Protein	[148]	[92]	[147]	*	Protein	[148]	[92]	[147]	*
A2MG		X		X	**MltB**	M	X	AB	X	**YdiK**		X		X
AcfD		X		X	**MltC**			AB	X	**YeaY**	M	X	B	X
AcrA	H	X	AB	X	**MltD**		X	B	X	**YecR**		X	B	X
AcrE		X	AB	X	**NlpA**	H	X	AB	X	**YecT**		X		
AmiD		X	B	X	**NlpC**		X	C	X	**YedD**	H	X	B	X
ApbE			AB	X	**NlpD**	M		AB	X	**YegR**			B	X
BamB	H	X	B	X	**NlpE**	H		AB	X	**YehR**		X	B	X
BamC	H		AB	X	**NlpI**		X	AB	X	**YfbK**		X		X
BamD	H	X	B	X	**OsmB**		X	AB	X	**YfeY**		X	B	X
BamE				X	**OsmE**	H	X	C	X	**YfgH**		X	B	X
Blc	M	X	AB	X	**Pal**	H	X	AB	X	**YfhG**			C	X
BorD	M	X	B	X	**PgaB**		X	B	X	**YfiB**		X	C	X
BsmA			B	X	**PqiC**				X	**YfiL**			B	X
ChiQ		X	B	X	**RcsF**	H	X	B	X	**YfiM**				X
CsgG		X	AB	X	**RlpA**		X	AB	X	**YfjS**			B	X
CusC		X	B	X	**RseC**		X		X	**YgdI**	H	X	B	X
CyoA		X	AB	X	**RzoD**				X	**YgdR**		X	C	X
DcrB	H		C	X	**RzoR**				X	**YgeR**			B	X
DigH				X	**Slp**	H		AB	X	**YghG**		X	B	X
EcnA		X	AC	X	**SlyB**	H	X	AB	X	**YgiB**	H			
EcnB		X	AC	X	**Wza**		X	B	X	**YhdV**		X		X
EmtA			AB	X	**YaeF**			B	X	**YhfL**		X	B	X
FlgH		X	AB	X	**YafL**					**YiaD**	H	X	B	X
GfcB		X	B	X	**YafT**		X	B	X	**YiaF**	H			X
GfcE		X	B	X	**YafY**			B	X	**YidQ**			C	X
HslJ			B	X	**YaiW**		X	C	X	**YidX**		X		X
LoiP	H		B	X	**YajG**	H		B	X	**YifL**				X
LolB	H	X	ABC	X	**YajI**			B	X	**YiiG**		X	C	X
LpoA	H				**YbaY**	M	X	C	X	**YjbF**			B	X
LpoB	H		B	X	**YbfP**		X	B	X	**YjbH**				X
Lpp	H	X	ABC	X	**YbhC**		X	B		**YjeI**			B	X
LptE	H		ABC	X	**YbjP**	H	X	B	X	**YnbE**		X	B	X
MdtE		X	B	X	**YcaL**			B	X	**YnfC**			B	X
MdtP		X		X	**YceB**	H	X	B	X	**YoaF**		X	B	X
MdtQ				X	**YceK**		X		X	**YpdI**				X
MepS			B	X	**YcfL**		X		X	**YqhH**			B	X
MetQ	H	X	C	X	**YcjN**		X	B	X	**YraK**			B	X
MlaA	H	X	B	X	**YdbJ**				X	**YraP**	H		B	X
MliC		X		X	**YdcL**	H	X		X	**YtcA**				X
MltA	H	X	AB	X	**YdhY**			B						

Experimentally determined via click chemistry, H = high confidence, M = medium confidence [148]*, DOLOP prediction* [92]*, Experimentally determined lipoproteins by A = previously identified, B = labelled with radioactive palmitate, C = globomycin inhibition* [147]*, * Prosite prediciton*

Roles of bacterial lipoproteins

Lipoproteins have a wide range of functions that are not limited to envelope biogenesis as briefly discussed before. Figure 14 represents an overview of some lipoproteins and their functions in *E. coli*. Here, lipoproteins will be discussed according to their role except for Lpp which will be given greater attention. Lpp is the most abundant lipoprotein in *E. coli* and the determinant for the essentiality of lipoprotein modification pathway and is a key factor in the work described in Chapter III. A focus will be on lipoproteins predicted or shown in *E. coli*.

Lpp

Lpp (Braun's lipoprotein) was first discovered by Braun in 1973 and its structure became quickly known in great detail. Initially translated with a signal peptide directing it to the Sec translocation machinery in the inner membrane, Lpp is inserted into the membrane with its main part present in the periplasm. Once the immature lipoprotein has been modified by the LMP in the classical manner discussed previously, Lpp is an alpha-helix anchored into the lipid membrane by a triacylated cysteine residue. Under optimal growth conditions, Lpp is transported to the outer membrane via the Lol system. Lpp is generally present as a homotrimer [149, 150]. It is highly abundant with some reports estimating over 1 million copies per cell and making up around 10% of mRNA, which is proportionally very high compared to the abundance of the modification enzymes (approximately, Lgt = 1500, Lsp = 1200, and Lnt = 450 copies per cell) [151]. Of the Lpp present in the cell, roughly a third is covalently linked (bound-form) and the rest non-linked to PGN (unbound-form) [152]. It has been shown that these two occupy distinct compartments of the cell and that unbound Lpp can be exposed at the cell surface [153]. Lpp has a terminal lysine residue which is covalently attached to the meso-DAP on the PGN peptide stem by the enzymatic activity of L,D-transpeptidases [63].

Although Lpp is the most studied lipoprotein due to its presence in *E. coli*, Lpp is not widely conserved and is restricted to a subclade of γ-proteobacteria [154].

Deletion of Lpp has no noticeable effect on growth but cells are more sensitive to SDS. The role of Lpp is to stabilize the macrostructure of the cell by linking the OM with PGN and by controlling the width of the periplasmic space [62, 155]. Without Lpp, the membrane forms blebs, suggesting poor attachment of the OM to the cell wall [156]. It has also been shown to have a role in virulence with Δ*lpp* Salmonella having reduced host cell invasion and an *E. coli* mutant being more susceptible to serum killing [157, 158].

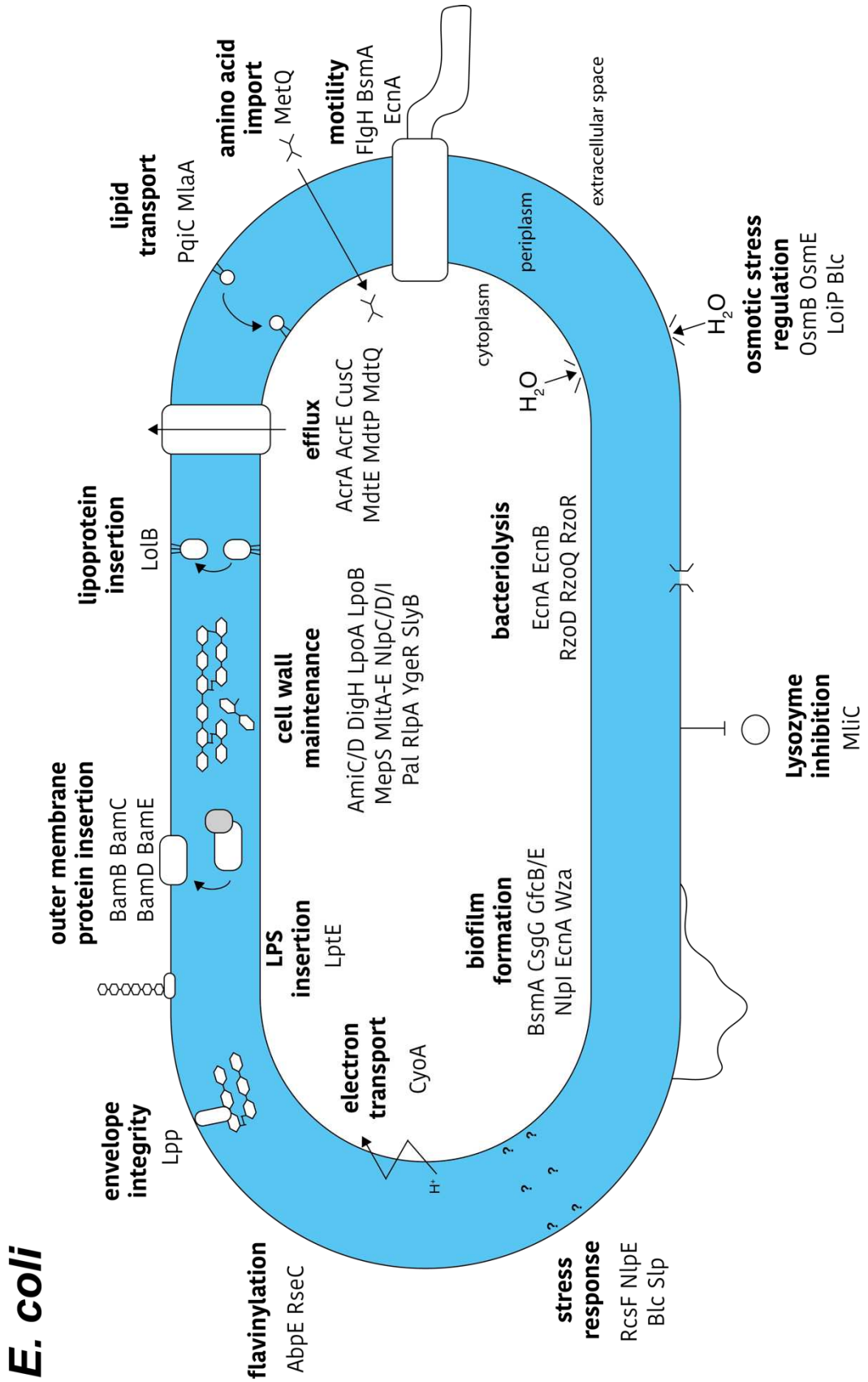


Figure 14. Lipoproteins of E. coli and their possible roles. *Displayed is a subset of confirmed and predicted lipoproteins that have been annotated with a possible function in the literature.*

Cell envelope biogenesis

Lipoproteins by nature reside in and around the periplasm of Gram-negative bacteria. As a key site for cell wall biogenesis, division and other cell envelope related functions, it is unsurprising that many lipoproteins have a role in cell envelope biogenesis.

The inseparable Bam, Lol and Lpt systems

Three of the essential systems in the cell envelope rely on lipoproteins, namely, the Bam system for the insertion of β-barrel proteins into the outer membrane; the Lpt systems responsible for transporting LPS to the OM and inserting it into the external leaflet; and the Lol system which transports lipoproteins to the OM.

The Bam complex is composed of a β-barrel protein (BamA), which is an integral membrane protein, and four associated lipoproteins (BamB-E) (Figure 4). BamA acts as the scaffold around which BamB-E function [159]. Of the five components, BamA and BamD are essential [159, 160]. BamC and BamE are not essential, but it has been postulated that they may play a role in the bacteria in a host environment [48]. The Bam complex is responsible for inserting outer membrane β-barrel proteins into the outer membrane.

The Lpt system which allows the hydrophobic Lipid A of LPS to be carried across the periplasm involves a single lipoprotein, LptE. LptE is associated to the OMP LptD and together play a role in extracting LPS from LptA (the periplasmic chaperone of the system) and inserting it into the membrane. As well as stabilizing LptD and possibly having a role in substrate recognition [161], LptE is thought to act as a plug and assists in the orientation of LptD towards the periplasmic components of the system [46]. As described, LptE is essential for cell viability [162].

The final essential lipoprotein is LolB. LolB is anchored into the outer membrane and is the receptor for the translocation of lipoproteins from the IM to the OM by the chaperone LolA [163].

It is interesting to note that these systems are co-dependent. The Bam system which inserts OMP into the membrane needs both an OMP (BamA) and lipoproteins (BamB-E). The lipoprotein translocation pathway requires lipoproteins already present in the OM (LolB) to be able to insert further lipoproteins there. The Lpt pathway requires the Bam pathway to insert the β-barrel LptD and the Lol pathway to insert the LptE into the OM. The crucial role of lipoproteins in these essential systems is clearly visible.

Cell wall synthesis, maintenance and division

Although there are no lipoproteins in the biosynthesis pathway of peptidoglycan, their role in cell wall biogenesis, maintenance and division are numerous.

Prominent lipoproteins in this pathway are LpoA and LpoB. Individually they are not essential but together they cannot be removed [28]. LpoA and LpoB form complexes with penicillin binding proteins (PBPs) PBP1a and PBP1b, respectively, for which their presence is essential for partner activity [28, 164]. Their association increases the transpeptidase activity and the glycosyltransferase activity of the PBPs clearing the way for new PGN to be inserted and as a result the synthesis of new PGN [165]. LpoA and LpoB have different functions and associate to different complexes. LpoB localizes to the septum of the cell and therefore is thought to have a role in cell division, whereas LpoA is considered to have a role in new PGN synthesis during cell elongation [164].

PGN is a dynamic macromolecule that is constantly being remodeled. A large part of this is the hydrolysis of various linkages which are performed by a number of different lipoproteins. As part of the LpoA/PBP1a complex, the lipoprotein NlpI acts as an adapter by binding to hydrolases and recruiting them to the site of PGN synthesis [166]. It has also been shown to play a role in cell division [167] and membrane integrity as its removal increases outer membrane vesicle (OMV) production [168]. One of the hydrolases regulated by NlpI is MepS whose activity appears specific for cleavage of PGN during cell elongation [169]. NlpC is another lipoprotein shown to have hydrolase activity [170].

There are a number of PGN transglycosylases which cleave the glycosyl bonds between the alternating NAM-NAG molecules of the PGN backbone. The biggest family of these lipoproteins are MltA-E which have shown transglycosylase activity [171-174].

A major OM lipoprotein Pal (peptidoglycan associated lipoprotein), forms part of the Tol-Pal system which has a key role in cell division. Pal is recruited to the septal ring of the dividing cell, where it binds non-covalently to PGN allowing constriction and invagination of the OM [175, 176]. A more recent proposal suggests that the Tol-Pal system also promotes PGN cleavage [177]. At the division site, other lipoproteins are present. For example, DigH is a recently discovered glycosyl hydrolase which, in the presence of the amidase AmiD, increase PGN degradation [177]. Lipoprotein amidases, AmiD and AmiC both have been shown to cleave PGN and are implicated at the septal ring as deletion of these amidases prevents successful cell division [178-181]. Although less well described, NlpD and YgeR are both implicated in cell wall division [178, 182]. The lytic transglycosylase, RlpA, has been shown to associate with FtsZ and is located at the septal ring highlighting a role in cell division [183-185]. Another lipoprotein with a possible role in cell division is SlyB. A SlyB homolog in *B. multivorans* cells

were elongated and produced filaments [186]. Interestingly, a SlyB homolog in *K. pneumoniae* was upregulated in a polymyxin resistance strain, an antibiotic with activity on the cell membranes [187].

Phospholipid transport

Although widely debated, the manner by which phospholipids are transported across the periplasm likely involves lipoproteins. The most notable description of a phospholipid transport system is the Mla system. MlaA is a lipoprotein that is located entirely within the OM where is interacts with OmpC to assist in maintaining lipid asymmetry in the OM [188] and may have a role in lipid trafficking [189]. The PqiABC system, also speculated to be involved in phospholipid transport, contains an OM lipoprotein, PqiC which interacts with a periplasm spanning domain [190]. The mechanisms and regulation of these systems are intriguing to study, however, a complete picture is not yet known.

Virulence and evasion

Biofilm formation

Biofilms produced by bacteria pose a major obstacle for therapeutics. Due to the thick extracellular matrix, it is difficult to remove them from surfaces but also hard to access with antibiotics. The most notable example of a lipoprotein involved in biofilm formation is CsgG. CsgG is an integral membrane lipoprotein [191, 192] which is part of the Csg complex. This is the machinery responsible for the extrusion of curli from the cell, forming a major component of biofilms. As well as stabilizing CsgA and CsgB, it has a role in controlling the quantity of curli fibers being extruded [193]. CsgG has been shown to be regulated by NlpI [194]. Another major exporter of polysaccharide is Wza, another OM integral lipoprotein [195, 196]. The Gfc system exports extracellular polysaccharide, producing a capsule. Lipoproteins GfcB and GfcE are parts of this system [197, 198].

Other lipoproteins implicated in biofilm formation are EcnA [199], BsmA [200], and MdtE [201].

Motility

FlgH is an important component in the flagella machinery which enables a major form of motility [202]. It forms part of the OM ring structure [203] and depletion strains of lipoprotein modification showed a motility defect in *Salmonella* [204]. BsmA has been implicated in motility [200], as well as EcnA [199].

Efflux

Alongside biofilms, the efflux of antibiotics from the cell is a major factor in intrinsic resistance but also has natural roles such as extruding harmful substances from the cell [205]. The RND-family (resistance-nodulation-cell division) efflux pumps in Gram-negative bacteria are numerous and show redundancy [206]. AcrA, AcrE, CusC, and MdtE, and possibly MdtP and MdtQ are all lipoproteins involved in these processes. AcrA and AcrE are the best described and are periplasmic adapter proteins (PAP) connecting the OM channel TolC to the IM 'pump' AcrB [207]. Co-localisation of PGN and AcrA has been described with a possible role of Lpp in this process [208]. It is thought that AcrA and AcrE also have a role in selecting substrates for access to the pump [209]. CusC is involved in metal ion export [210] but is not a PAP but instead an integral membrane lipoprotein [145].

Virulence

Effects on adherence, intracellular survival and virulence of lipoproteins have been explored. MetQ in *N. meningitidis* is a surface exposed lipoprotein [211, 212] which plays a role in both cell adhesion to cervical epithelial cells but also survival in macrophages [211, 213]. Slp has been shown to be involved in adherence to human cells [214]. TraT is a surface exposed lipoprotein that may be associated to PGN [215, 216] that is responsible for resistance to complement and macrophages [217-219]. NlpI,

which seems to have multiple functions or is involved in multiple pathways, it has been shown to play a role in binding to host cells as well as evading serum killing [220]. The lipid transport MlaA mutants were also shown to be attenuated in a silkworm model [221].

MliC is a lysozyme inhibitor allowing survival inside host macrophages [222, 223].

Stress and other survival systems

Nutrient uptake and metabolism

Methionine transport is an important pathway for protein synthesis. Of this pathway, which acquires methionine, the lipoprotein MetQ is necessary [212, 224, 225]. ApbE and RseC is involved in thiamine synthase [226] [227] that catalyses flavinylation [228]. Flavins have a redox activity and are involved in electron transport pathways [229].

Stress responses

There are two key lipoproteins involved in signaling membrane stress. The first is RcsF which is thought to interact with BamA which displays it on the cell surface. If there are perturbations in the Bam system, mislocalisation of RcsF to the IM allows it to interact with IgaA, an inner membrane signaling molecule [230, 231]. NlpE is the other key lipoprotein involved in stress responses. Normally localized to the OM [232], perturbations in the lipoprotein modification pathway causes its mislocalisation to the IM [37], activating the Cpx stress response pathway [233].

Some lipoproteins are differentially expressed under stress conditions. Slp is expressed during carbon starvation [234]. Blc, which may have a role in lipid binding, is expressed under starvation conditions and high osmolarity [235, 236]. LoiP is upregulated under low osmolarity conditions and is a metalloprotease, regulated by the Rcs system [237]. OsmB and OsmE, both induced by the Rcs system in response to envelope stress, are induced under osmotic pressure [238, 239].

Although a whistle-stop tour of lipoproteins, hopefully it is apparent that lipoproteins have a role in many different aspects of the Gram-negative cell and inhibition of the pathway that produces them would have wide ranging effects. They have a lipid moiety in common but are otherwise very diverse in structure and function. Inhibiting the pathway may provide a target that has multiple downstream physiological effects and therefore incomplete inhibition may be sufficient to assist the immune system in clearing an infection.

Project aims

In light of the need for new therapies to treat bacterial infections this study looks to use the target-based approach to antimicrobial discovery. The chosen target is Lgt, an enzyme in the LMP, due to its essential role in modifying the lipoproteins in essential pathways as well as its influence on multiple key pathways that reduce a pathogens susceptibility to chemotherapy or affect its virulence. We hypothesise that Lgt is a good target for novel antibiotics and will therefore challenge this hypothesis.

There is still some outstanding understanding about Lgt. Although some sequence alignments have been shown in the literature, they are often poorly described and the differences between pathogenic bacteria has not been assessed. There are some outstanding questions regarding the mechanism of action of Lgt and the exact role of H^{103} in catalysis is questionable. Fortunately, a structure of *E. coli* has been solved but comparison with other bacteria has not been completed. The advent of prediction tools such as AlphaFold2 enables this comparison with some confidence. Beyond this, the essentiality of Lgt, particularly the role of the major lipoprotein Lpp has conflicting accounts in the literature. As Lpp is key for resistance to inhibitors of other enzymes in the pathway, a better understanding of Lgt and Lpp is important. Finally, some *in vitro* activity assays have been described in the literature but they have short-comings and have not been used in high-throughput applications. A simple, adaptable assay is key to screening large libraries of small molecules for inhibitors of Lgt.

This study has three core axes. The first is to explore the conservation of Lgt. The study of sequence and structural diversity between the priority pathogens will help us better understand the broadness of the effects targeting Lgt might have. The genetic context of *lgt* is yet to be well discussed and we would like to understand if the synteny is conserved and if *lgt* is found in similar gene clusters. We will assess the phylogeny of Lgt and how this may correlate to structure and synteny. Beyond *in silico* analysis we will seek to determine the functional conservation of Lgt from a selection of priority pathogens via experimental complementation studies.

The second axis will explore the essential nature of Lgt in *E. coli*. Initially major lipoprotein Lpp was thought to be a key factor in Lgt essentiality, but recent studies suggest this is not the case. As Lpp is a factor in the essentiality of down-stream acting enzymes Lsp and Lnt we will seek to understand the role of Lpp on bacterial viability in the absence of Lgt. We will also assess the effects of Lgt depletion on bacterial growth and morphology.

Finally, to study Lgt in greater detail we will seek to develop a quantitative activity assay and attempt to set up an assay in high-throughput format to be used as a screen for Lgt inhibitors.

Section II: Conservation of Lgt

Summary

The aim of this chapter is to explore the conservation of Lgt.

A selection of key pathogenic species of bacteria were selected after reviewing global initiatives for tackling AMR. From this selection, protein sequence alignments and phylogenetic analysis were carried out to obtain insight into the evolutionary conservation of Lgt. This analysis revealed residues previously thought to be key for Lgt catalysis to be varied in non-proteobacterial species but also confirmed that Lgt is generally well conserved. Further to this, the local genetic context of *lgt* was analysed and revealed that *lgt* is likely to be found in an operon with *thyA* in proteobacteria or *hprK* in firmicutes.

Structural comparisons were conducted using the X-ray crystal structure of *E. coli* Lgt with AlphaFold2 predictions of Lgt enzymes from the pathogen list. All the predicted structures show variation to the solved structure in a loop region (L6-7) which suggests it may be a flexible region adding to the hypothesis that this region has a role in providing space for substrate entry and exit. The structures are broadly very similar with the notable difference of a large so-called head domain exposed to the periplasm in Enterobacterales and Mycobacterium but smaller head groups in other orders.

To see whether variations in sequence or structure had a material effect, a selection of *lgt* genes were chosen from a range of pathogenic species and transformed into two different *E. coli* depletion strains and analysed for growth and viability upon *lgt* expression. We found that Lgt from *A. baumannii*, *P. aeruginosa* and *H. pylori* complement Δlgt^{P} (plasmid encoded $Lgt^{E.c}$) and Δlgt^{C} (chromosomal encoded $Lgt^{E.c}$) in liquid media but only *H. pylori* and to a lesser extent *P. aeruginosa* could restore viability of Δlgt^{C} on solid media.

We hypothesised that the head domain may have a role in protein-protein interactions (PPI), particularly with Lsp and Lnt. ColabFold predictions revealed no interaction between these enzymes. The head domain of Lgt from *E. coli* also showed minimal predicted protein-protein interactions contrary to those from *M. tuberculosis*, *H. pylori* and *S. aureus* which had high predicted PPI. To assess whether changes in the head domains were able to prevent activity, the head groups from *M. tuberculosis*, *H. pylori* and *S. aureus* were cloned into the *E. coli lgt* gene. Only the head domain from *H. pylori* was sufficient to rescue growth and viability in the two depletion strains, suggesting a functional role of the head-domain that may be specific for particular groups of organisms.

These data highlight that Lgt has a well conserved catalytic site and inhibition may have broad spectrum capabilities but the highly variable head domain may provide an avenue for narrow spectrum inhibition.

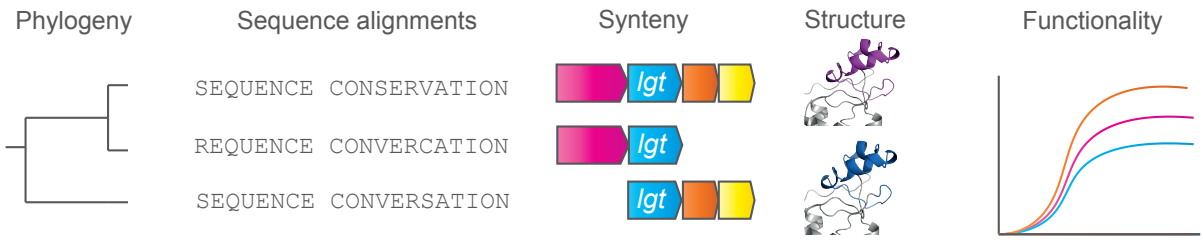


Figure 15. *Graphical abstract of Section II. Section II explores the conservation of Lgt at a sequence, structural and functional level.*

Results and discussion

Selection of Pathogens

In order to study the conservation of Lgt among species of interest we first compiled a list of pathogens that are difficult to treat due to their AMR properties. The composition of the list takes into account the WHO Priority Pathogens [9], ESKAPE pathogens[10], CDC Threat Report pathogens [11], and a remaining selection from the recent review into the impact of AMR [8] (Table 1 and 8).

The final list of 22 key AMR pathogens is composed of a range of organisms (Table 8, Figure 16) which infect a range of sites in the human body (Figure 16). The list contains a single Actinobacteria, *M. tuberculosis*, which is a major concern globally with increasingly high rates of multi-drug resistance (MDR) and extensive-drug resistance (XDR) and the causative agent of tuberculosis, an infectious disease affecting the lungs.

Seven firmicutes are on the list, including *C. difficile*, a healthcare acquired infection (HAI) causing severe diarrhoea which is particularly hard to treat due to its ability to form spores. *S. aureus*, commonly found in the nasal tract, was described as one of the first 'superbugs' due to the prevalence of methicillin resistant *S. aureus* (MRSA) and generally causes epidermal infections. Two *Enterococcus* species, *E. faecalis* and *E. faecium* are present and are considered serious threats. Found in the gastrointestinal (GI) tract as commensals, as opportunistic pathogens they can cause endocarditis, meningitis and urinary tract infections (UTIs). Three pathogenic *Streptococcus* species, *S. pneumoniae, S. agalactiae and S. pyogenes* present a wide range of symptoms and are found in varied environments. *S. pyogenes* (Group A Streptococcus) colonises the skin and upper respiratory tract (URT), *S. agalactiae* (Group B Streptococcus) colonises the gastro-intestinal (GI) tract and *S. pneumoniae* colonises the respiratory tract.

The remaining fourteen species are from the proteobacteria taxonomic group. More specifically, *C. jejuni* and *H. pylori* originate from the ε-proteobacteria are helical in shape and colonise the GI tract and stomach, respectively. A single β-proteobacterium, *N. gonorrhoeae* is a common cause of sexually transmitted infections (STIs) and genital tract infections. The remaining bacteria belong to the γ-proteobacteria, of which *E. coli*, *S. flexneri*, *C. freundii*, *S. enterica*, *M. morganii* cause GI infections. *A. baumannii* and *P. aeruginosa* causes bloodborne infections and respiratory infections. *S. marcascens* and *P. mirabilis* cause UTIs. *H. influenzae* and *K. pneumoniae* also infect the respiratory tract.

To reduce bias in the selection of specific strains from each species of interest, the first strain from each species listed in the SyntTax database was selected, unless a known reference strain was present. The SyntTax database provides a database of mostly annotated strains for the RefSeq and non-RefSeq GenBank databases and automatically provides local synteny [240]. These strains were validated for annotation of *lgt* and if it was not annotated, the next strain in the list was selected. To assess whether these strains were representative of both the sequence and genetic context, nine other strains of the same species were selected from the database at random and were compared. If the sequence or synteny had notable differences a wider assessment of up to 50 more genomes was compared and from this a suitably representative strain was selected. Specific strain names are found in materials and methods (Table 8). In each instance, only a single copy of *lgt* was found in each genome.

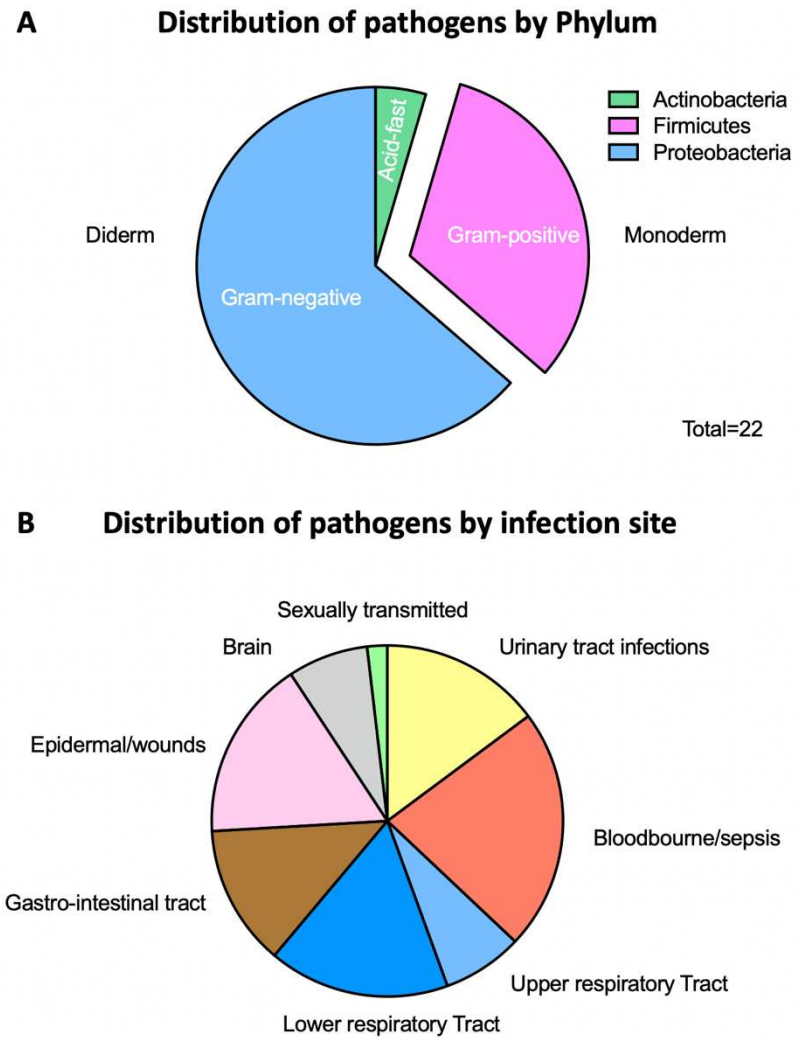


***Figure 16. Distribution of strains**. by A) phylum and B) common site of diseases caused by selected pathogens, some species infect multiple sites.*

Phylogeny

Multiple sequence alignments (MSAs) of the Lgt amino acid sequence of selected strains was conducted in Geneious Prime using the ClustalO method. After MSA, a phylogenetic tree was produced using Jukes-Cantor and UPGMA methodology in Geneious Prime with bootstrap analysis. To assess the degree to which the phylogeny of the Lgt protein differed from the global phylogeny of these strains, 16s rRNA analysis was conducted. This approach takes the 16s rRNA gene as a conserved maker gene that contains a hypervariable region to generate phylogeny[241].

The phylogenetic tree is a representation of the evolution of a particular gene, protein or species. It is made up of nodes and branches representing divergence events and the persistence of the element through time, respectively, with the length of branches also representing the proportional number of mutations [242]. Therefore, from analysing the phylogeny of Lgt in comparison to the commonly used 16s rRNA, we can see if the two elements have followed a similar evolutionary path or if there are differences or events of interest.

Observations of the percentage of sequence identity, i.e. residues that are identical between any two strains, reveals extremely high identity within the family *Enterobacterales* and high identity within the wider γ- and β-proteobacteria. The identity within ε-proteobacteria is high and within firmicutes is high but the percentage of identical residues between each of these two classifications of bacteria is much lower. Distinct groups can be clearly seen on the heatmaps (Figure 17).

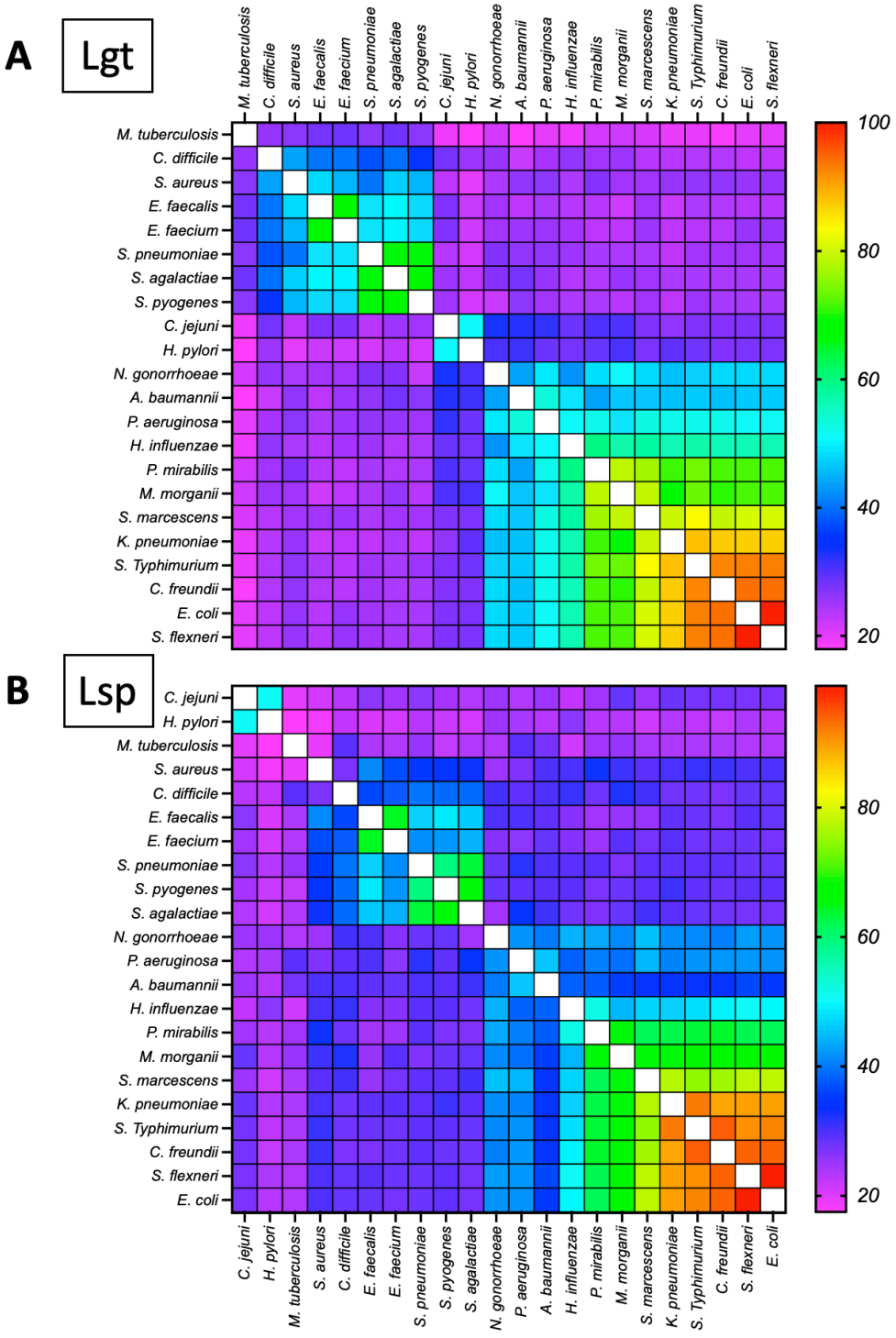


Figure 17. Percentage similarity of amino-acid sequences between selected pathogenic bacteria. *A) Lgt, B) Lsp*.

The Lgt phylogenetic tree has two distinct branches (Figure 18), the first containing the phyla actinobacteria and firmicutes and the second containing proteobacteria. The firmicutes subdivide into two branches containing the orders lactobacillales or bacillales and clostridales. The proteobacteria likewise subdivide into clades in alignment with their family taxa of ε-, β-, or γ-proteobacteria. Bootstrap analysis provides a measure of confidence by performing repeats of the analysis on samples of the alignment. In this instance the alignment was sampled 100 times and the figure shown by each node (Figure 18) represents the number of repeats where the same distribution was found. The closer to 100, the higher the confidence. From this metric we can see that the phylogeny of the Lgt protein sequence has high confidence throughout.

As with the Lgt protein sequence, the 16s rRNA phylogenetic tree separates the bacteria by their known taxa and is much the same as the Lgt tree. Some variation is observed within orders. *S. aureus* (Bacilalles) is clustered with *C. difficile* for Lgt but moves to be integrated into a clade of Lactobacilliales for 16s rRNA. There is also some variation in the Proteobacteria, particularly in the family of Enterobacterales. However, these are closely related species as observed by the short branch lengths.

The broad division between the Gram-negative proteobacteria and the Gram-positive firmicutes was previously observed by Banerjee *et al.,* (2013) who see the same division. Although they do not show the data, they state that they observe *M. smegmatis* between the two main groups and suggest that it represents an evolutionary intermediate of Lgt, particular for the change from H^{103} to Y^{103}. Although not the same species, *M. tuberculosis*, from the same genus as *M. smegmatis* is not observed here between the Gram-positives and Gram-negatives but more closely related to the monoderms. Although the focus of this study is on pathogenic bacteria, particularly those with high rates of resistance, a wider analysis of Lgt phylogeny may reveal a more comprehensive overview of Lgt evolutionary history.

As the lipoprotein modification pathway is yet to be reported absent in any bacterial species, we looked to see if Lsp shared the same phylogeny as Lgt, hypothesising that they co-evolved. After conducting the analysis for the protein sequence of Lsp in the manner described above we observed a major difference. Whereas for the most part, the tree takes the same form for Lsp as it does for Lgt, the ε-proteobacteria (namely, *C. jejuni* and *H. pylori*) now form their own branch, separate from Gram-positive, Gram-negative and Actinobacteria. This indicates a major evolutionary divergence. Interestingly, the similarity scores are not drastically lower when comparing all the bacteria against each other with many of the key residues of Lsp remaining intact (Appendix II). This suggest the Lsp of the ε-proteobacteria would still be functional and likely performs the same role but it may have been replaced by an *lsp* gene from a distant species at some point in time. Although interesting and would

provide a fascinating study, it is not within the scope of this thesis to explore the evolution of Lsp further. It can be concluded, however, that Lsp and Lgt share a similar phylogeny.

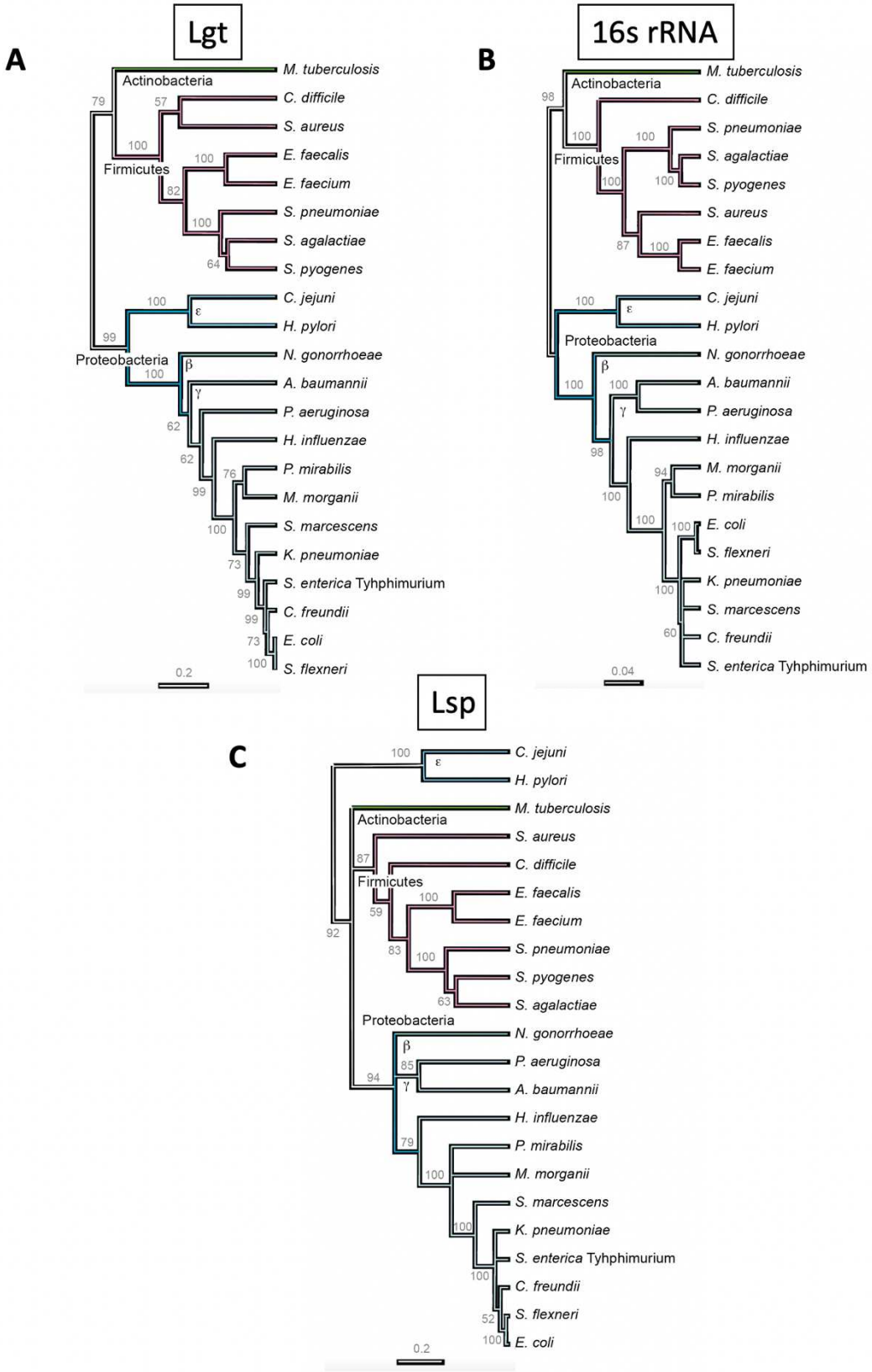


Figure 18. Phylogenetic trees *of A) Lgt protein sequence, B) 16s rRNA sequence, C) Lsp protein sequence. Numbers indicate confidence determined by bootstrap analysis.*

Sequence Conservation

Sequence alignments revealed 16 residues in Lgt conserved across all the selected pathogens (Y_{26}, R_{73}, G_{99}, G_{104}, D_{129}, G_{142}, R_{143}, G_{145}, N_{146}, E_{151}, G_{154}, F_{211}, Y_{235}, R_{239}, E_{243}, R_{246}) (numbering according to Lgt sequence of *E. coli*) (Figure 19, 20). W_{25}, F_{147}, and Q_{264} were conserved throughout proteobacteria and firmicutes. Although not fully conserved, A_{132}, N_{149}, L_{200}, and Y_{201} demonstrated >90% conservation. Pailler *et al.* (2012) completed alignments with proteobacteria, firmicutes and actinobacteria and is the best described conservation study in the literature (Figure 6). All of the fully conserved residues previously reported are found fully conserved in our analysis. Possibly due to the smaller scale of the alignment conducted in this study, there is a greater number of fully conserved residues than previously described. The majority of these residues are noted by Pailler *et al.* (2012) as being fully conserved in combinations of different phyla, i.e. in firmicutes and actinobacteria but not proteobacteria. This suggest that although they may not be fully conserved across hundreds of bacteria aligned previously, they are still very well conserved. As our study is focused on pathogenic bacteria, it is important to note that this subset of strains share a boarder range of conserved residues. The breakdown of strains in the previous studies is not presented so a direct comparison cannot be conducted.

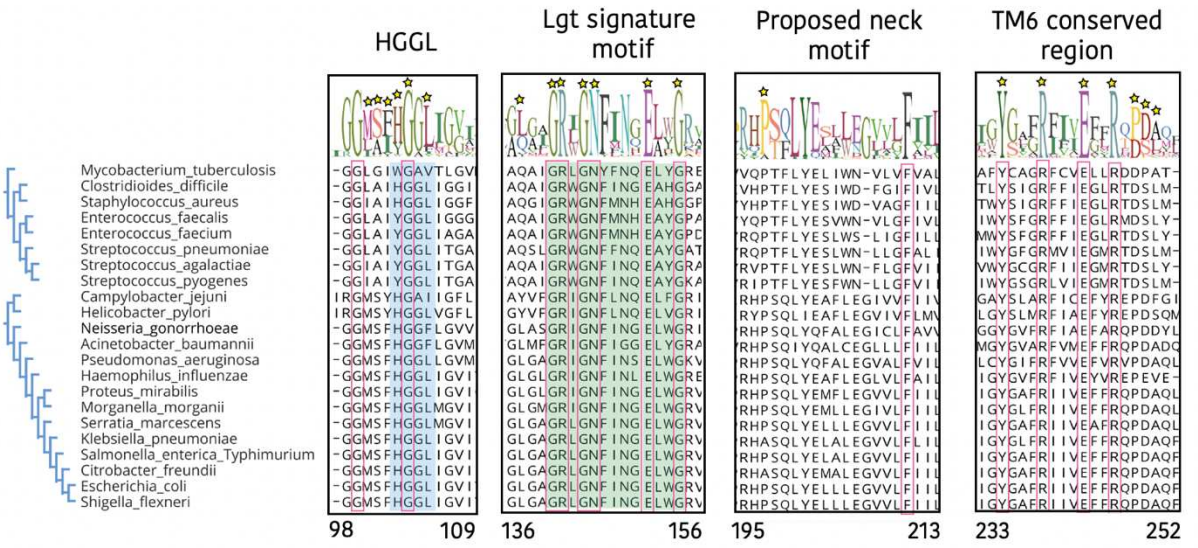


Figure 19. Key sections of Lgt sequence alignments. *HGGL motif highlighted in blue, Lgt signature motif highlighted in green, fully conserved residues in pink boxes, stars indicated experimentally determined essential residues in E. coli, numbers refer to amino acid position from the E. coli sequence (neck motif displayed on structure Figure 21).*

The H_{103}GGL motif, located in the periplasmic half of TM3 (Figure 21), is broadly conserved but with some notable divergences. H_{103} is a residue thought to be key in the catalytic activity of Lgt [77, 116]

and is conserved in proteobacteria and some firmicutes. However, we observe a tyrosine (Y) residue is present in *Lactobacillales* and tryptophan (W) is present in *M. tuberculosis*. This variation echos what has been seen previously at this site with different residues present at this position and an incomplete division between Gram-positive and Gram-negative bacteria (Table 3). Histidine and tyrosine are of similar size but it has previously been demonstrated that tyrosine could not be substituted for histidine in *E. coli* at this position [81] raising questions about the role of H_{103} in Lgt activity. Although H_{103}N, Q, A, R, Y substitutions have shown not to be viable, alongside ΔH_{103} [81, 82], one study report functional H_{103}Q [80]. It has previously been noted that *Mycoplasma gentilium* contains a glutamine (Q) as opposed to histidine in this position [82]. Although these studies used similar models to study Lgt (Δlgt^{P}) (Figure 7), they contain different complementing plasmids but the same induction systems, the difference in H_{103}Q viability may be due to this difference. The H to Y difference between firmicutes and proteobacteria has been previously reported but not further explored, although, the authors hypothesis that *M. smegmatis* (tyrosine) is a point of evolutionary divergence between Gram-positive and Gram-negative bacteria [78]. However, closely related *M. tuberculosis* has a tryptophan, suggesting that actinobacteria is not where the divergence occurred. G_{104} was fully conserved and G_{105} mostly so, with the exception of *M. tuberculosis* and *C. jejuni*. Finally, L_{106} was likewise highly conserved with few exceptions.

It may also be possible to extend this motif further. G_{98} is conserved in all but ε-proteobacteria and G_{99} is fully conserved, although they are not essential in *E. coli*. Following these residues and prior to H_{103}GGL, M_{100}SF is conserved in β- and γ- proteobacteria, M_{100}SY in ε-proteobacteria and $(L/I)_{100}$AI in firmicutes. The location of the additional residues extends the motif into the transition between TM3 and arm-2 (Figure 21). This region has been proposed to be a binding site of PG as 'site-1' [81] and therefore may provide phospholipid specificity in firmicutes compared to proteobacteria.

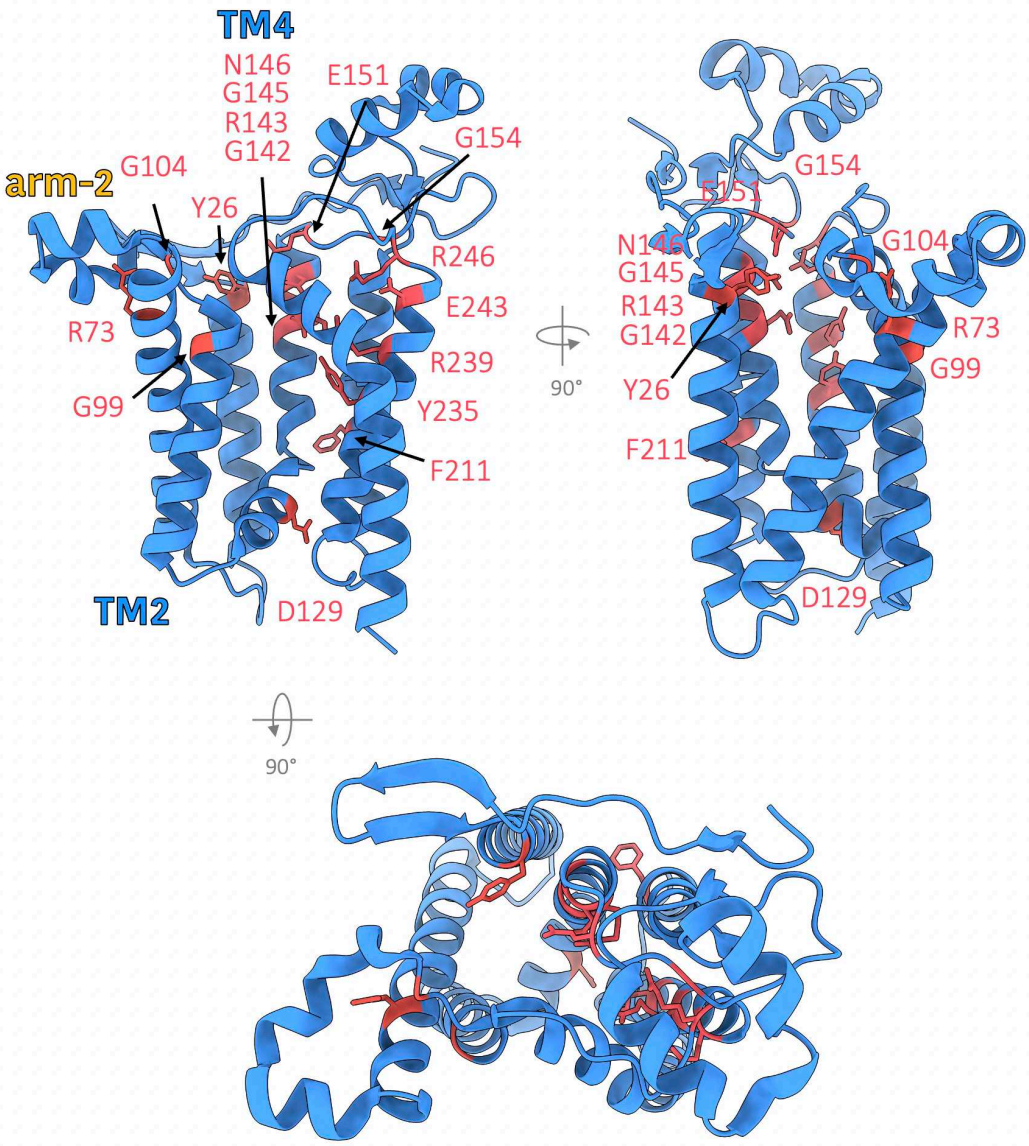


Figure 20. Conserved residues. *Red indicates conserved residues with side-chains shown mapped onto the X-ray crystal structure of Lgt* [81]

The Lgt signature motif, located at the periplasmic interface on TM4 (Figure 8, 21) and continuing into the transition to the head-domain, is widely conserved (Figure 19, 20). G_{142}, R_{143}, G_{145}, N_{146}, E_{151}, and G_{154} are fully conserved and have previously been shown to be essential in *E. coli* (Table 5) [80, 81], although the results were not entirely clear for R_{143} and E_{151} [80]. Other residues from the motif have not been analysed further with the exception of W_{153} which was functional with an alanine substitution (Table 5) [81]. F_{147} is conserved in all but *M. tuberculosis*, and F_{149} in all but *A. baumannii*. The remaining residues showed variation often with distinctions between different classes of the bacteria. The conserved residues are generally found on the inward facing turn of the helix (Figure 20) suggesting an important function in the central cavity.

Interestingly, Y_{80}, located at the arm-2 end of TM2 (Figure 8, 20) and adjacent to H_{103}, deemed essential in *E. coli* [81] but is only present in proteobacteria. G_{98} is fully conserved but is not essential when substituted for alanine or proline in *E. coli* (Table 5) [80, 81]. Beyond this, there are several residues that have been shown to be essential in *E. coli* [81] that are not widely conserved. For example, L_{139}, P_{248}, D_{249}, G_{264}, P_{269} are widely conserved in proteobacteria but are contrary to residues found in firmicutes. A_{250} and M_{263} were also deemed essential but are not widely conserved. Albeit the data presented for viability are simply spot dilutions without demonstration of gene expression or protein production and therefore effects on protein stability or membrane insertion are not explored.

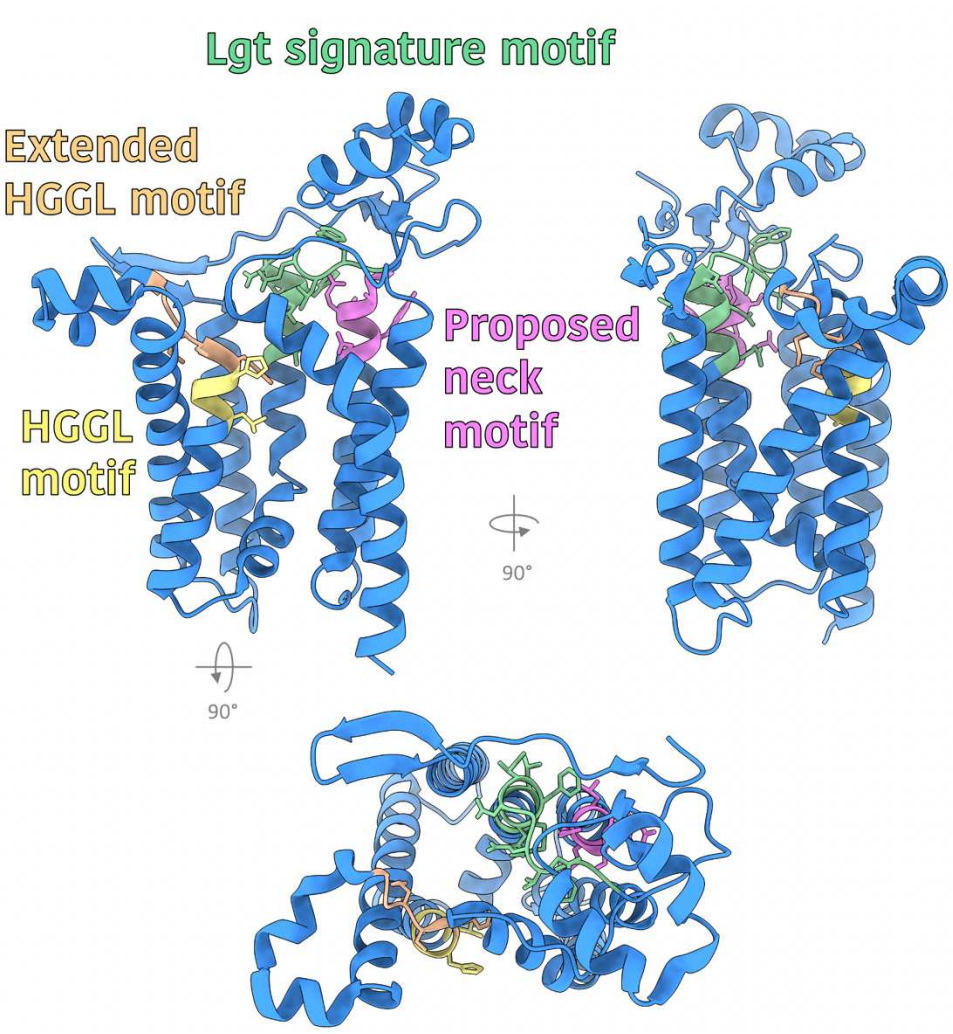


Figure 21. Lgt motifs. *X-ray crystal structure of Lgt [81] with key motifs highlighted.*

Another possible key region is found between P_{197} and E_{202}, located at the periplasmic side of TM5 adjacent to the Lgt signature motif and below the large head-domain (Figure 19, 21), where P_{197}SQLY (E/Q) is almost entirely conserved in proteobacteria and PTFLYES is fully conserved in firmicutes and *M. tuberculosis*. P_{197} has been shown essential in *E. coli* Lgt [81]. The location of this motif is found in

the central cavity and below the head domain (proposed ‘neck’ motif) and therefore may have a role in enzymatic activity. Exchanging the proteobacteria sequence for that of firmicutes may provide an interesting insight into these regions function and possible role in species specificity.

One notable region that had low sequence conservation was the so-called head domain.

Our analysis of the sequence conservation of Lgt concurs generally with that which has been reported previously but includes more details due to the larger number of genomes analysed. Further analysis of H_{103} in the HGGL motif and the proposed neck motif may provide greater understanding of species specificity.

Structural conservation

In order to assess structural differences between different Lgt proteins, structure prediction program AlphaFold2 was employed [243]. A few controls were conducted prior to full scale analysis. All PDB structures prior to 2018 were used to train the AlphaFold2 software and as the structure of Lgt had been solved [81] and deposited onto the PDB database prior to 2018, Lgt was part of this training. There are two basic conditions in which AlphaFold2 can be used, the first is without a structure template, the second is using the PDB70 database to search for templates. We ran the *E. coli* Lgt sequence through both conditions, compared the structures and found no noticeable difference in output. AlphaFold2 relies on an initial step of MSA and if there are sufficient sequences in the MSA, they state a template does not necessarily assist the model.

AlphaFold2 provides two confidence metrics, plDDT and predicted aligned error (PAE). plDDT measures the percentage of correctly predicted interatomic distances and PAE the expected positional error compared to a real structure. When analysing *E. coli* Lgt confidence metrics, there is generally high confidence (>90%) as stated by plDDT and low predicted error as defined by PAE (Figure 22). However, the region of the protein at the periplasmic half of TM7 and the loop from TM6 (L6-7), low plDDT and high PAE scores were observed which may indicate a region of flexibilty. The structure from AlphaFold2 was then visually compared to the solved structures (form-1, 5AZB, and form-2, 5AZC, Figure 22) and a clear difference between the structures is observed in this region (Figure 22). Form-1 (5AZB) represents a structure of Lgt bound to palmitate and detergent, and form-2 5AZC contains two phosphatidylglycerol substrate molecules in the central cavity (Appendix III). The solved X-ray crystal structure of Lgt has an extended α-helix which arcs just below the membrane plain whereas the predicted AlphaFold2 model has a shortened α-helix and loop extending upwards, through the membrane.

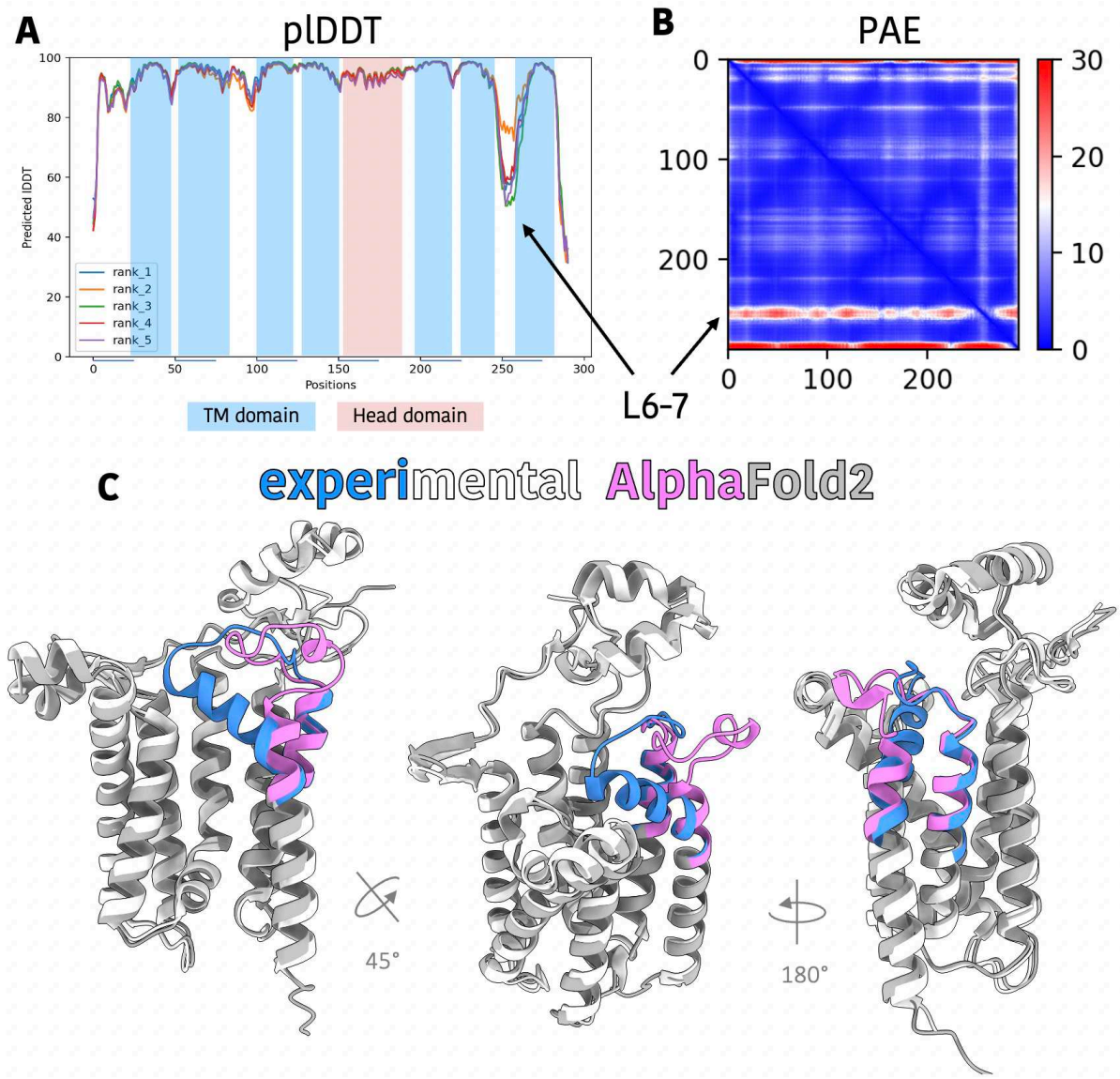


Figure 22. Comparisons of X-ray crystal structure of E. coli Lgt and AlphaFold2 predicted structure. *Confidence metrics provided by AlphaFold2 A) pLDDTs and B) predicted aligned error (PAE), ranks 1-5 denote the 5 predicted structures generated by AlphaFold2. C) white and blue depicts X-ray crystal structure, grey and pink depicts AlphaFold2 structures, coloured region = L6-7.*

The flexibility of L6-7 has been shown by minor changes between the two solved structures (Appendix III) [81] and in modelled structures with substrates (Figure 9) [79]. It should be noted that the rest of the predicted structure closely maps to the solved structure. For further analysis, due care was taken when analysing the L6-7 region of the protein, otherwise it is a suitable method to compare changes in the protein structures when a solved structure is not available.

For the remaining pathogens, the Lgt sequences were queried in AlphaFold2. *M. tuberculosis* presents an extended Lgt sequence. Beyond the noticeable Lgt structure, extended loops (Appendix IV) and

helices with exceptionally low confidence can be observed. Therefore, a shortened sequence, the length of *S. aureus* Lgt was used for analysis. The extended C-terminal part of Lgt from *M. tuberculosis* has no sequence similarity to other proteins. Consistent with that of *E. coli*, poor confidence was observed in L6-7 region of all other structures, with the exception of *C. difficile, S. aureus and P. aeruginosa*. All confidence metrics can be found in Appendix IV. In some instances, only medium confidence could be observed in the region between L2-3 (arm-2) (Figure 8), for example, in *E. faecalis*, suggesting the possibility of flexibility in this region. Arm-2 is close to site-1, a proposed PG binding site.

The most striking difference between structures is found in the head domain (Figure 23). β-proteobacteria and *Enterobacterales* all have a similarly large head domain whereas more distant proteobacteria and firmicutes have a smaller, less prominent head domain (Figure 23, Appendix V). *M. tuberculosis* has a large head domain with a different shape to that of the others. The large head-domains protrude into the periplasm whereas the smaller head domains are located closer to the membrane interface. The head domain is not discussed in the literature.

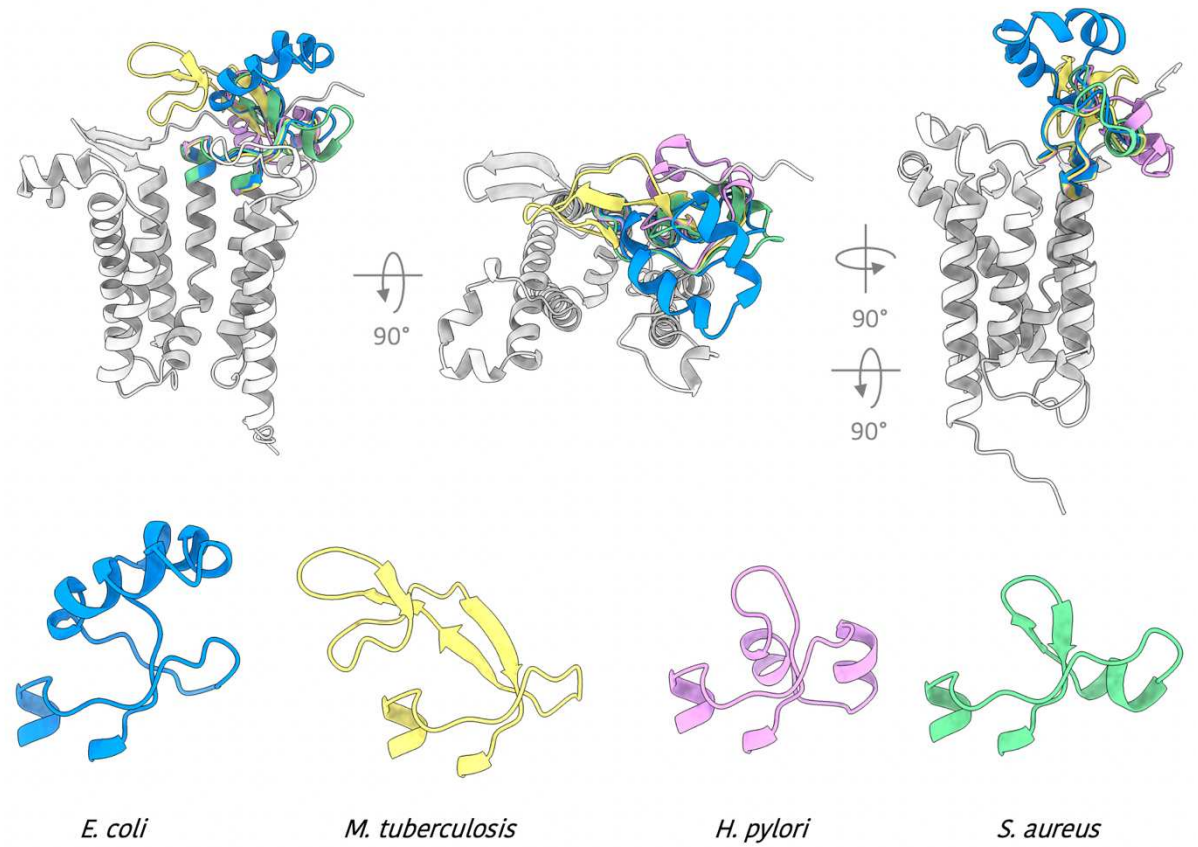


Figure 23. Head domains of Lgt. *Highlighted are the head domains from E. coli. M. tuberculosis, H. pylori and S. aureus attached to the body of E. coli Lgt sturcure predicted by AlphaFold2.*

Due to the proposed low abundance of the lipoprotein modification enzymes and the high modification rate of lipoprotein Lpp in *E. coli*, we hypothesised that the head-domain could have a role in co-localising the three enzymes in the pathway to each other. To explore this hypothesis we employed ColabFold [244], a derivative of AlphaFold2 which allows multiple sequence entries to predict protein-protein interactions (PPIs). The ColabFold predictions had little to no confidence that Lgt interacted with Lsp or Lnt (Appendix IV). As a control for these predictions, we assessed the PPI of Lgt and Lpp, a known substrate of Lgt. In this prediction, the signal peptide of Lpp is predicted to bind in the front-cleft of Lgt (Appendix IV). The location of the signal peptide is discussed in the models by Singh *et al.* (2019) and predict it bound to the front cleft. A recent review performed the same comparisons as shown in this study and their results concur [245] that Lpp is likely to interact at the front cleft. However, their predicted structures show Lpp in a slightly different orientation. Needless to say, both our prediction and that previously reported show Lpp highly likely to interact with the front cleft (Appendix IV).

We employed a second prediction tool (ScanNet) to assess whether the head-domain was likely involved in protein-protein binding [246]. This tool uses deep learning to predict the probability of protein binding to each residue in the queried structure. We queried Lgt from the UniProt database which contains the AlphaFold2 predicted structures. The two arm loops (arm-1 and arm-2) have a high protein binding probability, as does L6-7 (Figure 24). This aligns with the current understanding that these two regions, which make up the front and side clefts, are regions where substrates may enter and exit. However, the role of L6-7 in protein binding has not been noted.

The head-domain of *E. coli* Lgt showed a low probability for protein binding. As there are noticeable differences in the structure of the head domains, we assessed the predicted protein binding of an actinobacteria *M. tuberculosis*, a firmicute, *S. aureus* and distant proteobacteria, *H. pylori*. All the structures show high predicted protein binding in the arm domains and L6-7. Interestingly, we observed higher predicted protein binding in the head domains of *S. aureus*, *M. tuberculosis* and *H. pylori* particularly in the region that may be the interface with membrane (Appendix VI). This raises the possibility that the head domain in other species may have a role in PPI. It is also noted that in all instances, the periplasmic N-terminus has a high predicted PPI.

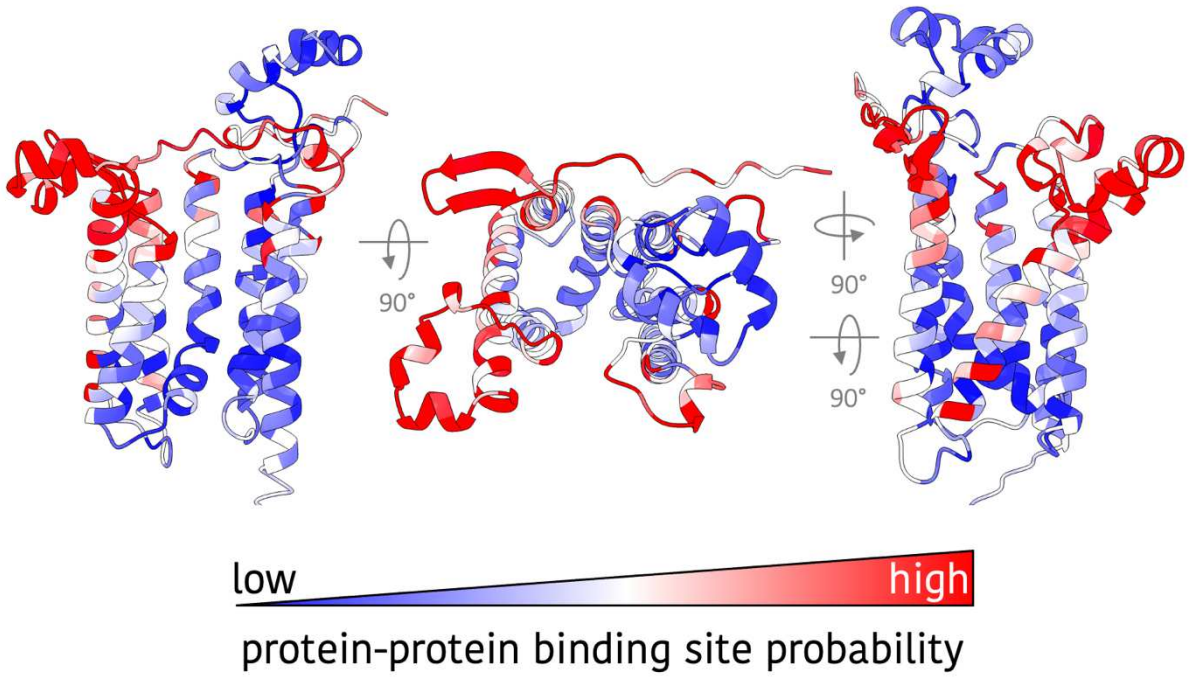


Figure 24. Predicted protein-protein bibding of E. coli Lgt. *Predicted protein-protein binding probability as determined by ScanNet [246], E. coli Lgt structure as determined by AlphaFold2 prediction*

Genetic context

To analyse differences in the genetic context of *lgt*, the SyntTax database was employed alongside two web-based operon prediction tools. SyntTax provides a visual representation of local genes [240], roughly 10 genes up- and down-stream of *lgt*. BioCyc [247] and MicrobesOnline [248] use different methods to predict co-translated regions. For example, MicrobesOnline considers the distance between genes, local conservation, expression correlation and function and BioCyc uses the intergene distances and functional classification. Initially the two operon prediction tools were compared with a great amount of concurrence between the two albeit with some exceptions (Figure 25).

The wider genetic region of *Enterobacterales* is well conserved apart from *M. morganii* and *S. marcescens* where variation is seen upstream of the *lgt* gene. In each instance, except for *M. morganii*, *lgt* is predicted to be coupled to *thyA*, a gene encoding thymidylate synthase, an essential enzyme in nucleotide synthesis. In all γ-proteobacteria, *thyA* is present and is transcriptionally coupled to *lgt* in *Enterobacterales* and *H. influenzae*. As Lgt has been speculated to reduce *thyA* translation [76] it may have become coupled to control regulation. *H. influenzae* has an extended operon of five genes that includes *nudH* and *tadA*, both involved in RNA metabolism, alongside a gene of unknown function. This unknown gene is also coupled to *lgt* in *P. aeruginosa* also where *thyA*, although present, is thought not to be co-translationally regulated. *M. morganii*, *A. baumannii*, and *C. jejuni* have *lgt* as a single translated unit. Interestingly, *thyA* is present next to *lgt* but reversed in *Morganella*, and in *N. gonorrhoeae lgt* is separated from *thyA* by two genes, a possible homolog of *leuA*, involved in leucine metabolism and a gene of unknown function, *yecA*. In ε-proteobacteria, *C. jejuni* has *lgt* housed alone whereas *H. pylori* is predicted to house *lgt* in a large unit of 7 genes: *rdxA* (oxidoreductase involved in resistance to nitroaromatic compounds), *rluA* (RNA processing), *waaA* (LPS biosynthesis), *folE2* (folate biosynthesis), *glyQ* (protein biosynthesis), and *gpsA* (lipid biosynthesis).

In firmicutes, *lgt* is invariably predicted to be coupled to *hprK*, a kinase involved in carbohydrate metabolism with the exception of *C. difficile*. In *Streptococcus*, two unknown genes are linked and *S. aureus* has a downstream O-acetyltransferase alongside a gene of unknown function as previously described [91]. MicobesOnline extends the translated region of *Enterococcus* to include *gspA* (glycosyl transferase involved in stress responses) and *galU* (nucelotidyl transferase possibly involved in stationary phase survival) concurring with what has been reported [90].

C. difficile has *lgt* linked to *rsmH,* a methyltransferase involved in RNA metabolism.

Interestingly, *M. tuberculosis* has *lgt* in a multigene operon encoding genes from the tryptophan synthesis pathway.

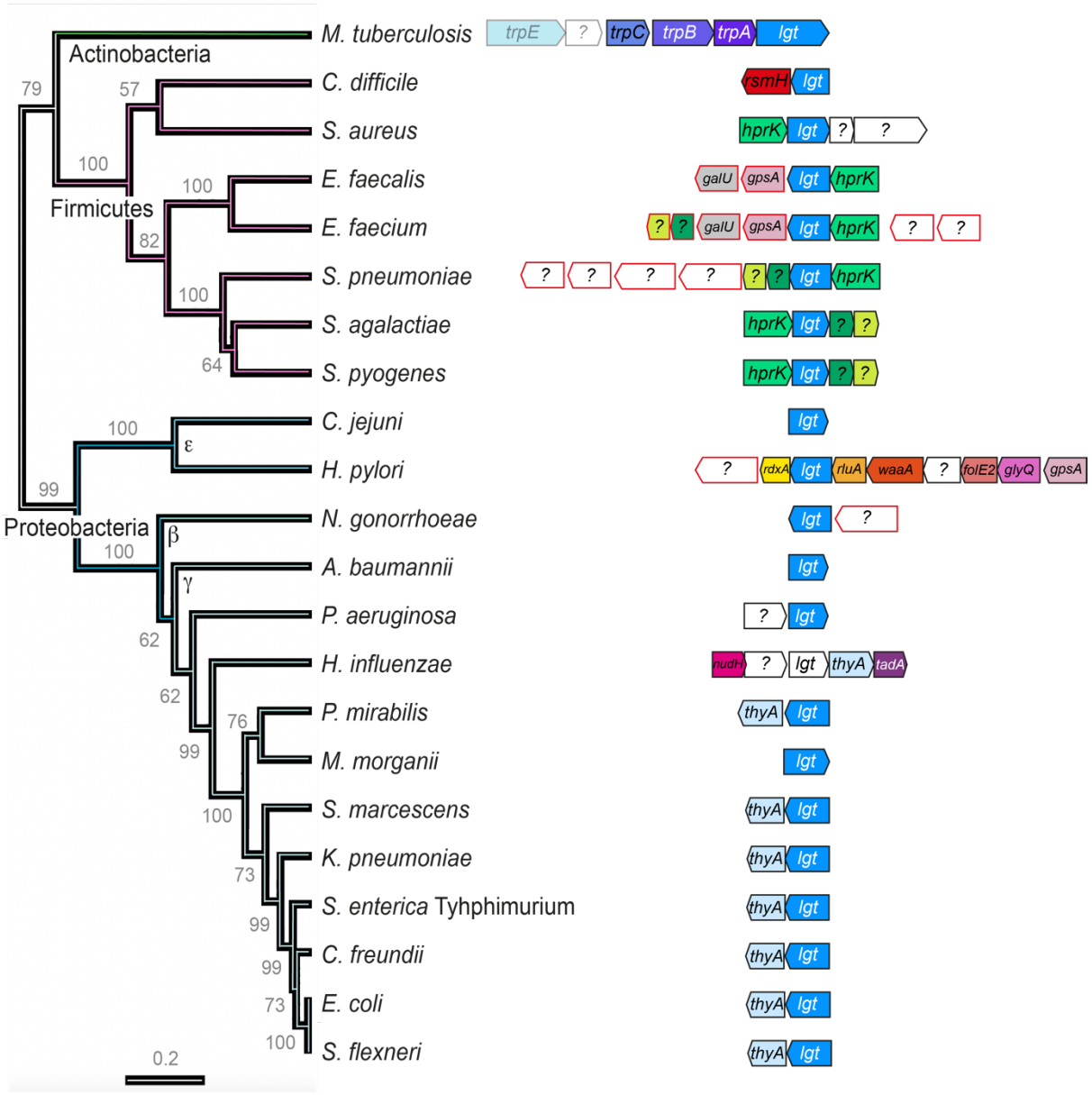


Figure 25. Predicted operons containing lgt. *Genes highlighted in black are found as consensus between BioCyc and MicrobesOnline operon predictions. Genes highlighted in red are predicted only by MicrobesOnline and grey are predicted only by BioCyc.*

Similarities in genetic context are seen within firmicutes and proteobacteria. However, Lgt does not appear to be near genes relating to lipoprotein modification or lipoproteins themselves. It is interesting that although Lgt is required for Lsp activity they are not conserved in an operon and to date no information about their regulation is known.

Functional conservation

As Lgt appears to be well conserved at a sequence and structural level but with some key differences, we sought to determine whether Lgt is conserved at a functional level. To this end, we employed two complementation strains (Figure 7). Both strains of *E. coli* have *lgt* replaced by a kanamycin resistance cassette and *lgt* restored under the control of an arabinose inducible promoter (P_{ara}). There are two notable differences, the first difference is that Δlgt^{P} has *lgt* present on a multi-copy vector [80] and Δlgt^{C} has *lgt* inserted onto the chromosome at the λ-attachment site, thereby housing a single copy of *lgt* [75]. The second is that $\Delta lgt^{P}$is in BW25113, a strain of *E. coli* modified to no longer have the ability to metabolise arabinose. Δlgt^{C} is in MG1655, a strain that retains this ability.

A plasmid containing an *lgt* gene for complementation studies was adapted from a previous study (pAM238-lgt^{Ec}) [80] to contain a *flag*-tag as opposed to a c-*myc*-tag. Lgt genes from a selection of key AMR pathogens were cloned into this vector to study their function in *E. coli* (Figure 7).

The *lgt* gene from *P. aureuginosa, A. baumannii* and *H. pylori* were successfully cloned into the pAM238 vector. Although possible successful clones of *S. enterica* Typhimurium, *E. faecalis*, *S. agalactiae*, *N. meningitidis* and *S. aureus* in pAM238 could be transformed into wild type *E. coli* and grew on solid LB media, they failed to grow in liquid LB media. These possible clones have not pursued further so it cannot be said as the reason for this.

Protein production in both Δlgt^{P} and Δlgt^{C} containing the pAM238 plasmid with *E. coli, P. aureuginosa, A. baumannii* and *H. pylori* was assessed under permissive conditions i.e. in the presence of L-arabinose to express the wild type Lgt and with or without IPTG expressing the complementing *lgt-flag*. An empty pAM238 vector was used as a negative control. Figures 26A and 27A present α-flag Western blots of each of the strains. Lgt-flag is produced in each strain and no Lgt-flag is observed in the pAM238 only control. $Lgt^{E.c}$ shows a greater level of production in comparison to the other proteins. Lower levels of protein production are also observed in the absence of IPTG suggesting basal level of expression from the P_{lac} promoter.

This basal level of expression acts as a 'low' expression system in comparison to the 'high' expression in the presence of IPTG. Appendix VII shows the entire Western-blots and multiple bands of approxiately 15 kDa are observed, possibly relating to protein degredation.

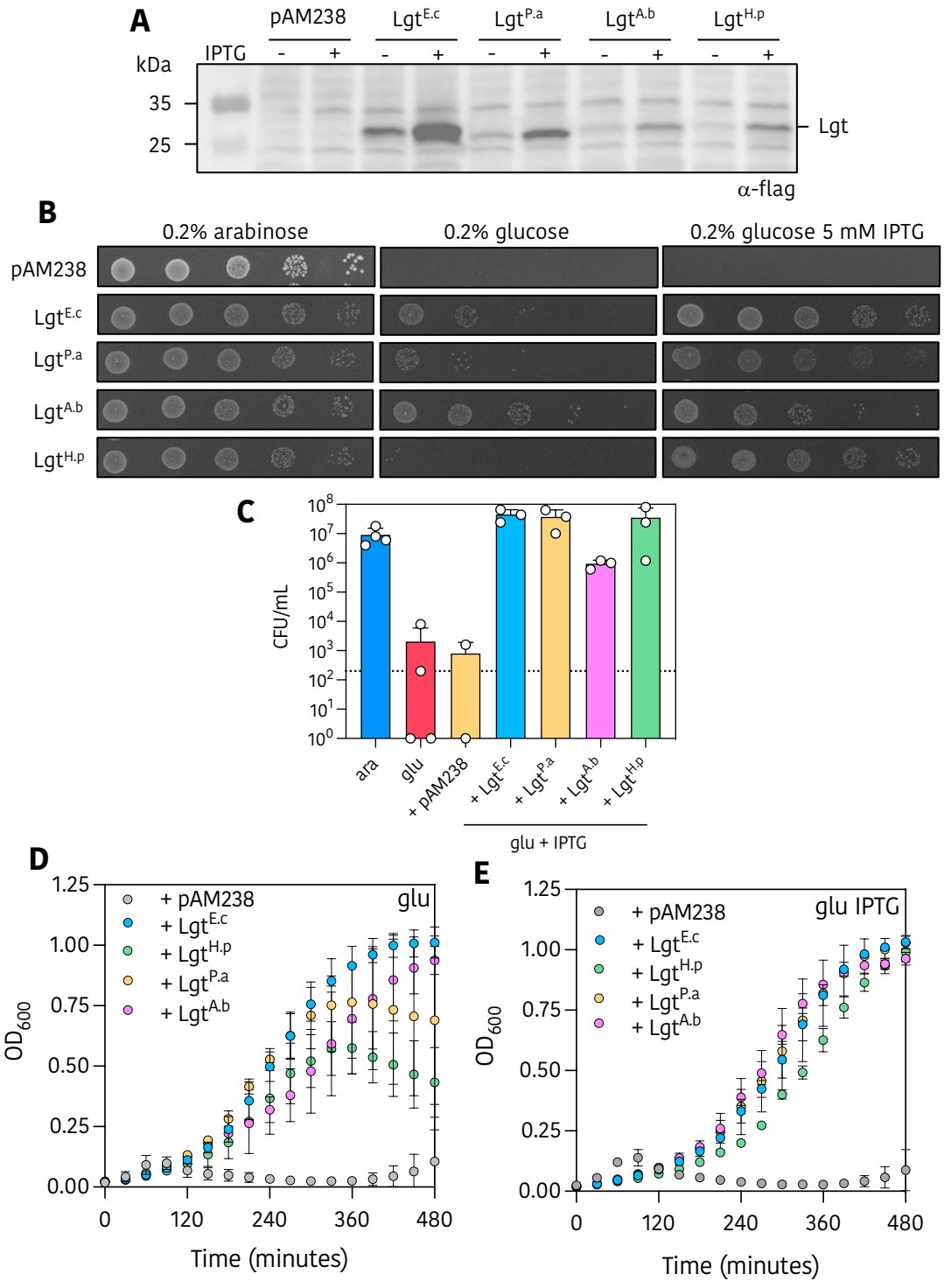


Figure 26. Complementation analysis in Δlgt^{P}. *A) Production of Lgt-flag from pAM238 plasmids in the presence of or absence of IPTG, analysed by Western-blot with α-flag antibodies. Total cell lysates corresponding to 0.1 OD units were applied to each lane. B) Spot dilution assay of depletion strains in the presence of arabinose, glucose or IPTG. C) Quantification of spot dilution assay with CFU/ml plotted. D) Growth kinetics in the presence of D-glucose. E) Growth kinetics in the presence of D-glucose and 5 mM IPTG. Growth and viability assays completed in at least duplicates, error bars indicate standard deviation from the mean.*

Both cell viability and growth kinetics analysis were performed for each of the two depletion strains. For cell viability, overnight cultures grown in the presence of L-arabinose were washed to remove residual L-arabinose. The cells were subjected to a period of Lgt depletion whereby they were grown for two hours in the absence of any inducer. The depleted cultures were then diluted and added to LB agar containing L-arabinose to express wild type *lgt*, D-glucose to repress wild type *lgt* or finally glucose and IPTG to express the complementing *lgt*.

For Δlgt^{P}, all strains grew in the presence of L-arabinose (wild type Lgt). The presence of empty pAM238 had no effect on viability (Figure 26B). When wild type Lgt was repressed by D-glucose, a partial restoration of viability is observed when Lgt from *E. coli, A. baumannii* and *H. pylori* are present on the complementing pAM238 plasmid. This effect is likely due to the basal expression from P_{lac} providing sufficient Lgt to allow growth of the cells. In the presence of IPTG, Lgt from *E. coli, A. baumannii, P. aeruginosa* and *H. pylori* all restored viability where an empty pAM238 did not. *A. baumannii* Lgt restored growth to a lesser extent than the others (Figure 26B & C) but production of $Lgt^{A.b}$ appears the lowest when compared to the others (Figure 26A).

For Δlgt^{C}, viability was not restored in the presence of glucose suggesting that low levels *lgt* expression from P_{lac} are not sufficient to restore growth when there is a single chromosomal copy of wild type *lgt* present (Figure 28). This suggests that in Δlgt^{P} the basal expression of *lgt* from P_{lac} was not the sole factor for restored growth when grown in glucose but basal expression from P_{ara} is also likely involved. Interestingly, viability was also not restored for $Lgt^{P.a}$ or $Lgt^{A.b}$ but was for $Lgt^{E.c}$ and $Lgt^{H.p}$ in the presence of IPTG, albeit to a lesser extent for $Lgt^{H.p}$. In all cases $Lgt^{E.c}$ is produced to a greater extent than the complementing Lgt proteins.

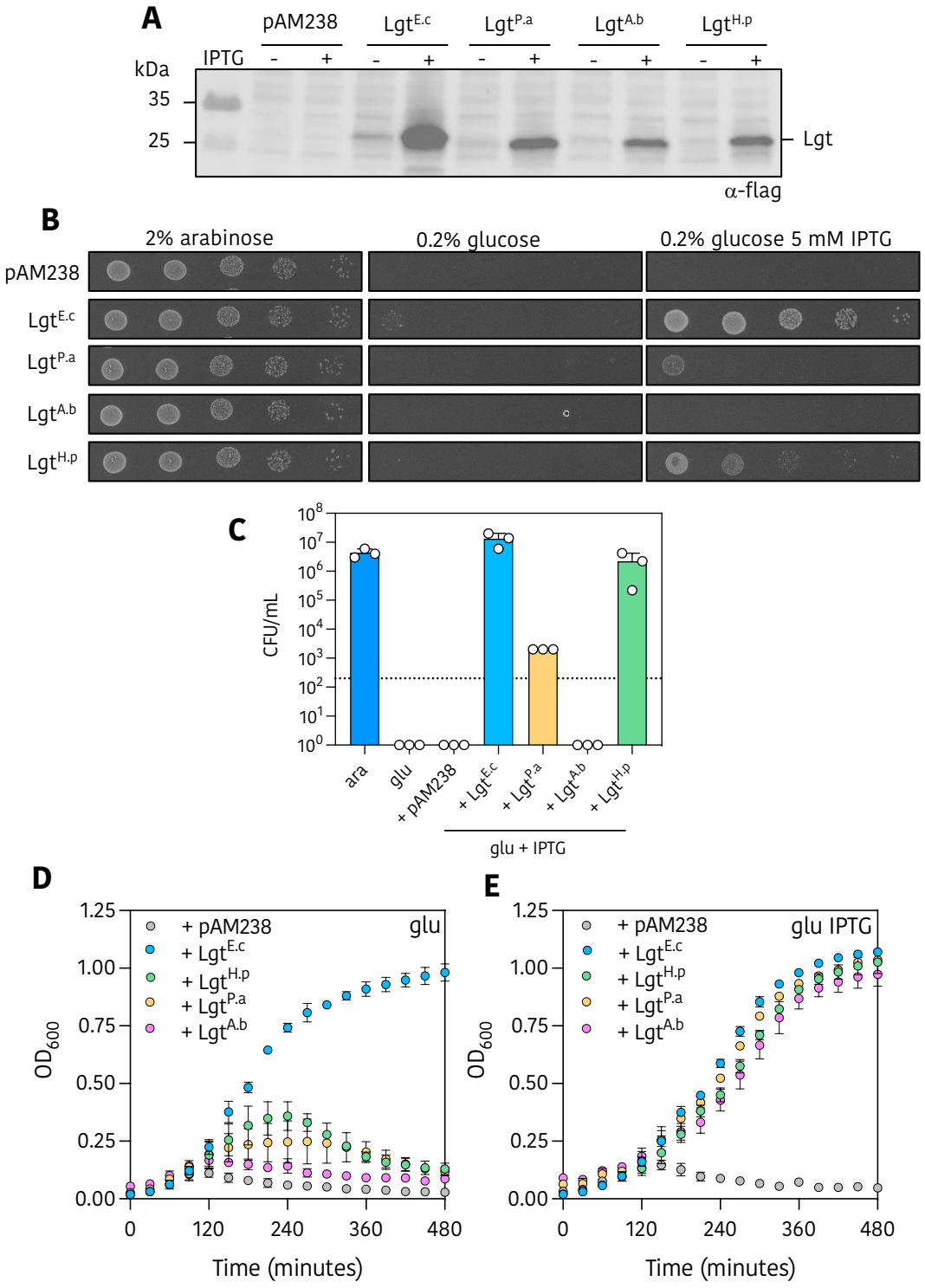


Figure 27. Complementation analysis in ΔlgtP. *A) Production of Lgt-flag from pAM238 plasmids in the presence of or absence of IPTG, analysed by Western-blot with α-flag antibodies. Total cell lysates corresponding to 0.1 OD units were applied to each lane. B) Spot dilution assay of depletion strains in the presence of arabinose, glucose or IPTG. C) Quantification of spot dilution assay with CFU/ml plotted. D) Growth kinetics in the presence of D-glucose. E) Growth kinetics in the presence of D-glucose and 5 mM IPTG. Growth and viability assays completed in at least duplicates, error bars indicate standard deviation from the mean.*

The second analysis of the complementing Lgt enzymes was their ability to restore growth in liquid media. For this assessment, the strains were depleted in the manner described for the cell viability assay. The cultures were then diluted to an optical density at 600 nm (OD600) of approximately 0.05 and grown under the three conditions as above.

As with viability, growth was restored in all Δ*lgt*^{P} strains when the complementing Lgt was actively expressed in the presence of IPTG (Figure 26E). There is also restoration of growth when they are not actively expressed (Figure 26D), further adding to the hypothesis that low level expression from P_{lac} is sufficient to partially restore growth. The growth under P_{ara} repressed conditions is more variable with Lgt$^{H.p}$ and Lgt$^{P.a}$ showing a reduction in OD after 360 minutes.

In the presence of the empty pAM238 plasmid in Δ*lgt*^{P}, growth begins to be restored after 450 minutes (Figure 26D/E). This phenomen is discussed in greater detail in Section III but is likely partly due to low level Lgt production of Lgt from P_{ara}.

In contrast to the viability assay, complementing Lgt enzymes in Δ*lgt*^{C} all restored growth in the presence of IPTG. In the absence of IPTG, growth appears partially restore until 120-240 minutes but the OD drops dramatically after this point suggesting the basal P_{lac} and P_{ara} expression of *lgt* is not sufficient for growth (Figure 27).

Due to the essential nature of Lgt, a simple knock-out strain was not possible and therefore both of the strains used require a copy of the wild type *lgt* and in each instance this is induced by arabinose. Although no production of Lgt from P_{ara} is observed at detectable levels when repressed by glucose (Section III, Figure 2) some leakage is still likely. This is therefore higher in Δ*lgt*^{P} due to the presence of P_{ara}-*lgt* on a high copy number plasmid and lower in Δ*lgt*^{C} as P_{ara}-*lgt* is present as a single copy on the chromosome. It is possible therefore to interpret the better restoration of growth and viability in Δ*lgt*^{P} as being due not wholly to the complementing Lgt but the ability of the complementing Lgt to provide sufficient diacylglycerol transfer ability in addition to that provided by basal expression from P_{ara}.

However, the results may contradict what has been described in the literature. The same depletion strain of Δ*lgt*^{C} generated in CFT073, an EPEC strain, was complemented by Lgt from *P. aeruginosa* and *A. baumannii* [75] whereas we observed the laboratory strain MG1655 was not complemented under the same conditions. Diao *et al.* (2021) measure CFU/ml from liquid media over a time course whereas our study measures CFU/ml at effectively time zero and we then follow the growth by measuring the OD. OD does not directly correlate to CFUs. However, as Lgt$^{E.c}$ was sufficient to complement Δ*lgt*^{C} in MG1655 and Lgt$^{H.p}$ did restore growth with similar levels of protein production as Lgt$^{P.a}$ and Lgt$^{A.b}$, it

seems likely that $Lgt^{P.a}$ and $Lgt^{A.b}$ are not sufficient to restore viability because they may differ more from $Lgt^{E.c}$ and $Lgt^{H.p}$. This result is surprising as $Lgt^{P.a}$ and $Lgt^{A.b}$ are more closely related to $Lgt^{E.c}$ in phylogeny.

Head domain exchanges

As the head domains showed the greatest variation in structure and some had high predicted PPIs we hypothesised that this region may have a species specific role. We therefore selected three species, *M. tuberculosis, S. aureus* and *H. pylori* as representatives of actinobacteria, firmicutes and distant proteobacteria, respectively (Figure 23). The head domains of these species were cloned into the pAM238 plasmid, replacing the head domain from Lgt$^{E.c}$. They were then transformed into Δ*lgt*^{P} and Δ*lgt* ^{C}. Protein production, cell viability and growth assay were conducted as described above. In each instance, Lgt-flag is produced upon induction by IPTG (Figure 28, 29).

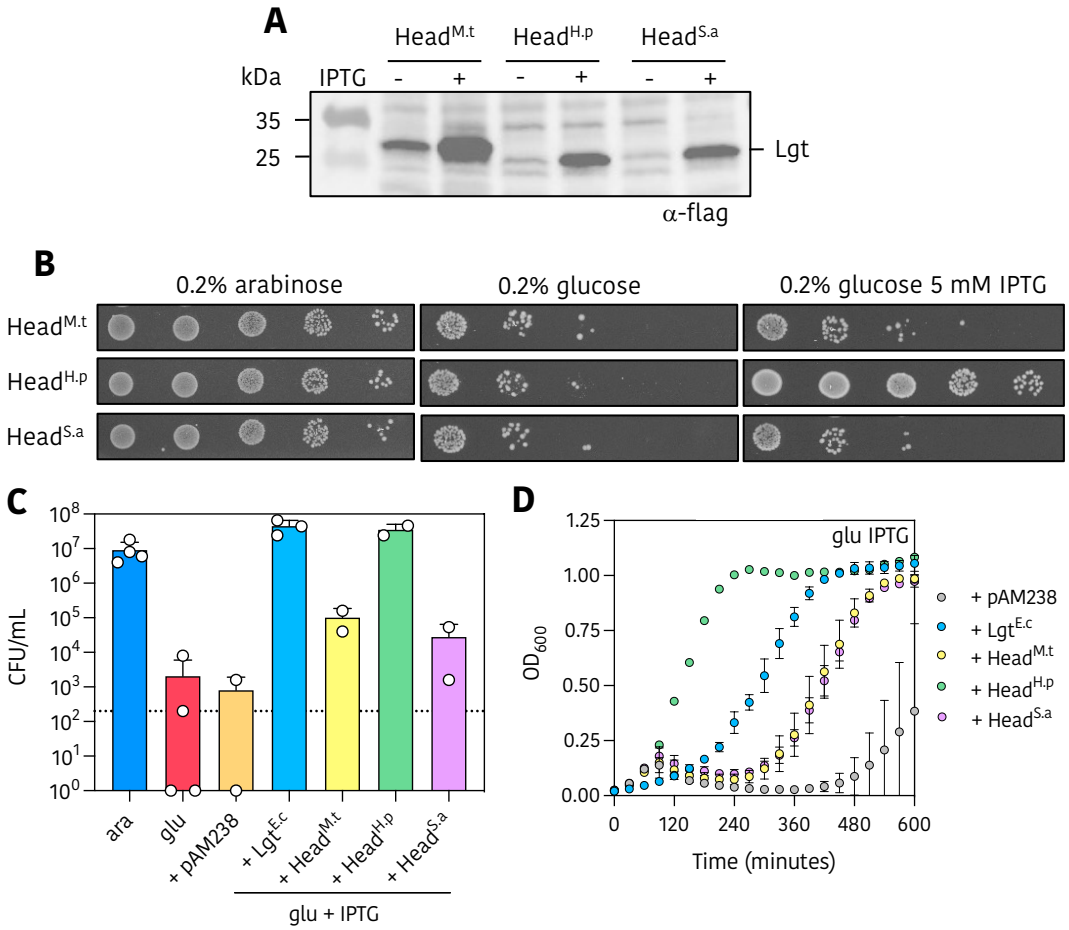


Figure 28. Head domain exchanges in Δlgt ^{P}***.*** *A) Production of Lgt-flag from pAM238 plasmids in the presence of or absence of IPTG, analysed by Western-blot with α-flag antibodies. B) Spot dilution assay of depletion strains in the presence of arabinose, glucose or IPTG. C) Quantification of spot dilution assay with CFU/ml plotted. D) Growth kinetics in the presence of D-glucose and 5 mM IPTG. Growth and viability assays completed duplicate, error bars indicate standard deviation from the mean.*

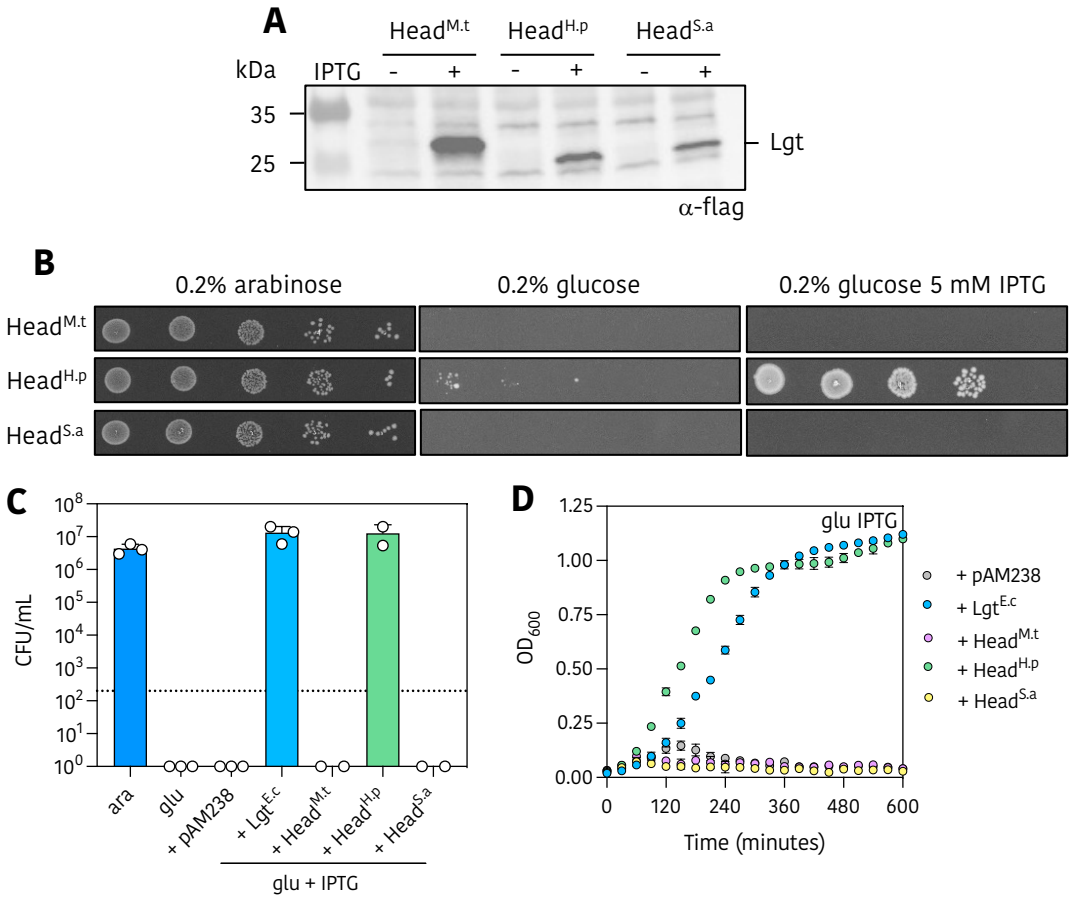


Figure 29. Head domain exchanges in Δlgt ^{C}. *A) Production of Lgt-flag from pAM238 plasmids in the presence of or absence of IPTG, analysed by Western-blot with α-flag antibodies. B) Spot dilution assay of depletion strains in the presence of arabinose, glucose or IPTG. C) Quantification of spot dilution assay with CFU/ml plotted. D) Growth kinetics in the presence of D-glucose and 5 mM IPTG. Growth and viability assays completed duplicate, error bars indicate standard deviation from the mean.*

Viability was partially restored in Δ*lgt*^{P} when the head domain from *M. tuberculosis* and *S. aureus* were present in the *E. coli* Lgt and fully restored when *H. pylori* head domain was present (Figures 28B). The same is observed with growth kinetics, Lgt containing the head domain of *H. pylori* has growth restored rapidly whereas *M. tuberculosis* and *S. aureus* head groups show a delayed growth. This delayed growth is the same in the presence of glucose only (Appendix VII), suggesting increasing expression of the modifed *lgt* genes does improve growth or viability.

In Δ*lgt* ^{C} viability and growth were restored only when the head domain from *H. pylori* was present.

Although *H. pylori* is the more closely related of the three strains selected, the structure of the head domain is more similar to that of *S. aureus*. It is therefore surprising that one can complement the *E. coli* depletion strains and the other cannot. The sequences of the two smaller head domains are not similar and therefore specific residues may be key to providing the answer. It should be noted that we have not assessed whether or not the complementing enzymes are correctly inserted into the membranes and therefore cannot be sure that they are correctly folded, located and orientated.

Interestingly, $Head^{H.p}$ grows better than $Lgt^{H.p}$ with basal P_{lac} and P_{ara} (i.e. in the absence of IPTG) in Δlgt^{C} which demonstrates that the head domain is not the sole factor in reduced ability to complement this depletion strain.

Appendix VII has further analysis of each strain grown in all conditions.

Conclusion

Lgt has been described as a well conserved enzyme that is present in all bacteria. Our phylogenetic analysis demonstrates Lgt from key AMR pathogens as following a similar conservation to that of 16s rRNA with strains grouping according to their known taxonomy. This is unlike lipoprotein modification enzyme Lsp which sees evolutionary diversion from the 16s rRNA phylogeny, with ε-proteobacteria differing greatly.

We find that, in concurrence with previous studies, Lgt is highly conserved at a sequence level with a number of residues conserved across all of the selected pathogens [78, 80, 81]. Among these residues is the Lgt signature motif with many residues thought to be key to Lgt activity [79, 81]. As previously noted, H^{103} of the HGGL motif is not broadly conserved yet is proposed to be a key residue in the predicted enzymatic activity [79]. The role of this residue in firmicutes or actinobacteria should be explored further. From previous studies the only complementation study of an Lgt without a histidine at position 103 is *C. glutamicum* Lgt (W^{103}) into *E. coli* which found that it could not complement growth in Δlgt^{P} but *E. coli* could complement lipoprotein localisation in a *C. glutamicum* depletion strain [111]. As H^{103} has been shown to be a key residue due to lack of growth after amino acid substitution [81, 82] to Q, N, R or Y in *E. coli*, this residue should be explored in species where H^{103} is not the natural residue. For example, is Y^{103} of *S. agalactiae* an essential residue, as H^{103} is in *E. coli* or can it be substituted for histidine. This may give a better understanding of the role of H^{103}.

We propose a new motif, the neck motif, found below the head domain and adjacent to the Lgt signature motif, with strict conservation within proteobacteria and within firmicutes. These residues may provide species specificity or have a functional role that has yet to be explored. Alongside this, there are a number of residues that have been shown to be essential in *E. coli* [81] but are not strictly conserved. A deeper look at the role of these residues may provide further information regard enzyme function.

Using AlphaFold2 structural predictions of the key AMR pathogen's Lgt enzymes, we observe that they are well conserved with little variation across the body of the enzyme. The one region that showed high variation is the head domain which is exposed to the periplasm. Three basic forms are observed, a large head domain as seen in *E. coli* and the Enterobacterialles, a smaller head domain observed in firmicutes and many proteobacteria and a larger head domain in *M. tuberculosis*. The head domain of *E. coli* had low predicted protein-protein interactions but the head domains of *S. aureus, H. pylori* and *M. tuberculosis* had higher predicted PPI. Initially, we hypothesised that the head domain had a role

in localising the different components of the modification pathway but ColabFold predictions deemed this unlikely. The role of the head domain is yet to be discussed in the literature.

We observe that the genetic context of *lgt* places it generally in an operon with *thyA* in proteobacteria and *hprK* in firmicutes. However, there are some variations to this but *lgt* is not found near genes considered to be related to lipoprotein modification or near the genes of lipoproteins. There is little known about the regulation of Lgt but more and more open-source data from transcriptomics and proteomics studies are becoming available and therefore may provide a source of information about the regulation of Lgt under tested conditions.

Finally, we assess the functional conservation of Lgt by employing two depletion strains, Δlgt^{P} and Δlgt^{C}. After some difficulty cloning *lgt* genes from firmicutes in an *E. coli* background, a phenomenon which requires further exploration, we cloned *lgt* from *A. baumannii, P. aeruginosa* and *H. pylori* in the depletion strains. Under IPTG induced production of the complementing Lgt enzymes, we observe growth in liquid media is restored by all the enzymes in both depletion strains. However, $Lgt^{P.a}$ and $Lgt^{A.b}$ failed to complement Δlgt^{C} when viability was analysed by spot dilution assay. Diao *et al.* (2021) showed that $Lgt^{P.a}$ and $Lgt^{A.b}$ could successfully complement Δlgt^{C} in a pathogenic strain. However, their assay is performed as CFU counts at different time points after growth in media containing IPTG. Therefore, it is not a direct comparison of our assay. The difference in the ability to grow on solid media compared to liquid media should be explored further.

We assessed whether the head domains from *S. aureus, H. pylori* and *M. tuberculosis* could be successfully cloned into *E. coli* Lgt and whether or not they would complement growth and viability. Interestingly, only $Lgt^{H.p}$ complemented growth fully. *H. pylori* is the more closely related of the species tested but contains a smaller head group more akin in structure to that of *S. aureus* albeit with a completely different sequence. These data suggest that head domain plays a role in the functionality of the enzyme and may provide species specificity. However, without further assessment as to the localisation and folding of the newly cloned hybrid Lgt enzymes, we cannot be sure that the change in the head domain is the sole cause of the loss in viability and growth.

Lgt is well conserved in its predicted catalytic domain across of the key AMR pathogens. If an inhibitor were to target this site it may have broad spectrum activity. However, our discovery that the head domains may have a functional role that is species specific could open avenues for more narrow spectrum inhibition.

Materials and Methods

Selection of strains for analysis

Species were selected due to their presence as a WHO Priority Pathogen [9], ESKAPE pathogen [10], CDC Threat Report pathogen [11], or in a recent review of global AMR impact [8]. From this list representative strains were selected from the SYntTax database [240]. The first on the list was initially selected unless a known reference strain was present. This sequence and synteny was then compared with 9 other strains using the SyntTax web tool [240] to confirm whether it was representative. If there was variation in synteny or sequence, up to 50 more strains were analysed. From this a suitable representative strain was selected.

Table 8. Selected key AMR strains

Mycobacterium_tuberculosis_H37Rv_aa1959552
Clostridioides_difficile_020474_aa35977751
Staphylococcus_aureus_014S_SA_aa168942851_C1
Enterococcus_faecalis_092160007_3_aa133446651
Enterococcus_faecium_116_aa182791451_C1
Streptococcus_pneumoniae_11A_aa28139551
Streptococcus_agalactiae_NCTC8184_aa900636751
Streptococcus_pyogenes_BSAC_bs1388_aa144959451
Campylobacter_jejuni_1_aa9125797051
Helicobacter_pylori_26695_1_aa8269851
Neisseria_gonorrhoeae_CT530_aa236117851
Acinetobacter_baumannii_2014BJAB1_aa170982052_C1
Pseudomonas_aeruginosa_PAO1_aa67651
Haemophilus_influenzae_10810_aa2108751
Proteus_mirabilis_HI4320_aa699651_C1
Morganella_morganii_11759_aa181414651
Serratia_marcescens_11_2010_aa134261551_C1
Klebsiella_pneumoniae_ATCC BAA 2146_aa228708051
Salmonella_enterica_subsp_enterica_serovar_Typhimurium_01ST04081_aa63849951_C1
Citrobacter_freundii_111_aa133902451_C1
Escherichia_coli_K_12_BW25113_aa155348551
Shigella_flexneri_113_aa97383051

Multiple sequence alignments and sequence conservation

All multiple sequence alignments were carried out on Geneious Prime (2022.1.1) using the ClustalO method. Percentage conservation of residues was calculated as the percentage of identical residues at a known residue position. Unless otherwise stated, residue number is in relation to that of *E. coli* Lgt. Percentage identity scores were exported into GraphPad Prism (9.4.1) to produce heat maps.

Phylogeny

Phylogeny was carried out on aligned sequences as described above. For Lgt and Lsp phylogeny, the Lgt or Lsp amino acid sequences were used. For 16s RNA analysis, BLAST search of the selected genomes with 16s rRNA from *E. coli* was used to determine the gene and MSA were carried out as above on the nucleotide sequence. The trees were produced using Geneious Prime (2022.1.1) and UPGMA alignment with bootstrap analysis and visually prepared in Adobe Illustrator.

Structural comparisons

The solved X-ray crystal structure of Lgt used in this study were from Mao *et al.* (2016), PBD = 5AZB (form-1), 5AZC (form-2). For structure predictions, AlphaFold2 was used [243]. Lgt sequences from the selected strains (Table 8) were queried in the ColabFold web tool. Initially, *E. coli* was run with and without the template setting to compare if this influenced output. All other structures were run without the template option. Predicted and solved Lgt structures were compared visually via ChimeraX (version 1.5). Confidence metrics were analysed for each prediction and regions of low confidence were noted.

ColabFold [244] was used in this study to compare *E. coli* Lgt interactions with Lsp , Lnt and Lpp.

ScanNet predictions

ScanNet [246] was employed to predict the likely protein-protein binding sites of Lgt from *E. coli, M. tuberculosis, H. pylori and S. aureus.* AlphaFold2 predicted structures were used as query structures.

Synteny

Local synteny was determined by SyntTax web tool [240]. Selected strains were queried against Lgt from *E. coli* and local genome output (approximately 10 genes up and down stream) were analysed.

Closer analysis of predicted operons was conducted via BioCyc and MicrobesOnline operon prediction tools [247, 248]. As specific strains from Table 8 were not always available, a reference strain was selected when required. The two outputs were compared.

Bacterial Strain and vector development

Standard growth conditions

Unless otherwise stated LB or LB agar (LBA) were used as growth medium. All strains containing the pAM238 plasmid were grown with additional spectinomycin (50 mg/mL, Spec50). Strains with pBAD18 were growth with additional chloramphenicol (30 mg/mL, Cm30). Δ*lgt* ^{P} was grown with 0.2% L-arabinose to induce WT *lgt* expression and Δ*lgt* ^{C} with 2% arabinose.

Construction of low-copy Lgt complementing plasmids

pCHAP9246 (pAM238-*lgt*$^{E.c}$-*c-myc*$_2$) [80] was used as a template to generate further low-copy (4-5 copies per cell) IPTG inducible complementation plasmids based on the pAM238 vector. The *c-myc*$_2$-tag was replaced with a *flag*$_3$-tag as follows: P12 was digested with XbaI and HindIII by incubation at 37°C for 2 hours. The digested vector was migrated on a 1% agarose gel for 45 minutes at 100 V and excised from the gel and purified with a Qiagen Gel Purification kit. Proligo Primers (FLAG) were prepared by the addition of 1 mM $MgCl_2$ in the forward/reverse primer mix and boiled for 10 minutes before cooling slowly to allow hybridisation. The flag insert was ligated into the linearised plasmid by incubation with T4 DNA ligase (NEB) in T4 DNA ligase buffer (NEB) overnight at 16°C. Ligated vectors were transformed into chemically competent *E. coli* BW25113 cells and selected for spectinomycin resistance on LBA-Spec50 plates incubated at 37°C overnight. Clones were checked by PCR with M13 F/R primers and sequenced to confirm correct insertion, the plasmid was named SLP14 (pAM238-*lgt*$^{E.c}$-*c-flag*$_3$).

To insert *lgt* genes from different species into pAM238, P14 was used as a template for gene insertion, which was digested with XbaI and EcoRI by incubation at 37°C for 4 hours. Digested vector was migrated on a 1% agarose gel for 45 minutes at 100 V, excised from gel and purified with a Qiagen Gel Purification kit. Individual genes were amplified by PCR from chromosomal DNA with primers listed in Table 9. PCR amplified inserts were purified with Qiagen PCR purification kit. Insertion of the insert was conducted by Gibson Assembly (NEB) and incubated for 15 minutes at 50°C. Gibson clones were transformed into chemically competent BW25113 cells and selected on LBA-Spec50 incubated at 37°C overnight. Clones were confirmed by PCR with M13 F/R primers and sequenced to confirm correct insertion of the corresponding *lgt* gene.

All clones were transformed into chemically competent SLEC21 (Δ*lgt* ^{P}) and SLEC67 (Δ*lgt* ^{C}) cells.

Construction of plasmid p14 derivatives with head-domain exchange

P14 was used as a template and was linearized by PCR with primers pr136 and pr137 (Table 9). The primers amplified the vector 'outwards' from the flanking regions of the head-domain of *E. coli* Lgt. Head domains from *M. tuberculosis*, *S. aureus* and *H. pylori* were amplified from the templates with primers pr136-143 (Table 9). Insertion of the insert was conducted by Gibson Assembly as described above.

Expression of *lgt* from pAM238

To ensure *lgt* expression from pAM238 clones in the Lgt depletion strain Δ*lgt* ^{P} and Δ*lgt* ^{C}, Lgt production was analysed by immunoblotting using detection with α-Flag antibodies (Sigma). Single colonies were grown overnight in 5 mL LB + L-arabinose at 37°C. Overnight cultures were diluted 1/100 in fresh LB + L-arabinose and grown for 4.5 hours at 37°C to early exponential phase. The cultures were centrifuged at 13,200 x *g* in a table-top centrifuge and the supernatant was removed. The pellet was resuspended in sample buffer (10% glycerol, 2.5% SDS, 100 mM Tris pH8, 10 mM DTT, phenol red) to a ratio of 0.01 OD_{600} units / μL. Samples were heated at 100°C for 10 minutes before loading onto a SDS-PAGE Stain-Free 4-15% gel (BioRad) and migrated at 20 mA. Prior to transfer of the proteins onto nitrocellulose membranes, gels were imaged under stain free conditions to ascertain quality of sample loading. Western blot was conducted on nitrocellulose membranes using a BioRad TurboBlot instrument. The blot was rinsed in water before incubation with PonceauS solution for 1 minute to verify transfer of the proteins and to mildly precipitate proteins due to the presence of low concentration of acetic acid. The PonceauS was washed away with water and the blot was incubated in blocking buffer (5% BSA, 1 x PBS, 0.5% Tween 20) for 1 hour at room temperature (RT) before being replaced by an α-Flag (1:10,000) antibody solution (1x TBS, 0.1% Tween 20, 1% BSA) incubated for 1 hour at RT. The blot was washed 2 x for 5 minutes and 3 x 10 minutes with 0.1% TBS-T. Secondary α-Mouse-HRP (Sigma 1:10,000) in buffer (1x TBS, 0.1% Tween 20) was incubated with the blot for 1 hour. 2x 5 minutes and 3x 10 minutes washes were completed before adding ECL Western blotting chemiluminescence substrate as per manufacturer instructions (Pierce). The blot was imaged on a ChemiDoc imager (Bio-Rad) and further analysed using ImageLab software (Bio-Rad).

Spot dilution assay to determine colony-forming units

Selected strains were grown overnight from a single colony in LB-Spec-Cm or LB-Spec as required with 0.2% (Δlgt^{P}) or 2% (Δlgt^{C}) L-arabinose. All cultures were washed 3x in LB by repetitive centrifugation and resuspension steps. Finally, they were diluted 1/100 in LB and grown for 2 hours to deplete Lgt. Optical density at 600 nm (OD_{600}) was measured and cultures were diluted to an OD_{600} of 0.1. Calibrated cultures were then serially diluted 1/10 in LB to 10^{-5} before 5 µL was spotted onto LBA plates. LBA plates were supplemented with 2%, 0.2% L-arabinose, 0.2% D-glucose or 5 mM IPTG. Plates were incubated overnight at 37°C and imaged on a ChemiDoc imager (Bio-Rad) before CFU/mL was determined.

Liquid growth kinetics in 96-well plate format

Selected strains were grown, washed and Lgt was depleted as described above. OD_{600} was recorded and cultures were diluted to an OD_{600} of 0.1 before 100 µL was added to each well of a clear, flat-bottomed 96 well-plate. 100 µL of LB supplemented with 4%, 0.4% L-arabinose, 0.4% D-glucose or 10 mM IPTG, as required, was added leaving final concentrations of 2%, 0.2% L-arabinose, 0.4% D-glucose or 5 mM IPTG in each well. Samples were prepared in duplicate. Plates were incubated in a TECAN plate reader overnight at 37°C and OD_{600} measurements were taken at a 15-minute interval for 12 hours with shaking between measurements. Results were collected and analysed on GraphPad Prism version 9.3.1.

Section II : Strains, plasmids and primers

Table 9. Strains, plasmids and primers for Section II

Strains	Alt name	Description	Reference
BW25113	SLEC30	*E. coli* K-12 $lacI^{q}$ $rrnB_{T14}$ $\Delta lacZ_{WJ16}$ *hsdR514* $\Delta araBAD_{AH33}$ $\Delta rhaBAD_{LD78}$	[249]
Δlgt^{P}	SLEC21, PAP9403	E. coli BW25113 Δlgt::Kan^{r} + pBAD18s-Cm-lgt^{WT}-c-myc_2	[80]
Δlgt^{P} + pAM238	SLEC33	Δlgt^{P} + pAM238	This study
Δlgt^{P} + $Lgt^{E.c}$	SLEC23	Δlgt^{P} + pAM238-$Lgt^{E.c}$-*flag*	This study
Δlgt^{P} + $Lgt^{P.a}$	SLEC24	Δlgt^{P} + pAM238-$Lgt^{P.a}$	This study
Δlgt^{P} + $Lgt^{A.b}$	SLEC27	Δlgt^{P} + pAM238-$Lgt^{A.b}$	This study
Δlgt^{P} + $Lgt^{H.p}$	SLEC28	Δlgt^{P} + pAM238-$Lgt^{H.p}$	This study
Δlgt^{P} + $Head^{M.t}$	SLEC80	Δlgt^{P} + pAM238-$Head^{M.t}$	This study

Δ*lgt* ^{P} + Head$^{H.p}$	SLEC81	Δ*lgt* ^{P} + pAM238-Head$^{H.p}$	This study
Δ*lgt* ^{P} + Head$^{S.a}$	SLEC82	Δ*lgt* ^{P} + pAM238-Head$^{S.a}$	This study
Δ*lgt* ^{C} + pAM238	SLEC76	Δ*lgt* ^{C} + pAM238	[250]
MG1655		F-, λ-, *ilv*G- *rfb-50 rph-1*	Lab collection
Δ*lgt* ^{C}	SLEC67	*E. coli* MG1655 Δ*lgt*::Kan^{r} λattB-pBAD-*lgt*	[75]
Δ*lgt* ^{C} + Lgt$^{E.c}$	SLEC71	Δ*lgt* ^{C} + pAM238-Lgt$^{E.c}$-*flag*	[250]
Δ*lgt* ^{C} + Lgt$^{P.a}$	SLEC77	Δ*lgt* ^{C} + pAM238-Lgt$^{P.a}$	This study
Δ*lgt* ^{C} + Lgt$^{A.b}$	SLEC78	Δ*lgt* ^{C} + pAM238-Lgt$^{A.b}$	This study
Δ*lgt* ^{C} + Lgt$^{H.p}$	SLEC75	Δ*lgt* ^{C} + pAM238-Lgt$^{H.p}$	This study
Δ*lgt* ^{C} + Head$^{M.t}$	SLEC83	Δ*lgt* ^{C} + pAM238-Head$^{M.t}$	This study
Δ*lgt* ^{C} + Head$^{H.p}$	SLEC84	Δ*lgt* ^{C} + pAM238-Head$^{H.p}$	This study
Δ*lgt* ^{C} + Head$^{S.a}$	SLEC85	Δ*lgt* ^{C} + pAM238-Head$^{S.a}$	This study
Plasmids			
pAM238		pSC101 origin, Plac promoter, Spc^{r}	[251]
pAM238-Lgt$^{E.c}$-*myc*	pCHAP9246	pSC101 origin, Plac promoter, Spc^{r}, expressing *E. coli lgt* with myc$_2$ tag	[80]
pAM238-Lgt$^{E.c}$-*flag*	SLP14	pSC101 origin, Plac promoter, Spc^{r}, expressing *E. coli lgt* with flag$_3$ tag	[250]
pAM238-Lgt$^{P.a}$		pAM238-Lgt$^{E.c}$-*flag* with *E. coli lgt* exchanged for *lgt* from *P. aureigonsa* PA01	This study
pAM238-Lgt$^{A.b}$		pAM238-Lgt$^{E.c}$-*flag* with *E. coli lgt* exchanged for *lgt* from *A. baumannii* AYE	This study
pAM238-Lgt$^{H.p}$		pAM238-Lgt$^{E.c}$-*flag* with *E. coli lgt* exchanged for *lgt* from *H. pylori* 26695	This study
pAM238-Head$^{M.t}$		pAM238-Lgt$^{E.c}$-*flag* with *E. coli lgt* head-domain exchanged for *head domain* from *M. tuberculosis* H37Rv	This study
pAM238-Head$^{H.p}$		pAM238-Lgt$^{E.c}$-*flag* with *E. coli lgt* head-domain exchanged for *head domain* from *H. pylori* 26695	This study
pAM238-Head$^{S.a}$		pAM238-Lgt$^{E.c}$-*flag* with *E. coli lgt* exchanged for *lgt* from *S. aureus* RN4220	This study
Primers	**Template**	**Sequence 5' – 3'**	
upperFLAG		CTAGAGACTACAAAGACCATGACGGTGATTAT AAAGATCATGACATCGATTACAAGGATGACGA TGGTACCTAGA	Proligo
lowerFLAG		AGCTTCTAGGTACCATCGTCATCCTTGTAATCG ATGTCATGATCTTTATAATCACCGTCATGGTCT TTGTAGTCT	Proligo
pr48	*A. baumannii* AYE	AGCTATGACCATGATTACGAATTCATGCTGA CCTATCCTAATATCGATCCG	
pr49		ATGGTCTTTGTAGTCTCTAGTACTGTTCTTT TGAGGGCCCCA	
pr49		ATGGTCTTTGTAGTCTCTAGTGGCCGCCTTC GGC	

pr50	*P. aeruginosa* PA01	AGCTATGACCATGATTACGAATTCATGCTGA CGTATCCCCAGAT	
pr42	*H. pylori* HP26695	ATGGTCTTTGTAGTCTCTAGTTTGATTTTCC TTTATTTTTTTAGAATTTTTTGTAGCATACAA TAAAATCC	
pr43		AGCTATGACCATGATTACGAATTCatgAACGCT TGGAATACGATTTATGATCAAT	
pr136	pAM238-Lgt$^{E.c}$-*flag*	GCCCCACAATTCACCGTTAAT	
pr137		CCATCACAGCTTTACGAGCT	
pr138	*M. tuberculosis* H37Rv	ctcgtaaagctgtgatggTTGAACCACGAACGCCACC	
pr139		aacggtgaattgtggggcCGTGAAACCACCATGCCGT	
pr140	*S. aureus* RN4220	ctcgtaaagctgtgatggATGATAATATTGGCCGTTAA	
pr141		aacggtgaattgtggggcGGATCGGTGTCACGCGCTTT	
pr142	*H. pylori* HP26695	ctcgtaaagctgtgatggATAACGCAACTCATTATCCACCA	
pr143		aacggtgaattgtggggcAGAATTGTCCCCAAAGACAGC	
M13 F	pAM238	TAGTCTCTAGAGGAAAC	
M13 R		CAGGAAACAGCTATGAC	

Section III: Essentiality of Lgt in *E. coli*

Summary

The abundant lipoprotein Lpp has been shown to be a key factor in the essentiality of Lsp and Lnt. Depletion of these enzymes is lethal in presence of Lpp but not strictly so in its absence or when it lacks the ability to cross-link to peptidoglycan (PGN) while localised in the cytoplasmic membrane. A recent study suggested that this may not be the case for Lgt. This raises questions on the drug-ability of the lipoprotein modification enzymes and the efficacy of inhibition in antibiotic development. We therefore looked to understand the role of Lpp on Lgt essentiality from a physiological point of view. This work was recently published in the Journal of Bacteriology .

A defect in lipoprotein modification by Lgt leads to abnormal morphology and cell death in Escherichia coli that is independent of major lipoprotein Lpp

Legood, S., Seng, D., G. Boneca, I., Buddelmeijer, N. (2022)

Journal of Bacteriology 115(3): 356-365.

In a depletion strain harbouring a copy of *lgt* on a multicopy plasmid, bacterial growth was arrested in the absence of *lgt* expression but revertants readily appeared at 450 minutes of growth. After analysis of the modification state of Lpp, whole genome sequence analysis, and attempts to cure the *lgt*-containing plasmid from these revertants, it was concluded that inducer-independent low-level expression of *lgt* from the plasmid was sufficient to restore growth possibly involving partnership with another yet discovered factor. In this background, Lpp deletion restored growth under non-permissive conditions but did not completely restore a wild type morphology.

A second strain where a single copy of *lgt* was present on the chromosome, revertants did not appear and removal of Lpp did not restore growth under non-permissive conditions. However, growth kinetics were improved if low levels of *lgt* were present and Lpp was removed but cell division was affected and resulted in an increase in cell length.

We concluded that Lgt was essential in the absence of Lpp but Δ*lgt* did have improved survival at low concentrations of Lgt when Lpp was removed. We discuss the possible role of other lipoproteins in the essentiality of the lipoprotein modification pathway and discuss the potential of Lgt as a target for novel antibiotics.

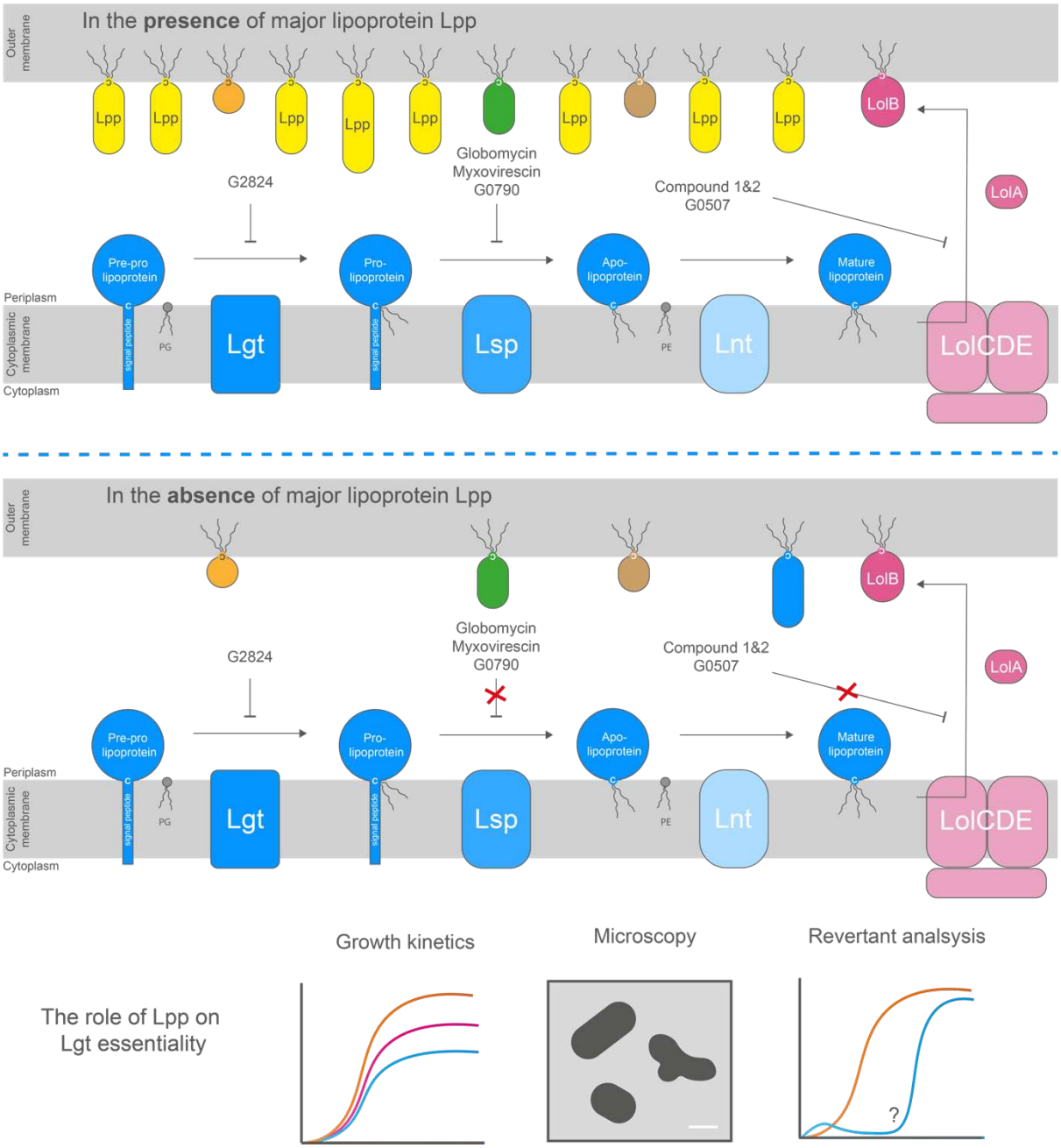


Figure 30. Graphical abstract of Section III.

A defect in lipoprotein modification by Lgt leads to abnormal morphology and cell death in *Escherichia coli* that is independent of major lipoprotein Lpp

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology®

RESEARCH ARTICLE

A Defect in Lipoprotein Modification by Lgt Leads to Abnormal Morphology and Cell Death in *Escherichia coli* That Is Independent of Major Lipoprotein Lpp

S. Legood,[a] D. Seng,[a] I. G. Boneca,[a] N. Buddelmeijer[a]

[a]Institut Pasteur, Université de Paris, CNRS UMR6047, INSERM U1306, Unité de Biologie et Génétique de la paroi bactérienne, Paris, France

ABSTRACT Lgt is an essential enzyme in proteobacteria and therefore a potential target for novel antibiotics. The effect of Lgt depletion on growth, morphology, and viability was studied in *Escherichia coli* to assess whether absence of Lgt leads to cell death. Two Lgt depletion strains were used in which *lgt* was under the control of an arabinose-inducible promoter that allowed regulation of Lgt protein levels. Reduced levels of Lgt led to severe growth and morphological defects that could be restored by expressing *lgt* in *trans*, demonstrating that only Lgt is responsible for the distorted phenotypes. In the absence of major lipoprotein Lpp, growth defects were partially restored when low levels of Lgt were still present; however, *lgt* could not be deleted in the absence of Lpp. Our results demonstrate that Lpp is not the main cause of cell death under conditions of Lgt depletion and that other lipoproteins are important in cell envelope biogenesis and cell viability. Specific inhibitors of Lgt are thus promising for the development of novel antibiotics.

IMPORTANCE Incomplete maturation and envelope mislocalization of lipoproteins, through inhibition or mutations in lipoprotein modification enzymes or transport to the outer membrane, are lethal in proteobacteria. Resistance to small-molecule inhibition or the appearance of suppressor mutations is often directly correlated with the presence of abundant outer membrane lipoprotein Lpp. Our results show that Lgt, the first enzyme of the lipoprotein modification pathway, is still required for growth and viability in the absence of Lpp and thus is necessary for the function of other essential lipoproteins in the cell envelope. This adds credence to the hypothesis that Lgt is essential in proteobacteria and an attractive target for the development of novel antibiotics.

KEYWORDS Lgt, viability, Lpp, cell envelope, *Escherichia coli*, cell viability

Lipoproteins play an important role in physiology and viability of bacteria and are involved in essential processes for the cell envelope, such as cell wall biogenesis, transport and insertion of membrane proteins and lipopolysaccharides, nutrient uptake, and efflux of toxic molecules (1–3). They are anchored in membranes through their fatty acid-linked amino termini, which also play an important role in virulence by signaling the innate immune response via interaction with Toll-like receptors (4). Lipoprotein maturation is a posttranslational process that takes place in the cytoplasmic membrane and involves three essential integral membrane proteins in proteobacteria. Lipoproteins are synthesized in the cytoplasm with a Sec or Tat membrane-targeting sequence composed of a hydrophobic signal peptide containing a specific sequence, the so-called lipobox, for recognition by the lipoprotein modification machinery. The first step in the pathway is catalyzed by phosphatidylglycerol:prolipoprotein diacylglyceryl transferase (Lgt), which transfers a diacylglyceryl moiety from the phospholipid phosphatidylglycerol to the sulfur group of the lipoprotein-specific cysteine residue of prolipoprotein, resulting in a

Editor Julie A. Maupin-Furlow, University of Florida

Address correspondence to N. Buddelmeijer, nienke.buddelmeijer@pasteur.fr.

The authors declare no conflict of interest.

Received 3 May 2022
Accepted 13 July 2022
Published 8 August 2022

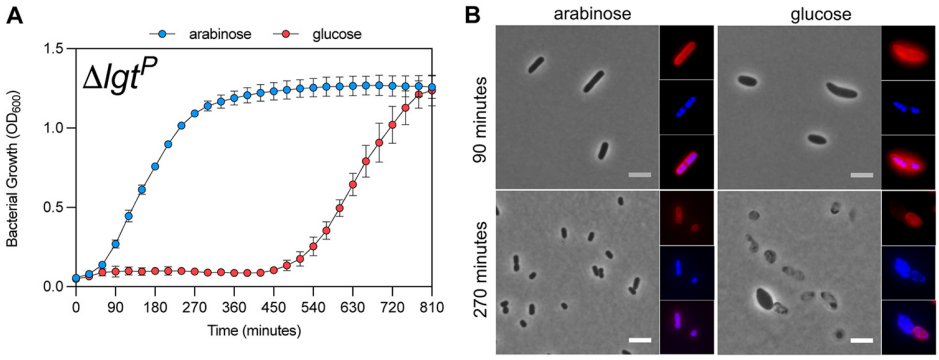


FIG 1 Lgt is essential for viability of *E. coli*. (A) Growth curve of *lgt* depletion strain (PAP9403, ΔlgtP) in 0.2% L-arabinose and 0.2% D-glucose. Time (t) 0 corresponds to 2 h of growth in LB medium without sugar. Graphed data represent duplicate OD_{600} measurements of biological triplicates. (B) Phase-contrast images of *lgt* depletion strain Δ*lgtP* grown in the presence of 0.2% L-arabinose and 0.2% D-glucose. Nucleoids were stained with Hoechst (blue), and the membrane was stained with FM4-64X (red). Bar, 5 μm.

thioether-linked diacylglyceryl-cysteine and glycerol-1-phosphate by-product (5, 6). The second step involves cleavage of the signal peptide catalyzed by signal peptidase II (Lsp), liberating the α-amine group of diacylglyceryl-cysteine (7). The third and final step in the pathway is catalyzed by apolipoprotein N-acyltransferase, which adds a third fatty acid from the *sn*-1 acyl group of phosphatidylethanolamine onto the α-amine, resulting in mature triacylated lipoprotein (8, 9). Lipoprotein modification is key for the biogenesis and maintenance of the cell envelope. In *Escherichia coli*, more than 90 lipoproteins, of which at least 80 are located in the outer membrane, have been confirmed by fatty acid acylation (10), and three outer membrane lipoproteins have been reported as essential, i.e., the lipoprotein outer membrane receptor LolB (11), the lipopolysaccharide (LPS) assembly component LptE (12, 13), and the β-barrel assembly complex component BamD (14, 15). Major outer membrane lipoprotein Lpp (Braun's lipoprotein) is not essential but is required for maintenance of the cell envelope. This highly abundant lipoprotein, with approximately 1 million copies per cell, is partly covalently cross-linked to the peptidoglycan (16) and surface exposed (17). Mislocalization of Lpp to the cytoplasmic membrane while being cross-linked to the peptidoglycan is lethal for *E. coli* (18).

Lgt catalyzes the first irreversible step in the sequential pathway of lipoprotein modification and is conserved in all bacteria (1). To better understand its role in bacterial physiology, we addressed its essentiality for viability of *E. coli*. Previous studies demonstrated that in conditions of incomplete maturation of lipoproteins, either by mutations (19) or upon inhibition of the modification enzymes by small molecules (20–23), mutations occur in *lpp* that lead to absence of the protein or the inability to cross-link to the peptidoglycan (20, 24). In this study, we showed that Lgt is essential in the absence of Lpp and that other lipoproteins play an important role in cell envelope biogenesis in *E. coli*, validating Lgt as a potential target for the development of novel antibiotics.

RESULTS

Lgt is essential for viability of *E. coli*. To address the role of major lipoprotein Lpp on the essentiality of Lgt and to confirm previous reports that showed a decrease in cell viability when *lgt* expression was reduced, we employed the reported *E. coli* Lgt depletion strain PAP9403, in which the chromosomal *lgt* gene is replaced by a kanamycin resistance cassette in a manner that does not disrupt the downstream *thyA* gene (25). To sustain growth, the *lgt* gene is present on a plasmid (pBAD18s-Cm-lgt^{WT}-c-myc_2, pLgt) under the control of an arabinose-inducible promoter (P_{ara}) (Δlgt^{P}). When PAP9403 (Δlgt^{P}) is grown in the absence of L-arabinose or in the presence of D-glucose, an initial phase of growth is observed for 90 min (5, 25), before the cultures exhibit a decrease in optical density suggesting growth arrest up to 450 min (about 20 generations) (Fig. 1). Upon depletion of Lgt, cells became wider, rounded up, lost DNA,

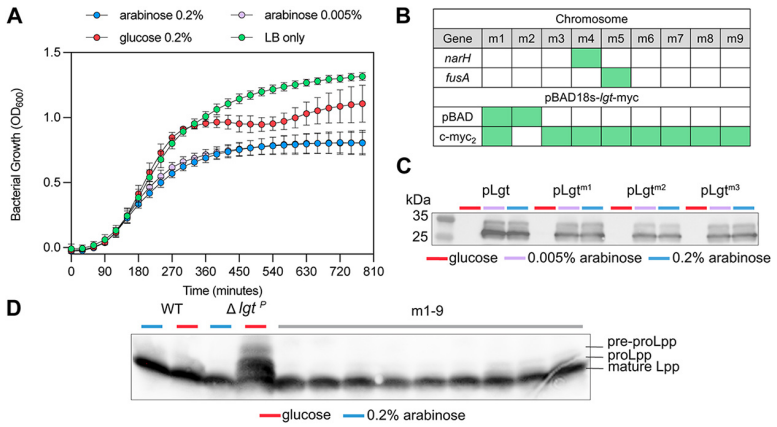


FIG 2 Mature form of Lpp is detected in revertants of the Lgt depletion strain. (A) Growth curve of revertants of strain PAP9403 (Δ*lgt*^{P}) obtained after extensive growth under restrictive conditions. Grouped growth curves of nine revertants in the presence of high (0.2%) or low (0.005%) L-arabinose, 0.2% D-glucose, or in LB medium without sugar, are presented. (B) Chromosomal and plasmid mutations identified by WGS in 9 revertant strains (Δ*lgt*^{Pm1} to Δ*lgt*^{Pm9}) are indicated as green boxes. In *narH* adenosine replaced thymine at position 322, resulting in Y108N in the protein, and *fusA* has thymine 1961 rather than adenosine, leading to I654N. For further details of pLgt mutations, see Fig. S1 in the supplemental material. (C) Detection of Lgt-*myc*$_2$ on Western blotting by anti-c-*myc* antibodies (Sigma) from wild-type strain BW25113 retransformed with plasmids from revertants m1 to m3, representing all combinations of mutations. Strains were grown in the presence of high (0.2%) or low (0.005%) L-arabinose, or 0.2% D-glucose. (D) Detection of Lpp by Western blotting using anti-Lpp antibodies. An equal amount of total cell extract (0.1 OD_{600} units) from cultures grown for 270 min was loaded per lane.

displayed a membrane defect, and finally lysed (Fig. 1). These results confirmed that *lgt* is an essential gene in *E. coli*.

Revertants of plasmid-encoded Lgt depletion are not affected in Lpp modification. We observed restored growth in the Lgt depletion strain under restrictive growth conditions at 450 min. We hypothesized that growth restoration occurred due to the absence of, or mutations in, Lpp that affected its interaction with the peptidoglycan. The arabinose-independent phenotype was maintained in all revertants from Δ*lgt*^{P} (m1 to -9), suggesting that the phenotype was passed on to future generations (Fig. 2). We performed whole-genome sequencing (WGS) to investigate mutations in *lpp* or other loci that could explain the revertant phenotype. Few single-nucleotide polymorphisms (SNPs) were identified in some of the revertants, either on the chromosome or on the pLgt plasmid (Fig. 2; see also Fig. S1 in the supplemental material). The mutations in *narH* and *fusA*, encoding nitrate reductase A subunit β and elongation factor G, respectively, were not directly involved in lipoprotein modification. No mutations were observed in *lpp* or *micL*, which encodes a small RNA regulating *lpp* (26), or in genes related to cell envelope biogenesis or the stress response. Strikingly, some revertants did not seem to have chromosomal mutations. Mutations were identified in pLgt, i.e., two revertants had an SNP in the arabinose promoter and eight had mutations in the 3′-end of the gene encoding the c-*myc* tag (see Fig. S1). This raised the possibility that mutations in the promoter of pLgt cause *lgt* expression independent of arabinose, resulting in Lgt production and therefore restoration of growth. pLgt from three revertants was isolated and transformed into wild-type strain BW25113, and clones were analyzed for Lgt production upon induction with L-arabinose and in the presence of D-glucose to repress expression of *lgt* from P_{ara} (Fig. 2). Whereas pLgt^{m1} had mutations in both the promoter region and the c-*myc* tag, pLgt^{m2} had a mutation in the arabinose promoter and pLgt^{m3} to pLgt^{m9} contained a mutation in the c-*myc* tag (see Fig. S1). Western blot analysis of Lgt-c-*myc*$_2$ revealed slightly reduced levels of Lgt from the revertant plasmids compared to the original plasmid in the presence of L-arabinose. Lgt could not be detected in any of the strains in the presence of D-glucose, possibly due to detection limits of the Western blot assay. To test whether the revertant plasmids are required for restored

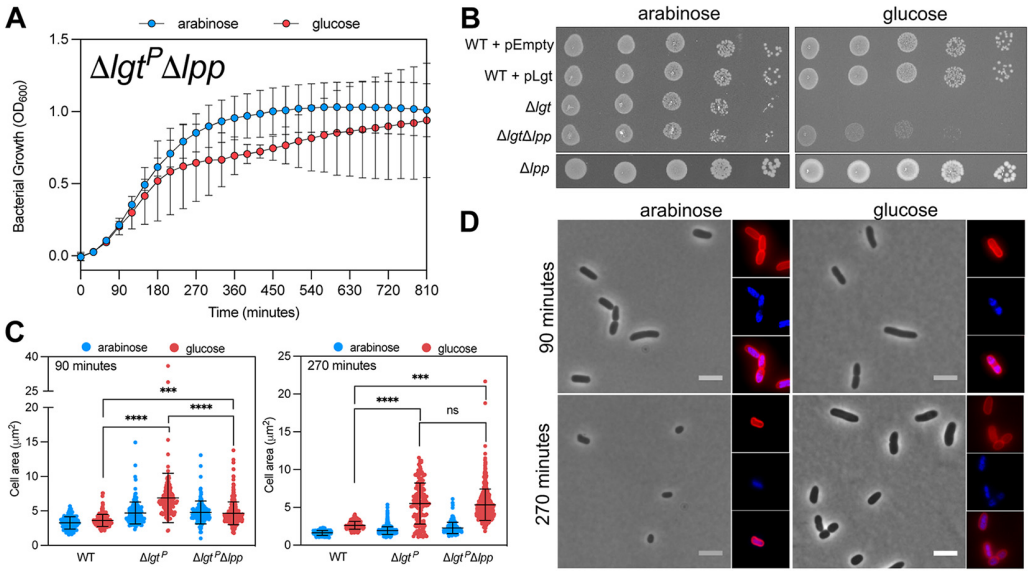


FIG 3 Lpp deletion partially restores growth of an Lgt depletion strain. (A) Strain PAP9403 *lpp*::Tn*10* (Δ*lgt*^{P} Δ*lpp*) was grown in LB medium containing 0.2% L-arabinose or 0.2% D-glucose. Graphed data represent duplicate OD_{600} measurements of biological triplicates. (B) Spot dilution assay of cultures of BW25113 + pEmpty, BW25113 + pLgt, Δ*lgt*^{P}, Δ*lgt*^{P} Δ*lpp*, and Δ*lpp* were washed with LB lacking sugar after overnight growth and grown for 2 h in LB. Cells were then spotted on plates as 5 μL of 10^{-1} to 10^{-5} dilutions. (C) Cell area measurements ($n > 100$ cells) from corresponding strains were performed on phase-contrast images using Microbe J. The WT is BW25113. ****, $P < 0.0001$; ***, $P < 0.001$. (D) Phase-contrast images of Δ*lgt*^{P} Δ*lpp* grown in the presence of 0.2% L-arabinose or 0.2% D-glucose. Nucleoids were stained with Hoechst (blue) and the membrane was stained with FM4-64X (red). Bar, 5 μm.

growth, we used the pFREE system (27). This system is based on the elimination of plasmids through expression of CRISPR-Cas9 that targets the origin of replication region in plasmids. The nine plasmids from the revertant strains could not be lost, whereas pLgt was removed from a wild-type strain (see Fig. S2). This suggests that the plasmid-linked mutations lead to arabinose-independent low-level expression of *lgt-myc*$_2$. Migration of Lpp on high-resolution gel electrophoresis revealed that all strains contained mature, triacylated Lpp, in contrast to Δ*lgt*^{P} grown with D-glucose, where accumulation of proLpp and pre-proLpp was observed (Fig. 2). This showed that maturation of Lpp was not affected in the revertants. Overall, the data suggested that very low levels of Lgt are sufficient to sustain growth, possibly in addition to an unidentified modification.

Deletion of *lpp* partially restores growth of an Lgt depletion strain. Lpp is the most abundant lipoprotein in *E. coli* and therefore a prominent substrate for the lipoprotein modification enzymes. Since very low levels of Lgt are sufficient for growth, we hypothesized that deletion of *lpp* would fully restore growth and morphology upon Lgt depletion. The double mutant Δ*lgt*^{P}Δ*lpp* grew in L-arabinose and D-glucose, but growth was highly variable and the optical density at 600 nm (OD_{600}) was lower than that for Δ*lgt*^{P} at 270 min (Fig. 3). In a spot dilution assay, colonies were smaller in Δ*lgt*^{P} Δ*lpp* compared to full expression of *lgt*, but CFU were constant, independent of L-arabinose concentration (see Fig. S3). The morphology of Δ*lgt*^{P}Δ*lpp* was less severe compared to Δ*lgt*^{P} in D-glucose, but cells were larger than wild-type cells upon entry into stationary phase (270 min) (Fig. 2). In stationary phase, the percentage of Lpp crosslinks to peptidoglycan increases (28), which might explain the observed morphology. Growth of the depletion strain was thus partially restored in the absence of Lpp and viability increased. To test whether Lgt is essential in the absence of Lpp, we first tried to cure pLgt from Δ*lgt*^{P} Δ*lpp* by using the pFREE system. Whereas pLgt was readily removed from a wild-type background, the plasmid could not be lost from Δ*lgt*^{P}Δ*lpp* (see Fig. S2). Attempts to delete *lgt* and replace it with a kanamycin cassette by P1 transduction in a *lpp*::Tn*10* strain were also unsuccessful. Together, these results

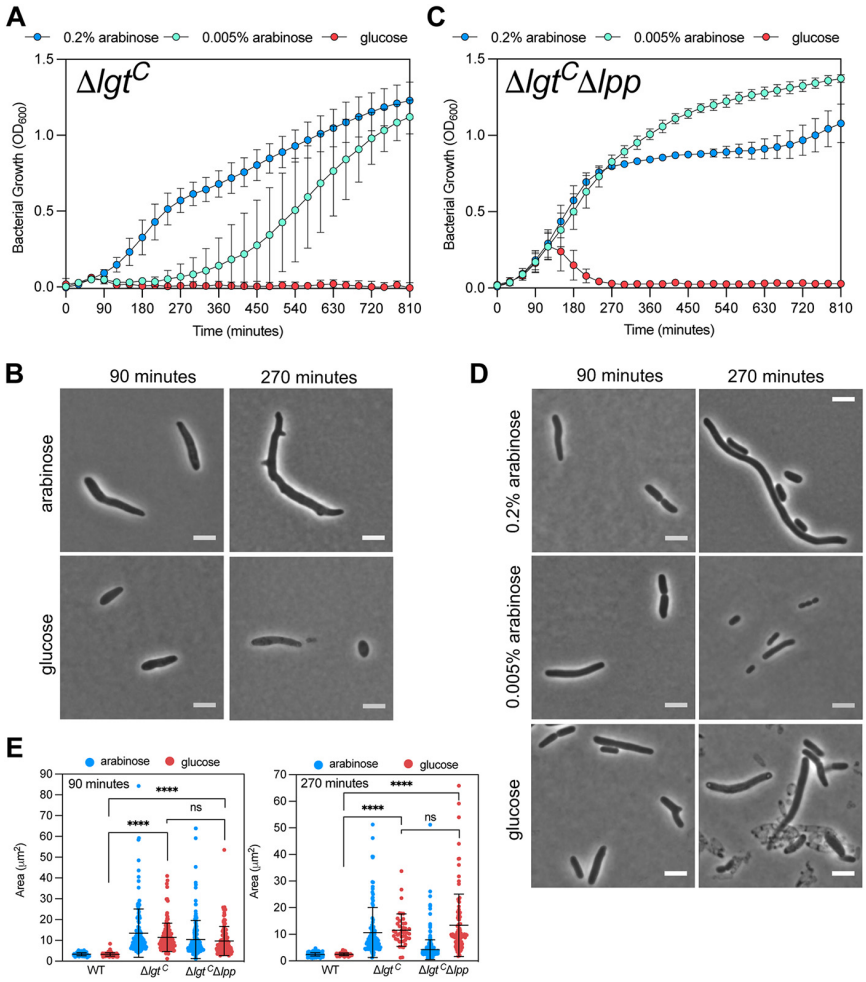


FIG 4 Deleting *lpp* does not rescue the distorted morphology caused by low Lgt levels. (A) Growth of chromosomal *lgt* depletion strain (MG1655Δ*lgt*) (Δ*lgt*^{C}) in 0.2% and 0.005% L-arabinose or 0.2% D-glucose. (B) Phase-contrast microscopy images of Δ*lgt*^{C}. (C) Growth of Δ*lgt*^{C} Δ*lpp* in high (0.2%) or low (0.005%) L-arabinose or in 0.2% D-glucose. (D) Phase-contrast microscopy images of Δ*lgt*^{C} Δ*lpp*. Bar, 5 μm. (E) Cell area measurements of cells ($n > 100$) from corresponding strains were performed on phase-contrast images using Microbe J. WT was MG1655.

suggest that basal levels of *lgt* expression from pLgt, alongside the absence of *lpp*, is sufficient to sustain growth.

Deletion of Lpp does not restore viability of the chromosomally encoded *lgt* depletion strain. As the basal level of expression of *lgt* from the multicopy pLgt may be a factor in the survival of the depletion strain, we sought to reduce the copy number of *lgt*. Diao et al. (29) recently showed that low levels of Lgt are sufficient to maintain cell viability. They described the construction of a MG1655 strain in which *lgt* was under the control of an arabinose promoter on the chromosome at the lambda attachment site (λatt; Δ*lgt*^{C}). When grown under the same conditions as Δ*lgt*^{P} in liquid media, Δ*lgt*^{C} was unable to grow in the presence of glucose and showed greater sensitivity to reduced arabinose concentrations (Fig. 4; see also Fig. S2). Interestingly, very low concentrations of L-arabinose (0.005%) were sufficient for growth, albeit after a dose-dependent lag period (see Fig. S2). Cells grown in the presence of 0.2% L-arabinose displayed distorted cell morphology with filamentous cells possessing multiple cell poles (Fig. 4). In the absence of Lpp, the depletion strain grew in D-glucose for 135 min before growth declined and was not further restored (Fig. 4). The cells grew as

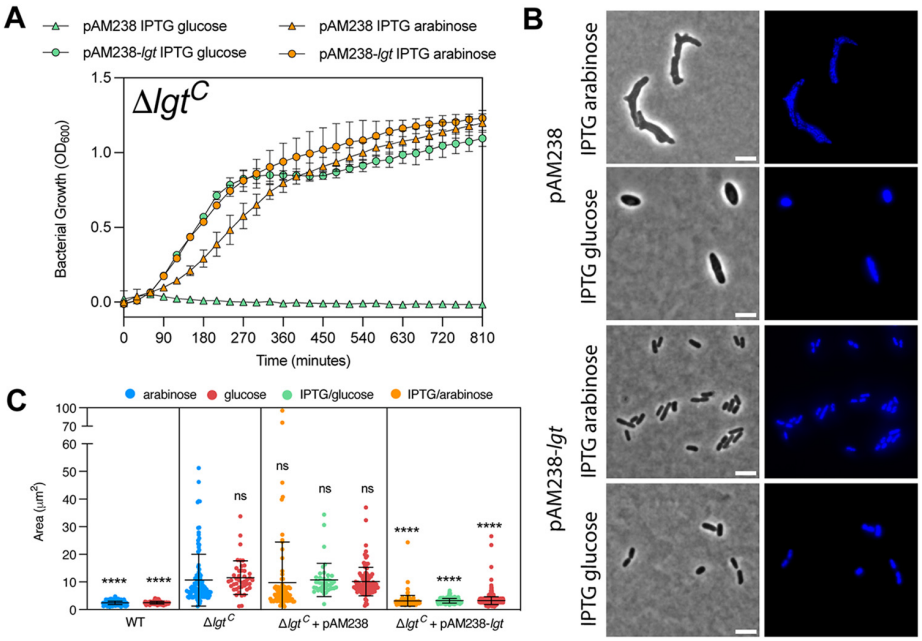


FIG 5 Complementation of MG1655Δ*lgt* by Lgt from pAM238-*lgt*-*flag*$_3$. (A) Growth of chromosomal *lgt* depletion strain (Δ*lgt*^{C}) containing empty plasmid pAM238 (triangle; SLEC70) or pAM238-*lgt*-*flag*$_3$ (round; SLEC71) in the presence of 0.2% L-arabinose with 5 mM IPTG (orange) and 0.2% D-glucose with 5 mM IPTG (green). (B) Phase-contrast images of SLEC71 and SLEC70 from 270 min. DNA was stained with Hoechst. Bar, 5 μm. (C) Cell area measurements of cells (n = 100) from corresponding strains at 270 min were performed on phase-contrast images using Microbe J. Statistical comparisons were in relation to Δ*lgt*^{C} in the presence of 0.2% L-arabinose. ****, $P < 0.0001$; ns, no significance.

filaments that lysed after 270 min. These results showed that Lpp is not the only cause of cell death in the absence of Lgt.

Complementation of Δ*lgt*^{C} is observed upon *lgt* expression in *trans* from a low-copy-number plasmid. Our results suggested that Lgt levels were too low to restore a wild-type morphology in Δ*lgt*^{C} grown in L-arabinose. Indeed, expression of *lgt* from an isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible promoter on a low-copy-number plasmid (pAM238-*lgt*-*flag*$_3$) restored cell size and morphology both in the presence of L-arabinose and D-glucose with clearly distinguishable nucleoids (Fig. 5). This indicated that only Lgt is responsible for the aberrant phenotypes observed under depletion conditions.

DISCUSSION

The correlation between lipoprotein modification and the cellular localization and function of the major lipoprotein Lpp in *E. coli* has been known for many years. Henry C. Wu and colleagues identified the *lgt* gene in *Salmonella enterica* serovar Typhimurium through analysis of a library of temperature-sensitive mutants and showed that deletion of *lpp* rescued the temperature-sensitive phenotype of an *lgt*(ts) mutant (5). The observation that Lgt is highly conserved in bacteria while Lpp is conserved in many but not all proteobacteria (30) raised the question of whether Lpp is important for the essentiality of Lgt in bacterial viability. Furthermore, since the lipoprotein modification pathway is unique to bacteria and the enzymes have domains exposed to the periplasm or exterior of the cell that seem readily accessible to small molecules, this pathway is potentially a good target for the development of novel antibiotics.

Recent phenotypic screenings have identified inhibitors targeting signal peptidase II (Lsp) (23) and downstream processes that involve transport of lipoproteins to the outer membrane (Lol) (21, 22). In all studies, absence of Lpp led to a decrease in efficacy of inhibition. The recent paper by Diao et al. demonstrated that inhibition of Lgt

by a synthetic cyclic peptide leads to morphology changes, membrane blebbing, and reduced viability (29). The effect of Lgt inhibition is, however, not reversed by *lpp* deletion. Cell size increase is modest in the presence of inhibitor compared to the Lgt depletion strain grown under restrictive conditions. These findings raise the question of whether the cyclic peptide directly targets Lgt or other off-target cell envelope components; however, inhibiting enzyme activity is different from lacking the enzyme, which probably leads to a more drastic phenotype. We took advantage of two Lgt depletion strains to study the role of Lpp on cell morphology and viability when cells were depleted for Lgt. In both strains, *lgt* expression is under the control of the arabinose-inducible P_{ara} promoter but differs in copy number of the *lgt* gene, i.e., one strain depends on a multicopy plasmid to control *lgt* expression, while the other strain contains a single chromosomal *lgt* gene. Deletion of endogenous *lgt* was performed similarly in both strains and did not affect downstream expression of *thy*A (25, 29). Depletion of Lgt led to a distorted morphology, leakage of DNA, cell lysis, and loss of viability. Wu and colleagues observed swollen, oval, and lysed cells in the *lgt*(ts) strain of *Salmonella* when grown at nonpermissive temperature (5). Strikingly, the morphology of the chromosomally controlled Lgt depletion strain was affected when grown under permissive conditions, in the presence of L-arabinose, and cells displayed branched cell poles. This phenotype was also observed, although more severely, in a mutant of *E. coli* lacking *dac*A, which encodes D,D-carboxypeptidase penicillin binding protein 5 (PBP5) (31). A similar *lgt* conditional null mutant in a uropathogenic *E. coli* strain led to retraction of the cytoplasmic membrane, membrane blebbing, and permeabilization of the outer membrane upon Lgt depletion, which was different from results with the K-12 wild-type strain used in that study, where DNA leakage and cell lysis were clearly observed (29). The genome has multiple pathogenicity islands inserted into its genome that may affect nucleoid organization and gene expression in general (32). Prolonged growth under restrictive conditions led to a revertant phenotype in the plasmid-controlled Lgt depletion strain. The survival was likely due to arabinose-independent low expression of *lgt*, since chromosomal mutations were not observed and the plasmid could not be lost from the strain, although the presence of an additional factor leading to growth restoration could not be excluded. Diao et al. showed a greater loss in viability of Δ*lgt*^{C} after 5 h of growth in the absence of Lpp when low concentrations of arabinose were present. After a short period of depletion, we observed no change in viability between Δ*lgt*^{C} and Δ*lgt*^{C} Δ*lpp* (see Fig. S3 in the supplemental material), but we did observe improved growth in liquid medium for the depletion strain when Lpp was removed and Lgt was expressed in low concentrations of arabinose (Fig. 4). In each instance, we showed that deletion of Lpp did not rescue growth of the chromosomally controlled Lgt depletion strain when Lgt was not expressed and that Lgt was therefore essential in *E. coli* independently of Lpp. Altogether, our findings show that other lipoproteins are important for cell viability in *E. coli*. Detailed roles have been assigned to the essential lipoproteins LolB, BamD, and LptE. LolB serves as a lipoprotein receptor in the outer membrane and is part of the Lol machinery (11). BamD is one of the four lipoproteins that constitute the Bam machinery involved in assembly of β-barrel proteins in the outer membrane (14, 15). LptE is important for correct assembly of LPS on the cell surface (12, 13). Lpp is the most abundant protein in the cell envelope of *E. coli*. In the absence of Lpp, substrate competition for these essential lipoproteins by Lgt is thus reduced, leading to fitness upon Lgt depletion, as previously suggested by Wu and colleagues (33). Over recent years, functional insight has been obtained about other lipoproteins involved in cell envelope biogenesis, including cell division, outer membrane biosynthesis, cell wall (peptidoglycan) biosynthesis, and envelope integrity; however, these lipoproteins do not have clearly distinguished roles in each of these processes (Fig. 6). For example, Lpp, the first outer membrane lipoprotein identified (34) that covalently links the outer membrane to the cell wall (35), regulates mechanical properties of the *E. coli* cell envelope (36), determines the space between the outer membrane and peptidoglycan (37), and has a

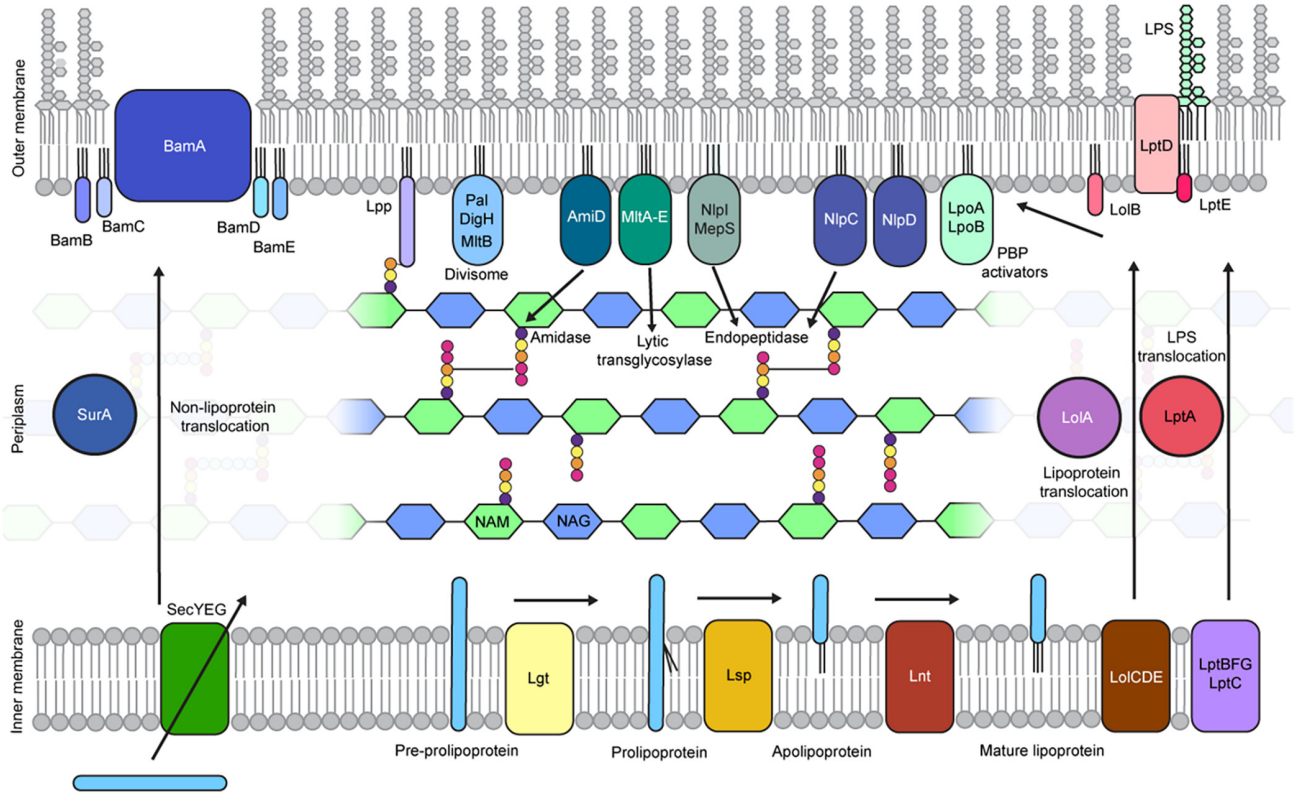


FIG 6 Essential cell envelope biogenesis processes that depend on correct outer membrane localization of lipoproteins. About 20 lipoproteins are involved in essential envelope biogenesis processes in *E. coli*, including peptidoglycan biogenesis and cell division. Three lipoproteins are essential, i.e., insertion of lipoproteins in the outer membrane is dependent on the Lol machinery (LolB), insertion of outer membrane β-barrel proteins (OMPs) is dependent on the Bam machinery (BamD), and assembly of LPS depends on the Lpt system (LptE). The implications for lipoproteins in cell wall biogenesis and cell division are discussed in the main text.

role in positioning large envelope-spanning complexes (38). A recent report described the role of Lgt in the correct function of DcrB, a lipoprotein in the cytoplasmic membrane which, together with small integral membrane protein YciB, functions in maintenance of membrane fluidity (39). Cell division requires the coordination of cell wall synthesis and membrane invagination. Two outer membrane lipoproteins, LpoA and LpoB, stimulate enzymatic activity of PBP1A and PBP1B, respectively, and their combined loss is lethal (40–42). In fact, when LpoB is localized to the cytoplasmic membrane, rod shape is lost (43). Amidases cleave the bond between the *N*-acetylmuramic acid sugar moiety and L-alanine of the peptide stem in peptidoglycan. AmiC is activated through interaction with lipoprotein NlpD and splits the peptidoglycan between daughter cells (44). Lipoprotein AmiD is an amidase with broad specificity, cleaving both muropeptides and intact peptidoglycan (45). NlpI is a general adaptor protein for peptidoglycan hydrolases (46) and interacts with, among other lipoproteins, MepS, an outer membrane lipoprotein with D,D-endopeptidase activity (47). Although not essential by itself, the lipoprotein Pal plays an important role as part of the Tol-Pal complex in maintaining cell envelope integrity, targeting of cell division components to midcell, and outer membrane constriction (48–51). Several lytic transglycosylases (MltA to -E) are lipoproteins involved in peptidoglycan biogenesis (52–56). MltB and DigH, a novel glycosyl hydrolase (57), are directly linked to the lipoprotein Pal, which noncovalently interacts with peptidoglycan (58) upon its dissociation from TolB (49). Thus, lipoproteins have a central function in cell wall biogenesis. Slowly, evidence has been obtained on how the different complexes operate and are coordinated. Finally, two major envelope stress response systems in *E. coli* depend on outer membrane lipoproteins: the Rsc system on RcsF (59) and the Cpx system on NlpE (60). Mislocalization of

RscF and NlpE leads to activation of downstream two-component systems and expression of repair genes (61). In conclusion, the increase in knowledge on lipoproteins with a function in cell envelope biogenesis shows that this very diverse class of fatty acid acylated proteins are important for bacterial physiology. Thus, removal or reduction of Lgt and therefore reduced lipoprotein processing have a multifaceted effect on the cell envelope. As the first enzyme of the posttranslational modification pathway, Lgt is essentially independent of the major lipoprotein Lpp and therefore a promising target for the development of novel antibiotics.

MATERIALS AND METHODS

Bacterial growth conditions. Bacterial cells were grown in Luria-Bertani broth (LB) or agar (LBA) at 37°C. When required, kanamycin (Kan, at 40 mg/liter), chloramphenicol (Cm, at 30 mg/liter), spectinomycin (Spec, at 50 mg/liter), or tetracycline (Tet, at 10 mg/liter) was added. A complete list of strains used in this study can be found in Table S1 in the supplemental material. Both Δ*lgt*^{P} and Δ*lgt*^{C} strains were grown in the presence of 0.2% L-arabinose unless stated otherwise. Overnight and initial growth were conducted as follows: single colonies of the required strains were selected and grown overnight in LB. Overnight cultures were washed three times in fresh LB and diluted 1/100 in LB without L-arabinose or IPTG to deplete Lgt.

Generation of *lpp* mutants. *lpp* deletion mutants were created by P1 transduction of *lpp*::Tn*10* from PAP8505 (19). The donor strain was grown in LB with 0.2% L-arabinose until an OD_{600} of 0.1 was reached. Different dilutions of P1 phage stock were added, and strains were grown until cell lysis was observed. Cleared P1 lysate from the lowest concentration of phage was used for transduction. Recipient strains were grown in liquid medium containing 0.2% L-arabinose overnight. Cells (100 μL) were supplemented with 10 mM $CaCl_2$ and dilutions of P1 phage were added and incubated without shaking at 37°C for 30 min. Sodium citrate was added to a final concentration of 150 mM to chelate Ca^{2+}. LB medium (1 mL) was added, and cells were grown at 37°C for 1 h while shaking. Cells were plated on Tet LBA containing Na-citrate to select for *lpp*::Tn*10* transductants. Colonies were restreaked once on Na-citrate Tet plates to rid P1 phage.

Construction of plasmid pAM238-*lgt-flag*$_3$. The upperFLAG and lowerFLAG primers (Proligo) were annealed and inserted into plasmid pCHAP9246 digested with restriction enzymes XbaI and HindIII to create pAM238-*lgt-flag*$_3$.

Growth kinetics. For growth kinetics, overnight and initial growth were conducted as described above for bacterial growth conditions. Cultures were diluted to an approximate final OD_{600} of 0.05 in 96-well plates with L-arabinose (0.2% and 0.005%), D-glucose (0.2%), and/or IPTG (5 mM). The plates were incubated at 37°C, and the OD_{600} was measured in a Tecan plate reader. Each condition was conducted with three biological replicates with at least two technical replicates.

Spot dilution assay. For spot dilution assays, overnight and initial growth were conducted as described above (Bacterial growth conditions). Cultures were diluted to a final OD_{600} of approximately 0.1. The adjusted cultures were serially diluted 10-fold in LB, and 5 μL was spotted onto LB agar plates containing decreasing concentrations of L-arabinose (0.2%, 0.05%, 0.0125%, 0.0031%, 0.008%, 0.004%, 0.002%), D-glucose (0.2%), or nonsupplemented LBA. Where required, LBA-Cm was used. Plates were incubated overnight at 37°C and imaged, and CFU counts per milliliter were recorded.

Microscopy. Overnight and initial growth were conducted in batch cultures. Cultures were diluted to a final OD_{600} of approximately 0.05. Cultures were grown for a further 270 min at 37°C with 0.2% L-arabinose, 0.005% L-arabinose, 0.2% D-glucose, or 5 mM IPTG (or a combination of conditions). Samples were taken at 90 min and 270 min, and the cells were fixed in 2.8% paraformaldehyde, 0.04% glutaraldehyde and incubated at room temperature (RT) for 15 min. Fixed cells were washed 3 times by centrifugation at 13,200 × *g* and resuspended in phosphate-buffered saline (PBS). For fluorescent labeling, 10 mg/liter Hoechst 33342 and 0.2 mg/liter FM 4-64FX (Invitrogen) were added, and mixtures were incubated for 15 min at RT in the dark. Cells were washed 3 times in PBS. Microscope slides were prepared with agarose pads (1% agarose in H_2O), and 5 μL of cells was added. Imaging was performed on a Zeiss Axio Observer microscope, and images were analyzed with the ImageJ plugin MicrobeJ. Final images were adapted for publication with ImageJ and Adobe Illustrator.

Collection of revertants. Cultures from growth kinetics of Δ*lgt*^{P} grown in 0.2% D-glucose were restreaked on LBA-Cm and incubated overnight at 37°C. A single colony from each of the nine samples collected was selected and stored (Δ*lgt*^{P} m1 to Δ*lgt*^{P} m9).

Whole-genome sequencing. DNA samples were fragmented with a Covaris M220 focused ultrasonicator (Covaris Ltd., Brighton, United Kingdom) using microTUBE-15 to 350 bp. The TruSeq DNA PCR-free libraries prep kit (Illumina) was used, following the instructions of the kit manufacturer. Sequencing was carried out on an Illumina NextSeq 500 platform. Raw reads were processed with an in-house bioinformatics pipeline for quality (62). The genome of the strain *Escherichia coli* Δ*lgt*^{P} was assembled using SPAdes (63). An automatic annotation was completed with Prokka (64) for the strain *E. coli* BW25113 (accession number CP009273.1) as a reference. Then, we used the variant calling pipeline (v0.11.0), which is available online (https://github.com/sequana/variant_calling). Analysis of variants was further performed using Geneious Prime.

Glycine-SDS and Tricine-SDS-PAGE and immunoblotting. For the detection of Lgt-c-myc$_2$ from pLgt and pLgt^{m1} to pLgt^{m3}, plasmids were extracted from Δ*lgt*^{Pm1} to Δ*lgt*^{Pm3} using a Qiagen Miniprep kit,

transformed into chemically competent BW25113 cells, and grown overnight on LBA-Cm. Single colonies were selected and grown overnight in LB-Cm. Overnight cultures were diluted 1/100 in LB-Cm and supplemented with L-arabinose (0.2%, 0.005%), or D-glucose (0.2%). Cultures were grown for a further 3 h at 37°C. Whole-cell pellets were taken and resuspended in sample buffer (10% glycerol, 2.5% SDS, 100 mM Tris [pH 8], 10 mM dithiothreitol, phenol red) equivalent to 0.01 OD_{600} units/μL before heating at 100°C for 10 min. Samples were loaded onto 4 to 15% Mini-Protean TGS stain-free precast protein gels (Bio-Rad), and proteins were separated by migration at a constant amperage (20 mA). Proteins were transferred to nitrocellulose membranes by using a Bio-Rad TurboBlot. Membranes were briefly stained in Ponceau S before washing in H_2O. Membranes were then blocked overnight in PBS–5% milk. Primary THE c-Myc tag monoclonal antibodies (Genscript) were diluted 1:10,000 in PBS–1% bovine serum albumin–0.05% Tween 20 and incubated for 1 h at RT. The blots were washed in PBS–0.1% Tween 20 twice for 5 min and three times for 10 min. Secondary anti-mouse-horseradish peroxidase (HRP) antibodies were diluted 1:10,000 in PBS–5% milk, and blots were incubated for 1 h. The blots were washed as described above. SuperSignal West Femto maximum sensitivity substrate was used to detect c-myc_2-tagged proteins when imaged under chemiluminescence and colorimetric conditions on a ChemiDoc imager (Bio-Rad).

For the detection of Lpp, cultures were diluted to a final OD_{600} of approximately 0.1. Cultures were grown for a further 270 min at 37°C in LB with 0.2% L-arabinose, 0.2% D-glucose, or nonsupplemented. Proteins from 1-mL cell cultures were precipitated in 10% ice-cold trichloroacetic acid and incubated on ice for 30 min. Precipitated samples were centrifuged at 13,200 $\times$ *g* for 5 min and washed twice in ice-cold (−20°C) acetone. After centrifugation, the pellet was air dried and resuspended in Tricine sample buffer (Bio-Rad) to an equivalent of 0.01 OD_{600} units/μL before heating at 100°C for 10 min. Samples were loaded onto Invitrogen Novex 16% Tricine protein gels and migrated at constant voltage (100 V) for 4 h. Western blots were blocked in Tris-buffered saline (TBS)–5% milk–0.1% Tween 20, primary anti-Lpp antibodies were diluted 1:20,000 in TBS–1% milk, and secondary anti-rabbit-HRP was diluted 1:10,000 in TBS–0.1% Tween 20. Detection was conducted as describe above.

pFREE plasmid curing. pFREE plasmid curing was performed as previously described (27). Briefly, overnight cultures of BW25113 pLgt, Δ*lgt*^{P} Δ*lpp*, or Δ*lgt* revertants were diluted 1/100 in LB and grown for 1 h at 37°C. Cultures were then made chemically competent and 50 ng of pFREE was transformed. After 2-h incubation in LB at 30°C, the cultures were induced by diluting 1/20 in LB containing 0.2% rhamnose, 200 ng/μL anhydrotetracycline, 50 mg/liter Kan, and further incubated overnight at 37°C. Uninduced strains were used as negative controls. The next day, cultures were plated onto LB and LB-Cm and the curing efficiency was calculated by determining the ratios in colony numbers. When colonies were susceptible to Cm, presence or absence of pLgt was confirmed by PCR with pBAD_F and pBAD_R primers (see Table S1).

Statistical analysis and data presentation. All numerical data were analyzed with GraphPad Prism (version 9.3.0). Microscopy images presented here were prepared with ImageJ and Adobe Illustrator. For comparisons of cell areas, distribution was determined by a Shapiro-Wilk test, and one-way analysis of variance with multiple comparisons by the Kruskal-Wallis test was employed.

Data availability. WGS data have been submitted to the NCBI SRA database under accession number PRJNA860548.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 1.3 MB.

ACKNOWLEDGMENTS

This work was financed by Institut Pasteur and the Pasteur Paris University (PPU)-Oxford International PhD Program with a fellowship to S.L.G. Our research received funding from the Institut Carnot Infectious Diseases Global Care (16 CARN 0023-01 Project iLiNT) and from the French Government's Investissement d'Avenir program, Laboratoire d'Excellence "Integrative Biology of Emerging Infectious Diseases" (grant ANR-10-LABX-62-IBEID). We thank G. H. Haustant, L. Lemée, and M. Monot from the Biomics Platform, C2RT, Institut Pasteur, Paris, France, who are supported by France Génomique (ANR-10-INBS-09) and IBISA. We thank Sharookh Kapadia for strain MG1655Δ*lgt* and Hajime Tokuda for Lpp antibodies. We are grateful for constructive comments and suggestions from all BGPB lab members.

REFERENCES

1. Buddelmeijer N. 2015. The molecular mechanism of bacterial lipoprotein modification: how, when and why? FEMS Microbiol Rev 39:246–261. https://doi.org/10.1093/femsre/fuu006.
2. Legood S, Boneca IG, Buddelmeijer N. 2021. Mode of action of lipoprotein modification enzymes: novel antibacterial targets. Mol Microbiol 115:356–365. https://doi.org/10.1111/mmi.14610.
3. Nguyen MT, Matsuo M, Niemann S, Herrmann M, Gotz F. 2020. Lipoproteins in Gram-positive bacteria: abundance, function, fitness. Front Microbiol 11:582582. https://doi.org/10.3389/fmicb.2020.582582.
4. Kovacs-Simon A, Titball RW, Michell SL. 2011. Lipoproteins of bacterial pathogens. Infect Immun 79:548–561. https://doi.org/10.1128/IAI.00682-10.

5. Gan K, Gupta SD, Sankaran K, Schmid MB, Wu HC. 1993. Isolation and characterization of a temperature-sensitive mutant of *Salmonella typhimurium* defective in prolipoprotein modification. J Biol Chem 268:16544–16550. https://doi.org/10.1016/S0021-9258(19)85453-4.
6. Tokunaga M, Loranger JM, Wu HC. 1984. Prolipoprotein modification and processing enzymes in *Escherichia coli*. J Biol Chem 259:3825–3830. https://doi.org/10.1016/S0021-9258(17)43170-X.
7. Yamagata H, Daishima K, Mizushima S. 1983. Cloning and expression of a gene coding for the prolipoprotein signal peptidase of Escherichia coli. FEBS Lett 158:301–304. https://doi.org/10.1016/0014-5793(83)80600-0.
8. Jackowski S, Rock CO. 1986. Transfer of fatty acids from the 1-position of phosphatidylethanolamine to the major outer membrane lipoprotein of *Escherichia coli*. J Biol Chem 261:11328–11333. https://doi.org/10.1016/S0021-9258(18)67387-9.
9. Hillmann F, Argentini M, Buddelmeijer N. 2011. Kinetics and phospholipid specificity of apolipoprotein N-acyltransferase. J Biol Chem 286:27936–27946. https://doi.org/10.1074/jbc.M111.243519.
10. Tokuda H, Matsuyama S, Tanaka-Masuda K. 2007. Structure, function and transport of lipoproteins in *Escherichia coli*, p 67–79. *In* Ehrmann M (ed), The periplasm. ASM Press, Washington, DC.
11. Matsuyama S, Yokota N, Tokuda H. 1997. A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of *Escherichia coli*. EMBO J 16:6947–6955. https://doi.org/10.1093/emboj/16.23.6947.
12. Chng SS, Ruiz N, Chimalakonda G, Silhavy TJ, Kahne D. 2010. Characterization of the two-protein complex in Escherichia coli responsible for lipopolysaccharide assembly at the outer membrane. Proc Natl Acad Sci U S A 107:5363–5368. https://doi.org/10.1073/pnas.0912872107.
13. Wu T, McCandlish AC, Gronenberg LS, Chng S, Silhavy TJ, Kahne D. 2006. Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of Escherichia coli. Proc Natl Acad Sci U S A 103:11754–11759. https://doi.org/10.1073/pnas.0604744103.
14. Wu T, Malinverni J, Ruiz N, Kim S, Silhavy TJ, Kahne D. 2005. Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. Cell 121:235–245. https://doi.org/10.1016/j.cell.2005.02.015.
15. Malinverni JC, Werner J, Kim S, Sklar JG, Kahne D, Misra R, Silhavy TJ. 2006. YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. Mol Microbiol 61:151–164. https://doi.org/10.1111/j.1365-2958.2006.05211.x.
16. Braun V, Sieglin U. 1970. The covalent murein-lipoprotein structure of the *Escherichia coli* cell wall. The attachment site of the lipoprotein on the murein. Eur J Biochem 13:336–346. https://doi.org/10.1111/j.1432-1033.1970.tb00936.x.
17. Cowles CE, Li Y, Semmelhack MF, Cristea IM, Silhavy TJ. 2011. The free and bound forms of Lpp occupy distinct subcellular locations in Escherichia coli. Mol Microbiol 79:1168–1181. https://doi.org/10.1111/j.1365-2958.2011.07539.x.
18. Yakushi T, Tajima T, Matsuyama S-I, Tokuda H. 1997. Lethality of the covalent linkage between mislocalized major outer membrane lipoprotein and the peptidoglycan of *Escherichia coli*. J Bacteriol 179:2857–2862. https://doi.org/10.1128/jb.179.9.2857-2862.1997.
19. Robichon C, Vidal-Ingigliardi D, Pugsley AP. 2005. Depletion of apolipoprotein N-acyltransferase causes mislocalization of outer membrane lipoproteins in *Escherichia coli*. J Biol Chem 280:974–983. https://doi.org/10.1074/jbc.M411059200.
20. Lin JJ, Kanazawa H, Ozols J, Wu HC. 1978. An *Escherichia coli* mutant with an amino acid alteration within the signal sequence of outer membrane prolipoprotein. Proc Natl Acad Sci U S A 75:489–495.
21. Nickerson NN, Jao CC, Xu Y, Quinn J, Skippington E, Alexander MK, Miu A, Skelton N, Hankins JV, Lopez MS, Koth CM, Rutherford S, Nishiyama M. 2018. A novel inhibitor of the LolCDE ABC transporter essential for lipoprotein trafficking in Gram-negative bacteria. Antimicrob Agents Chemother 62. https://doi.org/10.1128/AAC.02151-17.
22. McLeod SM, Fleming PR, MacCormack K, McLaughlin RE, Whiteaker JD, Narita S, Mori M, Tokuda H, Miller AA. 2015. Small-molecule inhibitors of gram-negative lipoprotein trafficking discovered by phenotypic screening. J Bacteriol 197:1075–1082. https://doi.org/10.1128/JB.02352-14.
23. Pantua H, Skippington E, Braun MG, Noland CL, Diao J, Peng Y, Gloor SL, Yan D, Kang J, Katakam AK, Reeder J, Castanedo GM, Garland K, Komuves L, Sagolla M, Austin CD, Murray J, Xu Y, Modrusan Z, Xu M, Hanan EJ, Kapadia SB. 2020. Unstable mechanisms of resistance to inhibitors of Escherichia coli lipoprotein signal peptidase. mBio 11. https://doi.org/10.1128/mBio.02018-20.
24. Wu HC, Hou C, Lin JJ, Yem DW. 1977. Biochemical characterization of a mutant lipoprotein of Escherichia coli. Proc Natl Acad Sci U S A 74:1388–1392. https://doi.org/10.1073/pnas.74.4.1388.
25. Pailler J, Aucher W, Pires M, Buddelmeijer N. 2012. Phosphatidylglycerol::prolipoprotein diacylglyceryl transferase (Lgt) of Escherichia coli has seven transmembrane segments, and its essential residues are embedded in the membrane. J Bacteriol 194:2142–2151. https://doi.org/10.1128/JB.06641-11.
26. Guo MS, Updegrove TB, Gogol EB, Shabalina SA, Gross CA, Storz G. 2014. MicL, a new sigmaE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. Genes Dev 28:1620–1634. https://doi.org/10.1101/gad.243485.114.
27. Lauritsen I, Porse A, Sommer MOA, Norholm MHH. 2017. A versatile one-step CRISPR-Cas9 based approach to plasmid-curing. Microb Cell Fact 16:135. https://doi.org/10.1186/s12934-017-0748-z.
28. Glauner B, Holtje JV, Schwarz U. 1988. The composition of the murein of Escherichia coli. J Biol Chem 263:10088–10095. https://doi.org/10.1016/S0021-9258(19)81481-3.
29. Diao J, Komura R, Sano T, Pantua H, Storek KM, Inaba H, Ogawa H, Noland CL, Peng Y, Gloor SL, Yan D, Kang J, Katakam AK, Volny M, Liu P, Nickerson NN, Sandoval W, Austin CD, Murray J, Rutherford ST, Reichelt M, Xu Y, Xu M, Yanagida H, Nishikawa J, Reid PC, Cunningham CN, Kapadia SB. 2021. Inhibition of Escherichia coli lipoprotein diacylglyceryl transferase is insensitive to resistance caused by deletion of Braun's lipoprotein. J Bacteriol 203:e0014921. https://doi.org/10.1128/JB.00149-21.
30. Witwinowski J, Sartori-Rupp A, Taib N, Pende N, Tham TN, Poppleton D, Ghigo JM, Beloin C, Gribaldo S. 2022. An ancient divide in outer membrane tethering systems in bacteria suggests a mechanism for the diderm-to-monoderm transition. Nat Microbiol 7:411–422. https://doi.org/10.1038/s41564-022-01066-3.
31. Nelson DE, Young KD. 2001. Contributions of PBP 5 and DD-carboxypeptidase penicillin binding proteins to maintenance of cell shape in Escherichia coli. J Bacteriol 183:3055–3064. https://doi.org/10.1128/JB.183.10.3055-3064.2001.
32. Lloyd AL, Rasko DA, Mobley HL. 2007. Defining genomic islands and uropathogen-specific genes in uropathogenic Escherichia coli. J Bacteriol 189:3532–3546. https://doi.org/10.1128/JB.01744-06.
33. Gan K, Sankaran K, Williams MG, Aldea M, Rudd KE, Kushner SR, Wu HC. 1995. The *umpA* gene of *Escherichia coli* encodes phosphatidylglycerol:prolipoprotein diacylglyceryl transferase (*lgt*) and regulates thymidylate synthase levels through translational coupling. J Bacteriol 177:1879–1882. https://doi.org/10.1128/jb.177.7.1879-1882.1995.
34. Braun V, Rehn K. 1969. Chemical characterization, spatial distribution and function of a lipoprotein (murein-lipoprotein) of the *E. coli* cell wall. Eur J Biochem 10:426–438. https://doi.org/10.1111/j.1432-1033.1969.tb00707.x.
35. Braun V, Bosch V. 1972. Sequence of the murein-lipoprotein and the attachment site of the lipid. Eur J Biochem 28:51–69. https://doi.org/10.1111/j.1432-1033.1972.tb01883.x.
36. Mathelie-Guinlet M, Asmar AT, Collet JF, Dufrene YF. 2020. Lipoprotein Lpp regulates the mechanical properties of the E. coli cell envelope. Nat Commun 11:1789. https://doi.org/10.1038/s41467-020-15489-1.
37. Asmar AT, Ferreira JL, Cohen EJ, Cho SH, Beeby M, Hughes KT, Collet JF. 2017. Communication across the bacterial cell envelope depends on the size of the periplasm. PLoS Biol 15:e2004303. https://doi.org/10.1371/journal.pbio.2004303.
38. Gumbart JC, Ferreira JL, Hwang H, Hazel AJ, Cooper CJ, Parks JM, Smith JC, Zgurskaya HI, Beeby M. 2021. Lpp positions peptidoglycan at the AcrA-TolC interface in the AcrAB-TolC multidrug efflux pump. Biophys J 120:3973–3982. https://doi.org/10.1016/j.bpj.2021.08.016.
39. Mychack A, Janakiraman A. 2021. Defects in the first step of lipoprotein maturation underlie the synthetic lethality of Escherichia coli lacking the inner membrane proteins YciB and DcrB. J Bacteriol 203. https://doi.org/10.1128/JB.00640-20.
40. Paradis-Bleau C, Markovski M, Uehara T, Lupoli TJ, Walker S, Kahne DE, Bernhardt TG. 2010. Lipoprotein cofactors located in the outer membrane activate bacterial cell wall polymerases. Cell 143:1110–1120. https://doi.org/10.1016/j.cell.2010.11.037.
41. Typas A, Banzhaf M, van den Berg van Saparoea B, Verheul J, Biboy J, Nichols R, Zietek M, Beilharz K, Kannenberg K, von Rechenberg M, Breukink E, den Blaauwen T, Gross CA, Vollmer W. 2010. Regulation of peptidoglycan synthesis by outer-membrane proteins. Cell 143:1097–1109. https://doi.org/10.1016/j.cell.2010.11.038.
42. Lupoli TJ, Lebar MD, Markovski M, Bernhardt T, Kahne D, Walker S. 2014. Lipoprotein activators stimulate Escherichia coli penicillin-binding proteins

by different mechanisms. J Am Chem Soc 136:52–55. https://doi.org/10.1021/ja410813j.
43. Ranjit DK, Jorgenson MA, Young KD. 2017. PBP1B glycosyltransferase and transpeptidase activities play different essential roles during the de novo regeneration of rod morphology in Escherichia coli. J Bacteriol 199. https://doi.org/10.1128/JB.00612-16.
44. Uehara T, Dinh T, Bernhardt TG. 2009. LytM-domain factors are required for daughter cell separation and rapid ampicillin-induced lysis in Escherichia coli. J Bacteriol 191:5094–5107. https://doi.org/10.1128/JB.00505-09.
45. Kerff F, Petrella S, Mercier F, Sauvage E, Herman R, Pennartz A, Zervosen A, Luxen A, Frere JM, Joris B, Charlier P. 2010. Specific structural features of the N-acetylmuramoyl-L-alanine amidase AmiD from Escherichia coli and mechanistic implications for enzymes of this family. J Mol Biol 397:249–259. https://doi.org/10.1016/j.jmb.2009.12.038.
46. Banzhaf M, Yau HC, Verheul J, Lodge A, Kritikos G, Mateus A, Cordier B, Hov AK, Stein F, Wartel M, Pazos M, Solovyova AS, Breukink E, van Teeffelen S, Savitski MM, den Blaauwen T, Typas A, Vollmer W. 2020. Outer membrane lipoprotein NlpI scaffolds peptidoglycan hydrolases within multi-enzyme complexes in Escherichia coli. EMBO J 39:e102246. https://doi.org/10.15252/embj.2019102246.
47. Singh SK, Parveen S, SaiSree L, Reddy M. 2015. Regulated proteolysis of a cross-link-specific peptidoglycan hydrolase contributes to bacterial morphogenesis. Proc Natl Acad Sci U S A 112:10956–10961. https://doi.org/10.1073/pnas.1507760112.
48. Gerding MA, Ogata Y, Pecora ND, Niki H, de Boer PA. 2007. The trans-envelope Tol-Pal complex is part of the cell division machinery and required for proper outer-membrane invagination during cell constriction in E. coli. Mol Microbiol 63:1008–1025. https://doi.org/10.1111/j.1365-2958.2006.05571.x.
49. Petiti M, Serrano B, Faure L, Lloubes R, Mignot T, Duche D. 2019. Tol energy-driven localization of Pal and anchoring to the peptidoglycan promote outer-membrane constriction. J Mol Biol 431:3275–3288. https://doi.org/10.1016/j.jmb.2019.05.039.
50. Szczepaniak J, Holmes P, Rajasekar K, Kaminska R, Samsudin F, Inns PG, Rassam P, Khalid S, Murray SM, Redfield C, Kleanthous C. 2020. The lipoprotein Pal stabilises the bacterial outer membrane during constriction by a mobilisation-and-capture mechanism. Nat Commun 11:1305. https://doi.org/10.1038/s41467-020-15083-5.
51. Hale CA, Persons L, de Boer PAJ. 2022. Recruitment of the TolA protein to cell constriction sites in Escherichia coli via three separate mechanisms, and a critical role for FtsWI activity in recruitment of both TolA and TolQ. J Bacteriol 204:e0046421. https://doi.org/10.1128/jb.00464-21.
52. Lommatzsch J, Templin MF, Kraft AR, Vollmer W, Höltje J-V. 1997. Outer membrane localization of murein hydrolases: MltA, a third lipoprotein lytic transglycosylase in *Escherichia coli*. J Bacteriol 179:5465–5470. https://doi.org/10.1128/jb.179.17.5465-5470.1997.
53. Dijkstra AJ, Keck W. 1996. Identification of new members of the lytic transglycosylase family in *Haemophilus influenzae* and *Escherichia coli*. Microb Drug Resist 2:141–145. https://doi.org/10.1089/mdr.1996.2.141.
54. Fibriansah G, Gliubich FI, Thunnissen AM. 2012. On the mechanism of peptidoglycan binding and cleavage by the endo-specific lytic transglycosylase MltE from Escherichia coli. Biochemistry 51:9164–9177. https://doi.org/10.1021/bi300900t.
55. Ehlert K, Holtje JV, Templin MF. 1995. Cloning and expression of a murein hydrolase lipoprotein from Escherichia coli. Mol Microbiol 16:761–768. https://doi.org/10.1111/j.1365-2958.1995.tb02437.x.
56. Engel H, Smink AJ, van Wijngaarden L, Keck W. 1992. Murein-metabolizing enzymes from Escherichia coli: existence of a second lytic transglycosylase. J Bacteriol 174:6394–6403. https://doi.org/10.1128/jb.174.20.6394-6403.1992.
57. Yakhnina AA, Bernhardt TG. 2020. The Tol-Pal system is required for peptidoglycan-cleaving enzymes to complete bacterial cell division. Proc Natl Acad Sci U S A 117:6777–6783. https://doi.org/10.1073/pnas.1919267117.
58. Godlewska R, Wiśniewska K, Pietras Z, Jagusztyn-Krynicka EK. 2009. Peptidoglycan-associated lipoprotein (Pal) of Gram-negative bacteria: function, structure, role in pathogenesis and potential application in immunoprophylaxis. FEMS Microbiol Lett 298:1–11. https://doi.org/10.1111/j.1574-6968.2009.01659.x.
59. Cho SH, Szewczyk J, Pesavento C, Zietek M, Banzhaf M, Roszczenko P, Asmar A, Laloux G, Hov AK, Leverrier P, Van der Henst C, Vertommen D, Typas A, Collet JF. 2014. Detecting envelope stress by monitoring beta-barrel assembly. Cell 159:1652–1664. https://doi.org/10.1016/j.cell.2014.11.045.
60. May KL, Lehman KM, Mitchell AM, Grabowicz M. 2019. A stress response monitoring lipoprotein trafficking to the outer membrane. mBio 10. https://doi.org/10.1128/mBio.00618-19.
61. Laloux G, Collet JF. 2017. Major Tom to ground control: how lipoproteins communicate extracytoplasmic stress to the decision center of the cell. J Bacteriol 199. https://doi.org/10.1128/JB.00216-17.
62. Cokelaer T, Desvillechabrol D, Legendre R, Cardon M. 2017. 'Sequana': a set of snakemake NGS pipelines. JOSS 2:352. https://doi.org/10.21105/joss.00352.
63. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes de novo assembler. Curr Protoc Bioinformatics 70:e102. https://doi.org/10.1002/cpbi.102.
64. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

Supplementary Figures and Tables

Supplementary Figures and Table

Fig. S1

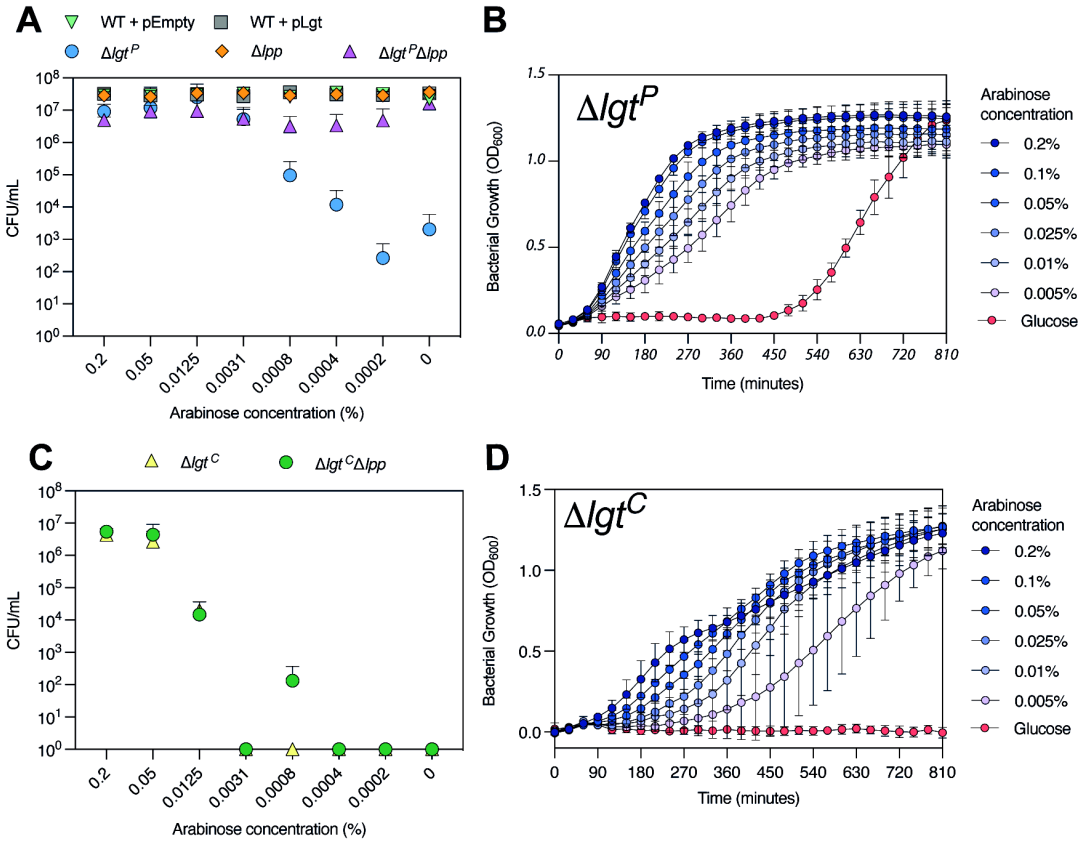


Fig. S1. Whole genome sequencing of Lgt revertant plasmids.

(A) Sequence of P_{ara} from pLgt^{m1-2} containing a single nucleotide polymorphism (SNP) G-A at residue 210. (B and C) Nucleotide and translated sequence of *myc*$_2$-tag region of pLgt^{m1-9}. Red box indicates stop codons, green box indicates deletions relative to pLgt.

Fig. S2.

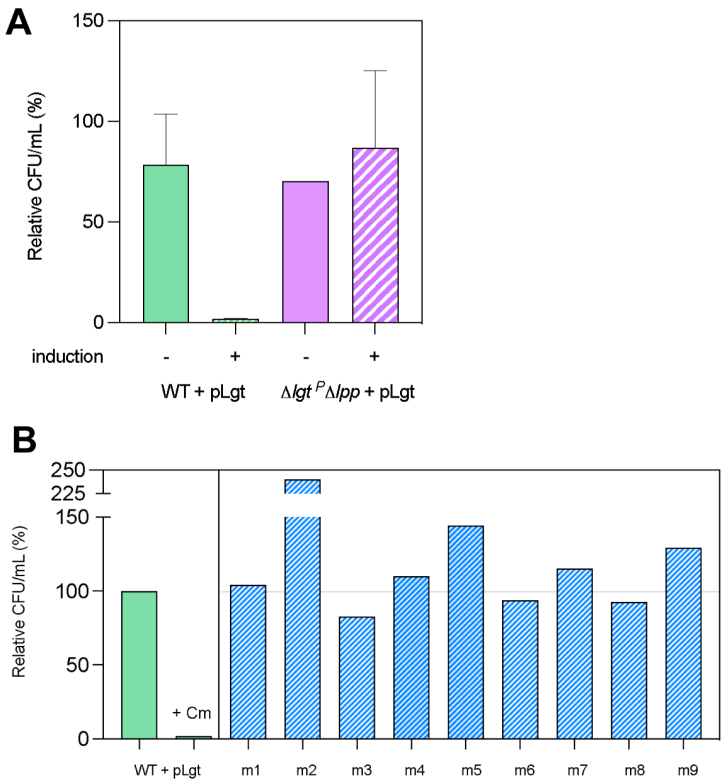


Fig. S2. Lgt complementing plasmid cannot be counter selected from *lgt* depletion strain lacking Lpp.

Cas9/gRNA encoding plasmid pFREE targeting the origin of replication of plasmids was transformed into wild type strain BW25113 containing (A) pCHAP9224 (pLgt) and Δ*lgt*^{P}Δ*lpp* or (B) Δ*lgt*^{P} revertants m1-9. Expression of *cas9*/*gRNA* was induced with 0.2% L-arabinose and clones are selected on non-selective plates and chloramphenicol (25 mg/L) containing plates. Colonies were counted in the two strains without and with induction of the Cas9/gRNA system. Loss of pLgt was expressed as ratio of number of colonies on non-selective and Cm plates. Where present, error bars indicate standard deviation from the mean at least two replicates, otherwise n = 1.

Fig. S3.

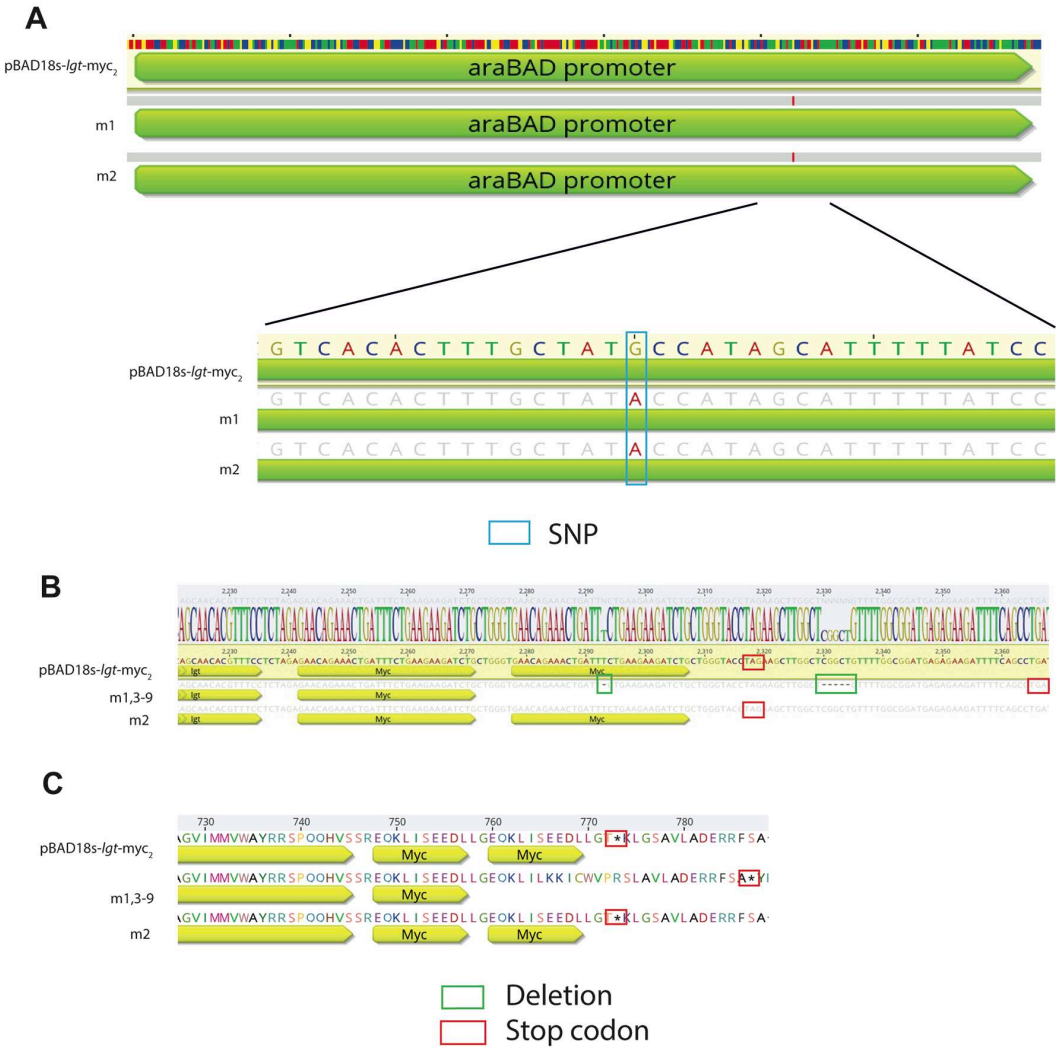


Fig. S3. Growth and viability of Lgt depletion strains.

(A) Colony forming units (CFU) of Δ*lgt*^{P} compared to BW25113 and Δ*lpp* derivative, and overexpression strain BW25113 containing pCHAP9224 versus empty pBAD18s-Cm; or (C) of chromosomal *lgt* depletion strain (Δ*lgt*^{C}) compared with Δ*lpp* derivative. (B) Growth kinetics of depletion strain Δ*lgt*^{P} ; or (D) of Δ*lgt*^{C} in varying L-arabinose concentrations and 0.2% D-glucose. Time t=0 corresponds to 2 hours of growth in LB medium without sugar. Graph represents duplicate OD_{600} measurements of biological triplicates (n=3).

Table S1. Strains, plasmids and primers used in this study.

Name	Alt Name	Description	Reference
Strains			
BW25113	SLEC30	*E. coli* K-12 *lacI*^{q} *rrnB*$_{T14}$ Δ*lacZ*$_{WJ16}$ *hsdR514* Δ*araBAD*$_{AH33}$ Δ*rhaBAD*$_{LD78}$	(1)
BW25113 Δ*lgt*^{P}	PAP9403	BW25113 Δ*lgt*::Kan^{r} + pBAD18s-Cm-*lgt*^{WT}-c-*myc*$_{2}$	(2)
BW25113 Δ*lgt*^{P}Δ*lpp*	SLEC44	BW25113 Δ*lgt*::Kan^{r} *lpp*::Tn10 + pBAD18s-Cm-*lgt*^{WT}-c-*myc*$_{2}$	This study
BW25113 pEmpty	SLEC45	BW25113 + pBAD18s-Cm	This study
BW25113 pLgt	SLEC22	BW25113 + pBAD18s-Cm-*lgt*^{WT}-c-*myc*$_{2}$	This study
PAP8505		BW25113 *ybeX*-(*kan-rpoCter-p*$_{araB}$)-*lnt* *lpp*::Tn10	(3)
BW25113 Δ*lpp*	SLEC65	BW25113 *lpp*::Tn10	This study
BW25113 Δ*lgt*^{P} m1-9		BW25113 Δ*lgt* revertants	This study
MG1655	SLEC69	F-, λ-, *ilv*G- *rfb-50 rph-1*	Lab collection
MG1655 Δ*lgt*^{C}	SLEC67	MG1655 Δ*lgt*::Kan^{r} λattB-pBAD-*lgt*	(4)
MG1655 Δ*lgt*^{C}Δ*lpp*	SLEC68	MG1655 Δ*lgt*::Kan^{r} λattB-pBAD-*lgt*Δ*lpp*::Tn10	This study
MG1655 Δ*lgt*^{C} pEmpty	SLEC70	MG1655 Δ*lgt*::Kan^{r} λattB-pBAD-*lgt* + pAM238	This study
MG1655 Δ*lgt*^{C} pLgt	SLEC71	MG1655 Δ*lgt*::Kan^{r} λattB-pBAD-*lgt* +	This study

		pAM238-*lgt-flag*$_3$	
Plasmids			
pBAD18s-Cm		pBR322 origin, P*ara* promoter (pBAD), Cm^r	(2)
pBAD18s-Cm-*lgt*^{WT}-c-*myc*$_2$	pCHAP9224, pLgt	pBR322 origin, P*ara* promoter (pBAD), Cm^r, expressing *E. coli lgt*	(2)
pLgt^{m1-9}		pBAD18s-Cm-*lgt*^{WT}-c-*myc*$_2$	This study
pAM238		pSC101 origin, Plac promoter, Spc^r	(5)
pAM238-*lgt-myc*$_2$	pCHAP9246	pSC101 origin, Plac promoter, Spc^r, expressing *E. coli lgt*	(2)
pAM238-*lgt-flag*$_3$	SLP14	pSC101 origin, Plac promoter, Spc^r, expressing *E. coli lgt*	This study
pFREE		gRNA and Cas9 expressing plasmid for plasmid curing	(6)
Primers	**Sequence (5' -> 3')**		**Reference**
upperFLAG	CTAGAGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGGTACCTAGA		Proligo
lowerFLAG	AGCTTCTAGGTACCATCGTCATCCTTGTAATCGATGTCATGATCTTTATAATCACCGTCATGGTCTTTGTAGTCT		Proligo
pBAD_F	AGATTAGCGGATCCTACCTG		Sigma
pBAD_R	CTCATCCGCCAAAACAG		Sigma

References

1. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Molecular Systems Biology 2.
2. Pailler J, Aucher W, Pires M, Buddelmeijer N. 2012. Phosphatidylglycerol::prolipoprotein diacylglyceryl transferase (Lgt) of Escherichia coli has seven transmembrane segments, and its essential residues are embedded in the membrane. J Bacteriol 194:2142-51.
3. Robichon C, Vidal-Ingigliardi D, Pugsley AP. 2005. Depletion of apolipoprotein N-acyltransferase causes mislocalization of outer membrane lipoproteins in *Escherichia coli*. J Biol Chem 280:974-83.
4. Diao J, Komura R, Sano T, Pantua H, Storek KM, Inaba H, Ogawa H, Noland CL, Peng Y, Gloor SL, Yan D, Kang J, Katakam AK, Volny M, Liu P, Nickerson NN, Sandoval W, Austin CD, Murray J, Rutherford ST, Reichelt M, Xu Y, Xu M, Yanagida H, Nishikawa J, Reid PC, Cunningham CN, Kapadia SB. 2021. Inhibition of Escherichia coli Lipoprotein Diacylglyceryl Transferase Is Insensitive to Resistance Caused by Deletion of Braun's Lipoprotein. J Bacteriol 203:e0014921.
5. Binet R, Wandersman C. 1995. Protein secretion by hybrid bacterial ABC-transporters: specific functions of the membrane ATPase and the membrane fusion protein. EMBO J 14:2298-2306.
6. Lauritsen I, Porse A, Sommer MOA, Norholm MHH. 2017. A versatile one-step CRISPR-Cas9 based approach to plasmid-curing. Microb Cell Fact 16:135.

Section IV: *In vitro* assay to study Lgt enzymatic activity

Summary

The aim of this chapter is to develop an *in vitro* enzymatic activity assay in order to study Lgt in greater depth and to screen for inhibitors.

To this end we successfully purified Lgt from *E. coli* using a two-step purification method of affinity and size exclusion chromatography. The purified enzyme was deemed pure with minimal contaminants.

To test the enzymatic activity of the enzyme we adapted a gel-shift assay whereby a reaction containing Lgt, a truncated Lpp peptide with a His_6-tag and biotin group and phosphatidylglycerol (PG) was incubated and analysed via Tricine-SDS-PAGE. A shift in migration indicated successful incorporation of diacylglycerol from PG onto the Lpp peptide. This assay confirmed our enzyme was active, but the gel-shift method is not adaptable for high throughput applications.

We therefore sought to see whether fluorescent phospholipids were substrates for Lgt but found poor modification under these conditions. A 'click-chemistry' based assay, adapted from an *in vitro* activity assay of Lnt was attempted. However, although the click-reaction was successful and we could detect the modified substrate by Tricine-PAGE, the assay development turned out difficult, likely due to the biotin-peptide substrate used for the reaction binding poorly to the streptavidin-coated plates.

Finally, we have explored the use of Mass Spectrometry in the study of Lgt enzymatic activity.

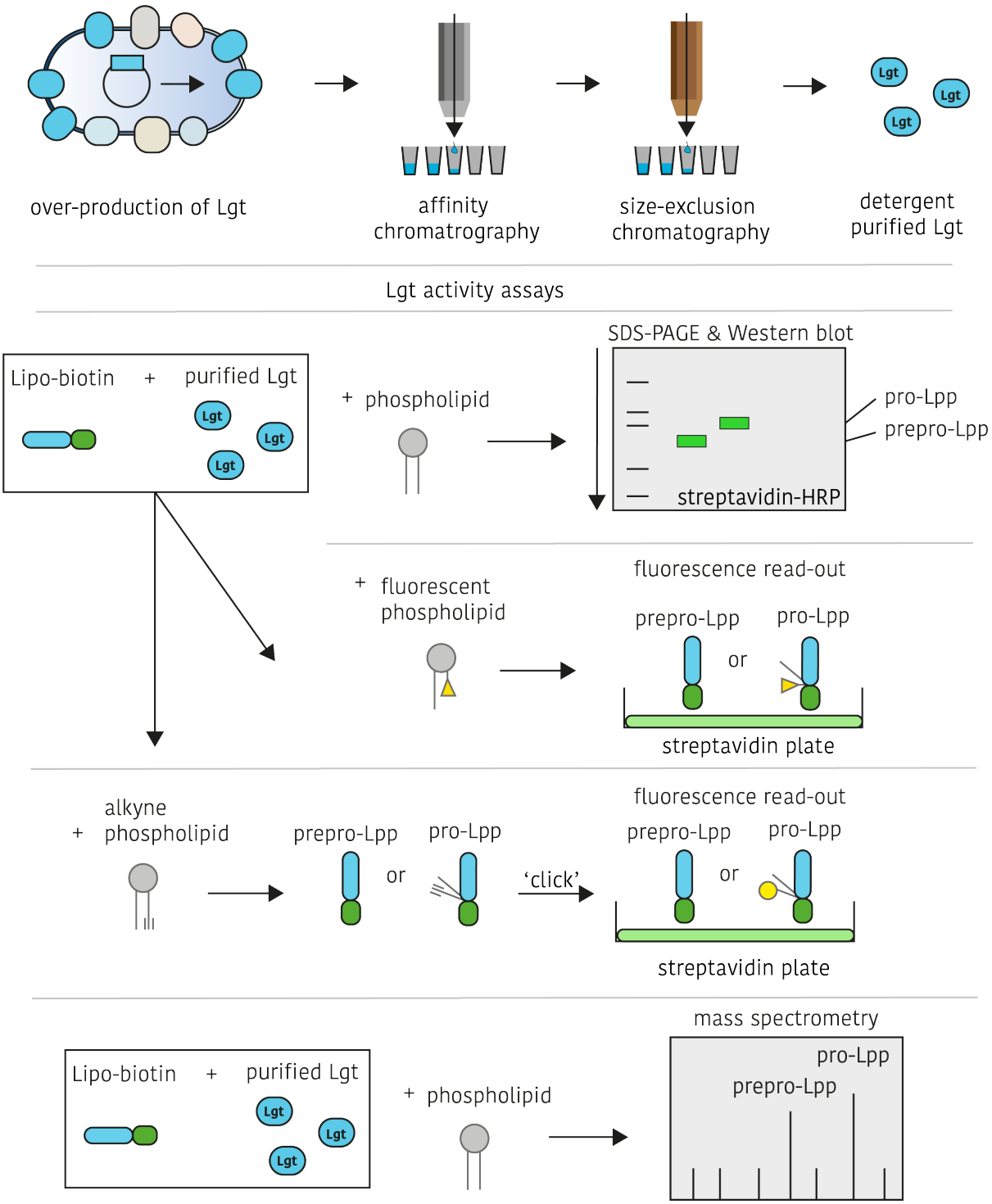


Figure 31. Section IV Graphical Abstract.

Results and discussion

Cloning, overproduction and solubilisation

The first step in developing an *in vitro* activity assay was to purify Lgt in an active form. Previous Lgt activity assays have used either inverted membrane vesicles (IMVs), single step ion-exchange purification or two-step affinity and size-exclusion detergent purification (Figure 11, Table 6)[75, 78, 81, 82, 93, 101, 117]. As the detergent purified Lgt was shown to be active and was the method used for the X-ray crystallography [81] we opted for this approach. IMVs do not provide a pure sample and contain native phospholipids. This is a disadvantage if analysis of phospholipid substrates is required. A second drawback to IMVs is the presence of multiple membrane proteins whose function and activity could interfere with the Lgt reaction or possible enzymes yet discovered with similar activity. The single-step ion-exchange method did produce active protein and is a viable alternative to the two-step purification method.

A pET28b overexpression vector template was used for expressing *lgt* from *E. coli.* The gene was cloned with a His_6 tag at the C-terminal end for affinity purification and detection by immunoblotting. The method for cloning *lgt* into the expression vector was similar to that described by Mao *et al.,* (2016) and Daley *et al.,* (2006). pET28b is a high-copy expression plasmid commonly used for protein overproduction in *E. coli* with the encoding gene under control of a T7 promoter. The plasmid has a kanamycin selection marker. Correct insertion of the gene into the vector was confirmed by sequencing (Figure 32).

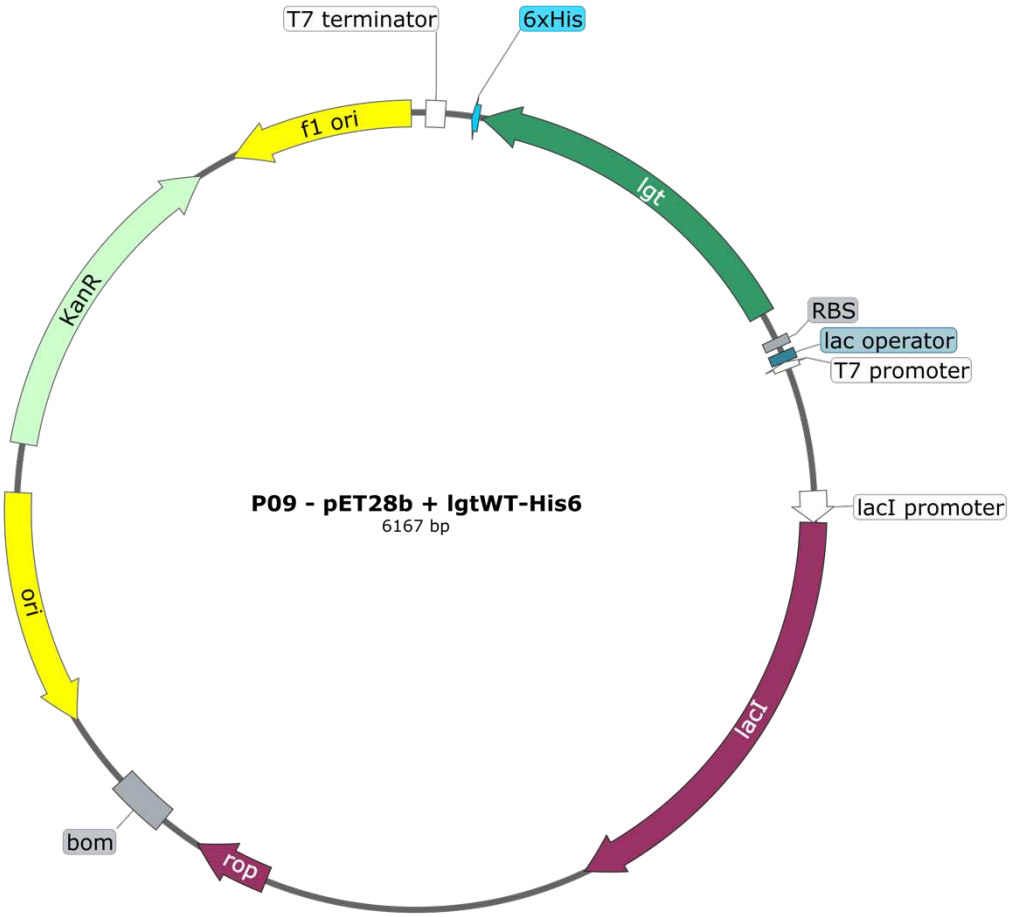


Figure 32. pET28 cloning vector for Lgt purification (P09).

pET238b-*lgt*-His_6 was transformed into chemically competent C43 cells for protein expression (SLEC20). The C43 strain is a derivative of *E. coli* BL21 (DE3) containing the T7 RNA polymerase gene under control of an IPTG-inducible P_{lac} promoter. Upon addition of IPTG, T7 RNA polymerase drives gene expression from plasmid pET28b, in this case *lgt-his*$_6$. To assess whether Lgt-*his*$_6$ was produced from P09 in C43 cells, the strain was grown in Terrific Broth in the presence or absence of IPTG (Figure 33). Lgt-His_6 (herein, Lgt) production is clearly visible upon induction with IPTG (Figure 33, Lane 1 & 2) and therefore a larger scale purification was conducted. Lgt is observed on α-His-HRP Western blot at around 26 kDa but has a predicted MW of 33 kDa. Lgt has been observed on acrylamide gel to run at 27-30 kDa [76] or 25 kDa [88]. Here, Lgt migrates within the bounds of these previous studies.

After growth and induction of *lgt* expression on a larger scale (2 L flask culture), IPTG induced production of Lgt was assessed and was in accordance with the preliminary assessment of Lgt production. After the cells were homogenized by high pressure in a CellD apparatus, an initial centrifugation step was conducted to remove unbroken cells and debris (Figure 33, Lane 3). The

supernatant was then cleared by ultracentrifugation to separate soluble proteins (Figure 33, Lane 4) from insoluble membrane proteins. Integral membrane proteins are not overproduced in large amounts, the detection by antibodies is often required to detect the protein. The vast majority of Lgt protein is in the membrane fraction and only a faint band is observed in the soluble protein fraction.

The membrane localisation of Lgt is in agreement with the hypothesis that it is an integral membrane protein and indeed it was first observed in the membrane fraction as opposed to the soluble fraction by Williams *et al.* (1989). It therefore raises questions as to the activity assays conducted by Sundaram *et al.* (2012) and more recently Sangith *et al.* (2019), who observe diacylglyceryl transfer activity in water soluble fractions. The possibility could be raised that only very low levels of Lgt are required for a reaction and therefore low quantities of the enzyme in the soluble fraction are sufficient to produce activity *in vitro*. Lgt purified by the same group but from the firmicute *L. lactis* required the detergent octylglucoside (OG) to solubilize the enzyme [78]. Our purification procedure includes buffers containing salt (300 mM) and detergent. A general procedure to exclude peripheral association of membrane proteins and confirm their integral insertion into the membrane is the use of high salt (750 mM) wash steps. Although not conducted here, this has previously been shown not to dissociate Lgt from the membrane [80].

The membrane pellets from the ultracentrifugation step were incubated with 0.5% decylmaltoside (DM) or 1% N-dodecyl-β-D-maltopyranoside (DDM) to solubilize proteins from the membrane. 1% lauryl-maltose-neopentyl-glycol (LMNG) and TritonX-100 were also tried to see if they solubiled Lgt and both were successful in solubilising (Appendix VIII). These non-ionic detergents are widely used in studies on membrane enzymes and compatible with X-ray crystallography studies such as for that of Lnt [81]. After incubation with detergents, another step of ultracentrifugation was performed to separate the newly solubilised proteins (Figure 33, Lanes 5-8). From the Western blot analysis, Lgt is clearly visualized in the fraction solubilized with 1% DDM, LMNG and TritonX-100 (Appendix VIII) but not that solubilized with 0.5% DM, as described by Mao *et al.,* (2016). However, DDM showed the highest proportion of solubilised protein. The reason for this difference in result is unknown as the protocols are identical up to this stage of processing. 1% DDM was used for further solubilization and in the purification buffers.

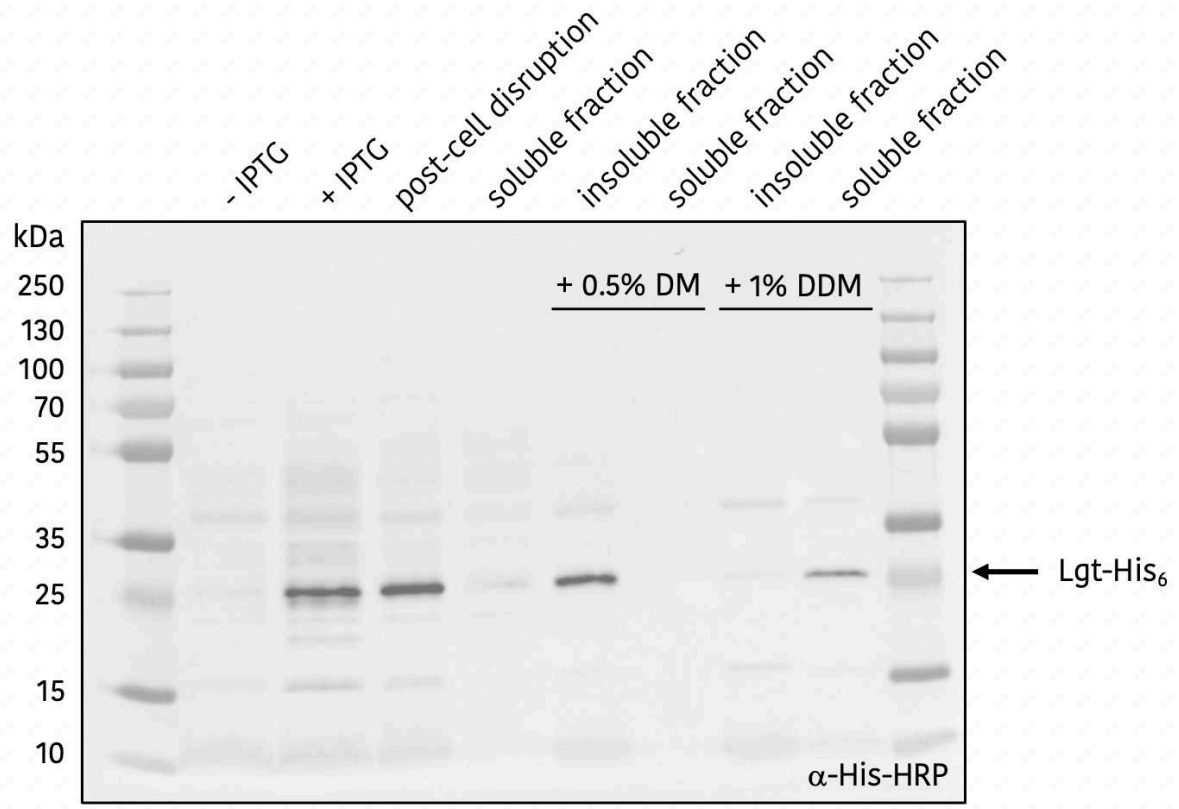


Figure 33. Western blot after SDS-PAGE with detection of His_6. *Lanes 1 & 2) SLEC20 grown with and without IPTG. Lane 3) post-cell disruption. Lane 4) after centrifugation of disrupted cells, supernatant contain soluble proteins. Lane 5 & 6 and 7 & 8) after incubation with DM or DDM, soluble fraction is supernatant after ultra-centrifugation, insoluble fraction is pellet. Lgt appears in the soluble fraction after incubation with DDM and not DM.*

Purification

The purification of Lgt involved a two-step procedure. The first step was conducted with a pre-cast His-Trap HP affinity column (Cytiva) compatible with the AKTA purifier system. During this step the concentration of DDM in the protein buffer was reduced to 0.05%. The critical micelle concentration (CMC) of DDM is 0.007% and therefore 0.05% was deemed sufficient to maintain protein solubility. Lgt was eluted with an imidazole gradient and eluted at 90-120 mM (30-40%) imidazole. Two peaks are observed, a large peak at 90 mM imidazole and small peak at 120 mM imidazole (Figure 34 B).

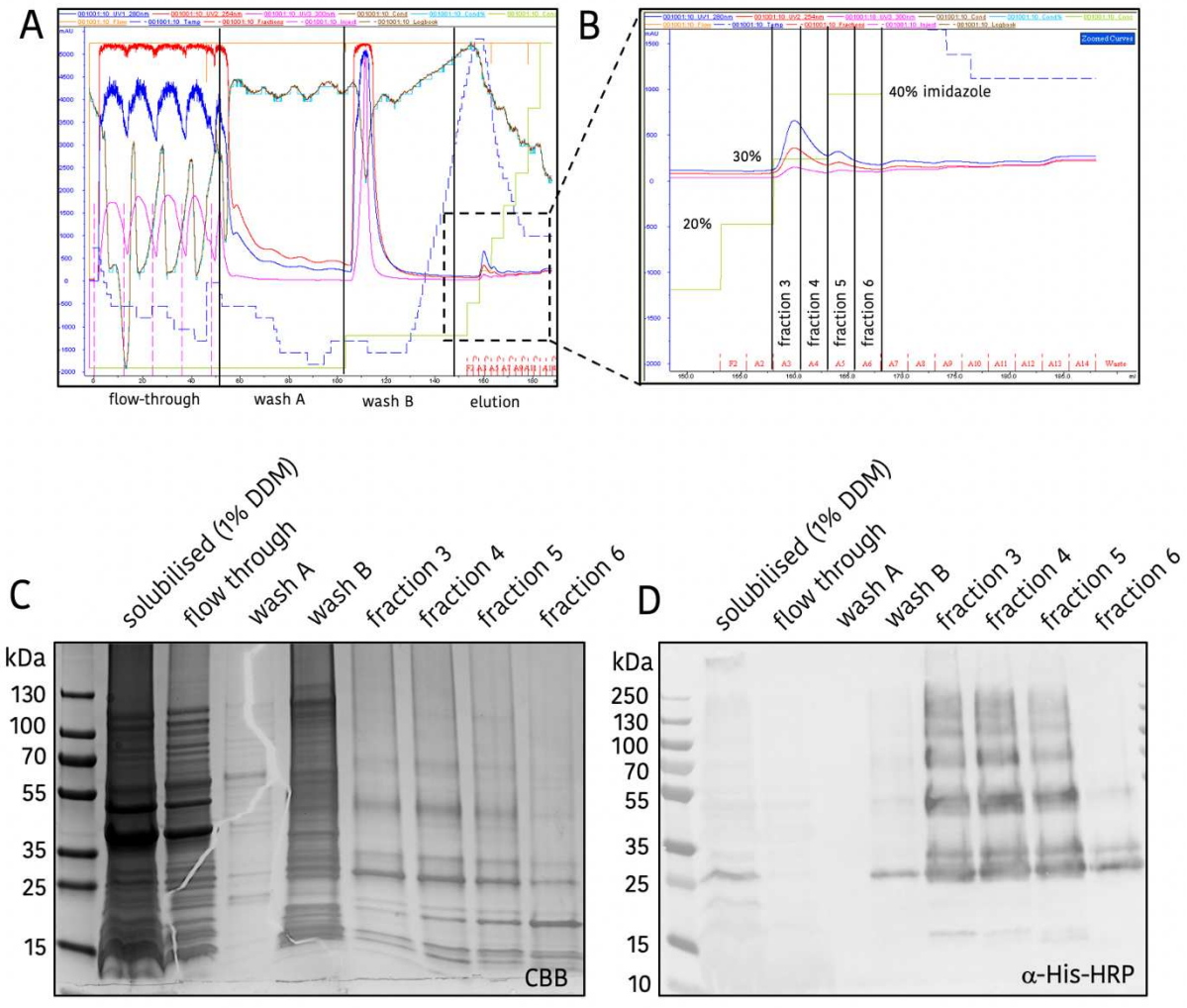


Figure 34. Affinity chromatography of Lgt-His purification. *A) Analysis of flow, B) zoomed region where Lgt is eluted. C) Coomassie Brilliant Blue gel and D) α-His-HRP Western-blot of samples, lanes relate to fractions noted in A).*

Lgt was present in each of the two peaks (Figure 34, Lanes 5-8) but a laddering effect could be observed on the α-His-HRP Western blot. It is possible that the bands correspond to multimers of Lgt as their MW are approximately multiples of the expected Lgt size of 26 kDa. Unfortunately, gels and Western blots are not presented in the work of Mao *et al.* (2016) so comparisons could not be made. The laddering was less visible in fraction 6 (Figure 34, Lane 8) but this corresponds to a region where the peak has flattened and therefore contains less total protein and possibly less Lgt. This does raise the possibility that a high concentration of Lgt in the sample is the cause for what is observed on the Western blot and that multimers are rare but are observed by the high sensitivity of α-His-HRP. Interestingly, a second band which migrates slower (30 kDa) than the expected Lgt band is visible. This raises the possibility that Lgt may contain a peptide substrate at this stage of the purification.

Samples from the Lgt containing fractions were pooled and concentrated for the second step of purification using a size exclusion chromatography column (Sephadex S75) to separate possible coeluted proteins. A single, large peak is observed (Figure 35 A & B) that eluted early in the process close to the void volume of the column. This suggests that most of the protein eluted at that time. Although other, smaller peaks from later in the elution were assessed, this early peak is the fraction where Lgt is visible (Figure 35 E & F).

The fractions containing Lgt were pooled and subjected to quality control analysis that was carried out by the PFBMI platform at Institut Pasteur. UV Spectrophotometry was used to determine non-protein contamination, protein concentration, and aggregation. Mass spectrometry MALDI-TOF was used to determine size and protein contaminants. The calculated size of the purified Lgt was 30 kDa and not the predicted 33 kDa. Dynamic light scattering (DLS) was used to determine aggregation and polydispersity of small species. The results from these analyses concluded that the purification of Lgt was successful and with few contaminants and aggregation. Further to this, LC-MS analysis was employed at the Proteomics platform at Institut Pasteur to determine if the protein was intact since the measured size differed from the theoretical size in MALDI-TOF. The whole protein and His_6-tag were confirmed present in the sample with few contaminants.

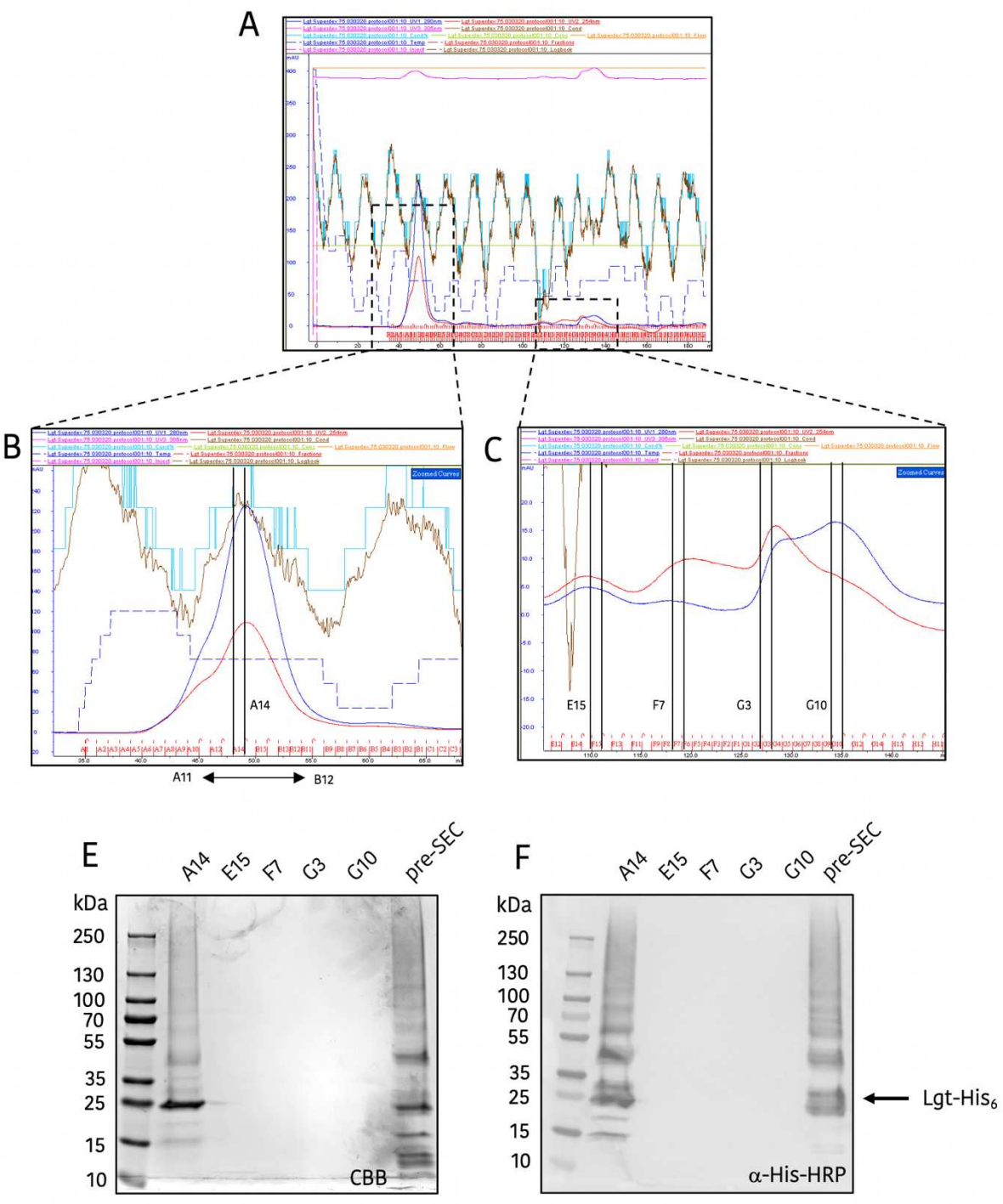


Figure 35. Size exclusion chromatography of Lgt-His purification. *A) flow of size exclusion chromatography key regions expanded; B &C. E) CBB of key fractions. F) α-His-HRP Western-blot of samples, lanes relate to fractions noted in B &C.*

Sample clean up

As the sample was deemed relatively pure by quality control we hypothesized that the multiple bands observed in Figures 34D and 35E were caused by multiple copies of Lgt being present in individual detergent micelles that were not resolved in monomers in the presence of SDS and reducing agent DTT. Initially, to confirm non-contaminating proteins in the purified sample, we removed the detergent from the samples using a chloroform-methanol extraction and precipitation method. This was applied to multiple fractions of the size exclusion purification to assess whether the same effect of laddering was consistent across the major peak. Indeed, although the Western blots and Coomassie stained gel looked cleaner, the laddering affect was still visible (Figure 36A &B). A difference in fraction density in fractions A15-B15 is likely due to protein loss in the extraction procedure. We hypothesized that DDM might not be the most efficient detergent of choice to purify Lgt. We therefore looked for other detergents that would avoid multimerization and would keep the enzyme in an active form.

Two methods of buffer exchange were used, Zeba Column (Thermo Scientific) and His-Select Gel (Sigma) (Figure 36C). Two detergents were selected, LMNG (0.25% and 0.05%) and OG (0.6% and 1%). LMNG was used in the purification of Lnt [252] and OG has been used in the past in various studies of Lgt. The results from the Western blot reveal that the presence LMNG removed most of the laddering effect when exchanged by either method (Figure 36C, Lanes 4 & 5, 8 & 9). OG did not appear to have the same result and the laddering was still visible (Figure 36C, Lanes 2 & 3, 6 & 7).

As to why the laddering appears after His-trap affinity chromatography we are unsure but the effect is reversed in the presence of LMNG suggesting that different detergents affect the proteins differently.

However, as the sample seemed pure and the quantity from the purification was low, we opted to continue with DDM to assess activity and avoid the risk of losing protein in the buffer exchange procedure.

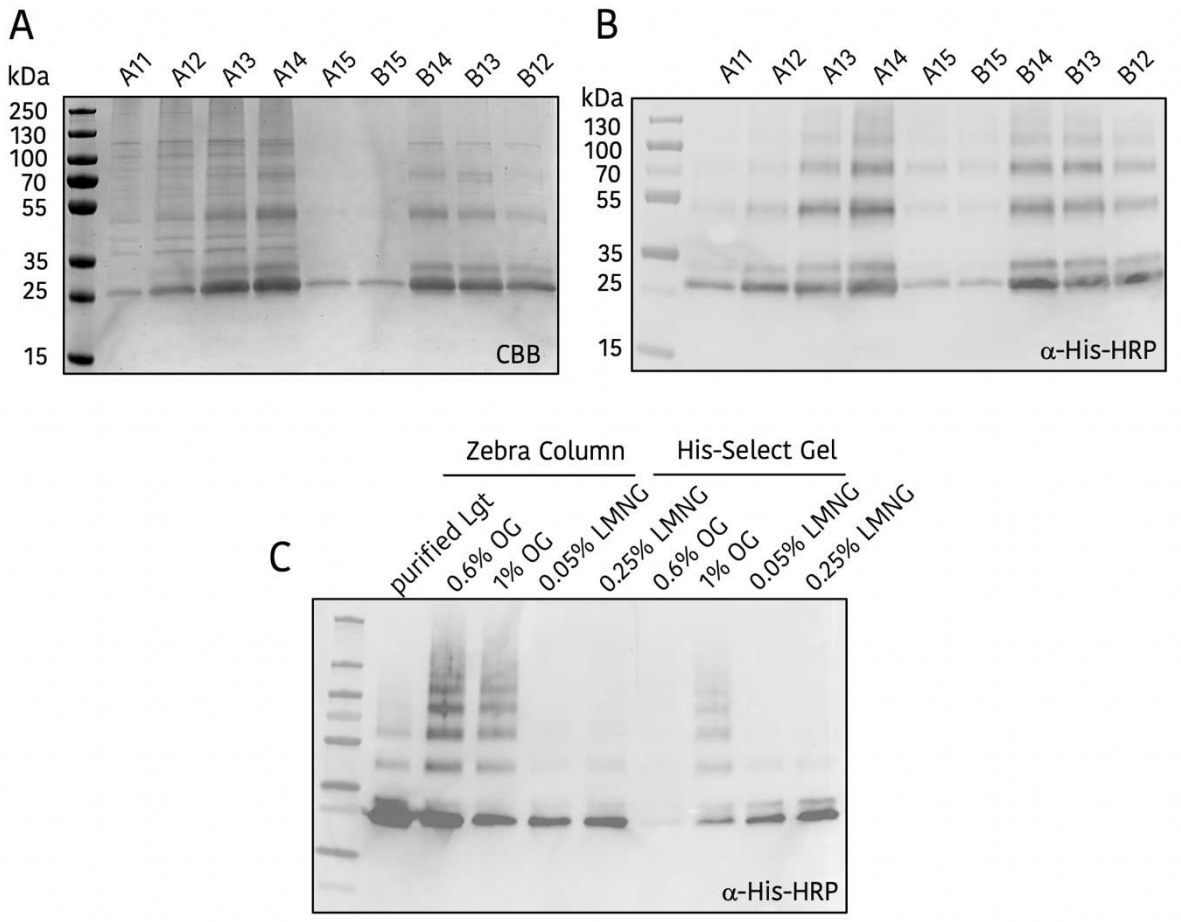


Figure 36. Troubleshooting Lgt purification. A*) CBB and B) Western-blot of purified Lgt in DDM after chloroform-methanol extraction. C) Western-blot after different buffer exchange protocols.*

In vitro Lgt enzymatic activity

The activity of Lgt has been studied previously but in most instances the enzyme has not been purified but instead inner membrane vesicles (IMVs) have been isolated and tested for activity. More recently, Mao *et al.* (2016) and Diao *et al.* (2021) have conducted assays with purified protein. After analysing reaction conditions, we concluded that an initial trial of activity with 1:20:100 ratio of Lgt, peptide and phosphatidylglycerol (PG) (μM) should be used. A peptide was designed on the basis of the abundant lipoprotein Lpp. It contained the signal peptide and four residues of the mature protein following the lipobox. A His_6-tag and biotin group were synthesised on the N- and C-terminals, respectively. These two groups were added for downstream applications. POPG was selected as the lipid substrate as it has been shown to be the primary substrate for Lgt [81].

The peptide and lipids were first sonicated to break up aggregates before adding to TED (Tris, EDTA, DTT) buffer. The substrates were pre-warmed together at 37°C before Lgt was added. As further controls, Lgt was either heat inactivated or MTSES, a cysteine alkylating agent, was added to block the thiol group of cysteine of the peptide. The reaction was further incubated overnight. The reaction was analysed in one of three ways, by gel shift, click-chemistry or mass spectrometry.

Gel Shift Assay

Analysis of Lgt activity has traditionally been conducted by analysing a change in migration on Tricine-SDS-PAGE. The addition of a diacylglyceryl (DAG) group from PG to the incoming peptide creates a mass change of 575 Da. As the peptide used has a biotin group attached, streptavidin-HRP could be used to visualize the peptide on Western blot. A schematic of the procedure is shown in Figure 37A. Previous studies used radioactive palmitate [77, 82, 93, 101] or amino-acids [89] and subsequent autoradiography as a method for visualizing a shift in migration patterns or by using fluorescent proteins [81]. However, due to the inherent risks of using radioactive compounds or the difficulties of seeing a > 1 kDa shift of a large protein on SDS-PAGE, our chemiluminescence approach was preferred. Due to the small size of the peptide (3768 Da), a Tris-Tricine 16% gel was preferred due to its ability to separate small proteins. Figure 37B displays the Western blot after gel electrophoresis. A clear mobility shift of the peptide substrate can be observed where Lgt was incubated with both substrates (Figure 37, Lane 7). A band is only observed in lanes where the peptide has been added, ensuring that the band is indeed the peptide. Heat inactivated Lgt or addition of MTSES does not induce a shift in migration, further adding to the conclusion that purified Lgt is active and that the DAG transfer reaction can be analysed by gel shift.

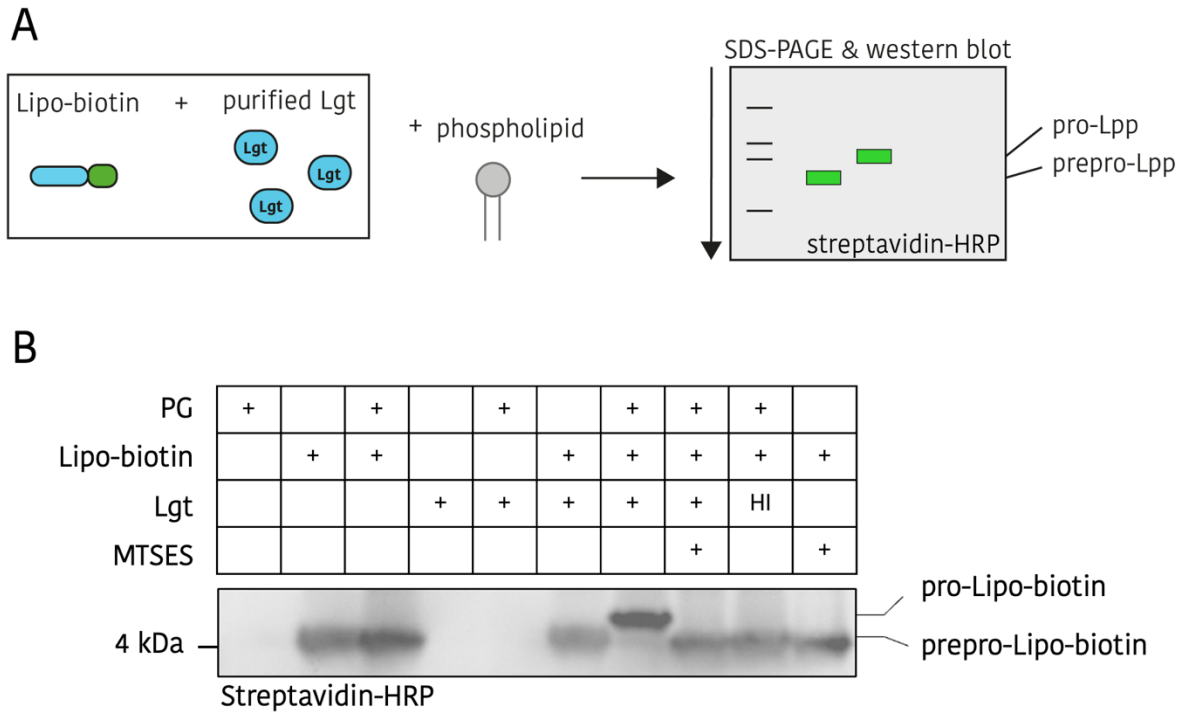


Figure 37. Activity of Lgt by gel-shift. *A) schematic of gel-shift protocol. B) Western-blot of Lgt reaction, PG = POPG, HI = heat-inactivated*

Fluorescence based assay

Although the gel shift assay is adequate for small scale analysis of Lgt activity, it is elaborative for quantitative analyses and not suited for high-throughput screening purposes. We therefore hypothesized that fluorescent phospholipids could be used in the assay in a similar way radio-labeled lipids have been used. The biotin group on the peptide enabled a binding step on streptavidin plates after the reaction and subsequent fluorescence measurement could be used to measure Lgt activity (Figure 38A). To this end, we purchased PG-NBD, a commercial phospholipid with a fluorescent group (Avanti Polar lipids) (Figure 38B). The options were limited, and PG-NBD has shorter lipid chains on the *sn*-2 position compared to standard POPG. The assay was conducted as described previously and analysed by Western-blotting to see whether PG-NBD was a substrate for the reaction.

Western blot analysis revealed that although the NBD group is visualized by Alexa 488 fluorescence, the reaction is not as efficient as with standard POPG (C16:0, C18:1 PG). A faint band is observed in the complete reaction (Figure 38Cii, Lane 7) that corresponds to a change of 575 kDa but this is also observed in the MTSES control. Due to the poor transfer of DAG-NBD, we deemed PG-NBD a poor substrate and not suitable for further analysis.

A second band migrating at approximately 2x the biotin peptides mass is also observed when PG-NBD is present, raising the possibility that NBD interacts with the peptide directly (Figure 38Ciii). This is not observed for POPG (Figure 37).

The differences in structure are notable between PG and PG-NBD. The reason why it is a poorer substrate has not been assessed here but it could be speculated upon. The central cavity is thought to house the phospholipid during the reaction has a hydrophobic base and the although the NBD group is hydrophobic and has been used as a probe for membrane applications [253] it may interfere with the reaction mechanism of the enzyme or possibly have poorer movement through the different sites of the enzyme. Future work should seek to model this probe inside the enzyme. Interestingly, PE-NBD was not a substrate for the *in vitro* activity assay of Lnt [252].

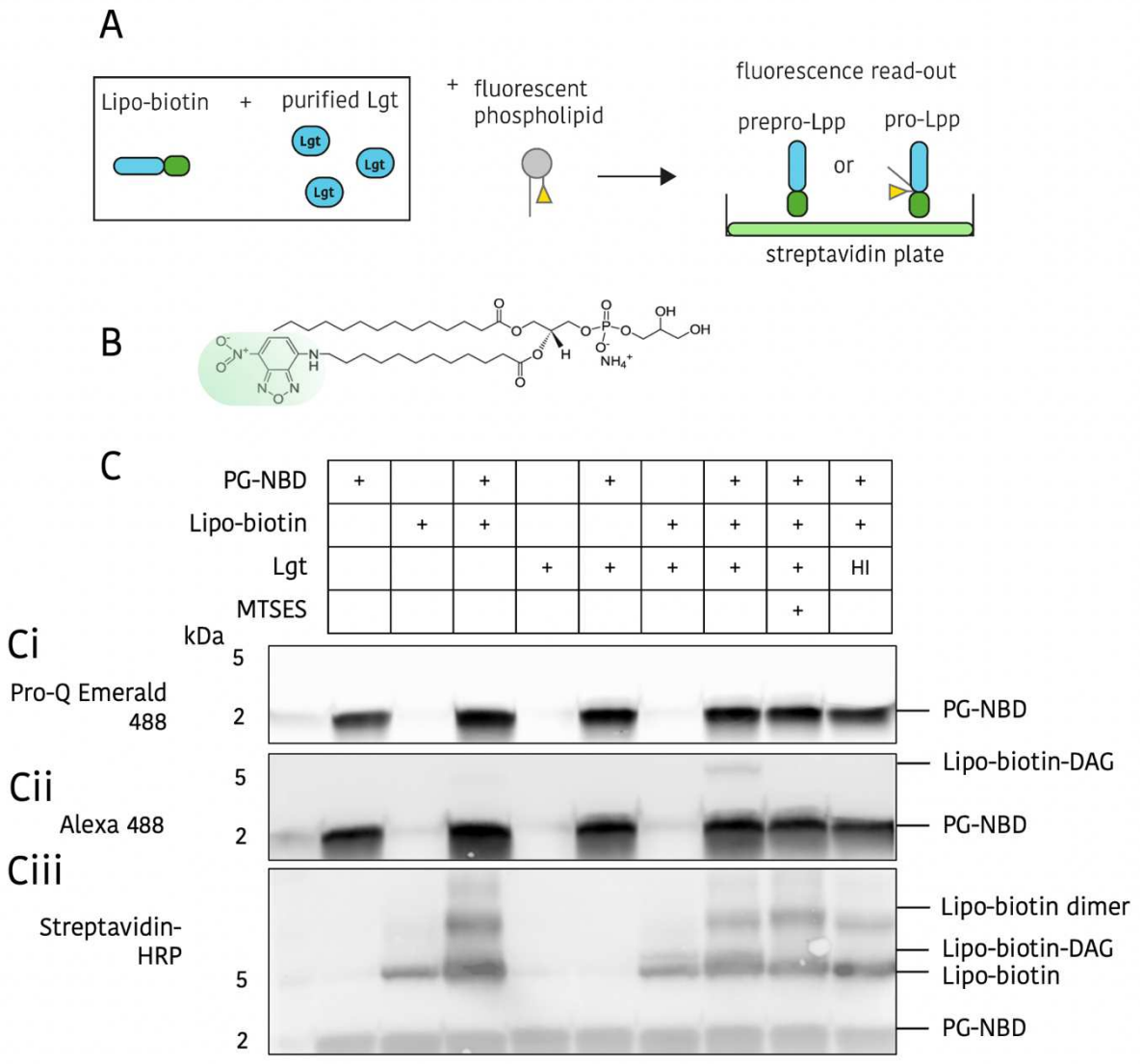


Figure 38. Activity of Lgt using fluorescent phospholipids. *A) schematic of fluorescent lipid (PG-NBD) protocol. B) PG-NBD, Ci) Fluorescence image of Tricine-SDS-PAGE, dark bands indicate NBD. Cii) Western-blot of Lgt reaction , dark bands indicate NBD, Ciii) Western-blot of Lgt reaction with streptavidin-HRP, dark bands indicate Lipo-biotin peptide. HI = heat-inactivated*

Click-chemistry based assay

As the fluorescently probed PG was not a suitable substrate we adopted a click-chemistry approach. This method takes advantage of the ability for azide and alkyne groups to react and form covalent bonds in the presence of the catalyst, copper. It is a widely used method in biochemistry and was employed to study the activity of lipoprotein modification enzyme Lnt [252]. We therefore sought to use this method to study Lgt activity. The reaction schematic can be seen in (Figure 39A). In brief, the Lgt reaction is conducted in the presence of PG containing an alkyne group on the *sn*-1 fatty acid (PG-alkyne) (Figure 39B). Once the reaction is completed an azide-tagged fluorophore is then added and the click reaction conducted. The reaction is added to a home-made streptavidin coated plate where the biotin group on the peptide binds to the plate. The unbound products are washed away and the peptide with or without DAG-tagged fluorophore remain attached. Formation of the fluorescent DAG-peptide-biotin product can be quantified by fluorescence measurements in a fluorescence spectrometer (TECAN) where fluorescence indicates incorporation of DAG onto the peptide and therefore Lgt activity.

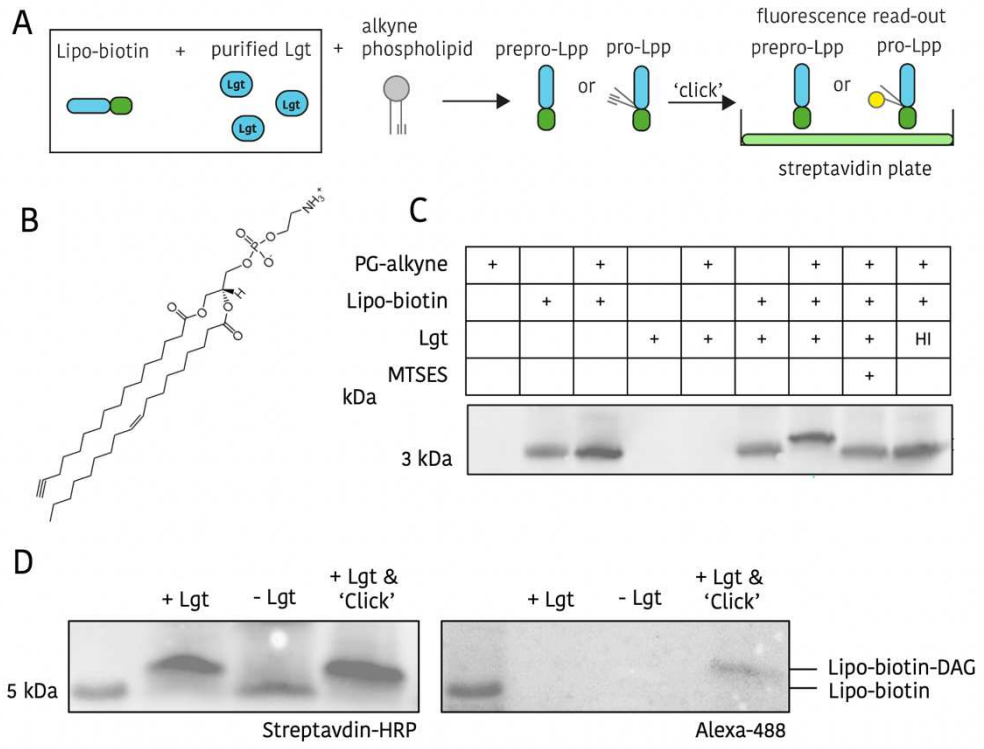


Figure 39. Click-chemistry reaction. *A) Schematic of click-chemistry Lgt assay. B) PG-alkyne. C) streptavidin-HRP Western blot of Lgt reaction. D) streptavidin-HRP Western blot of Lgt reaction after the click reaction and Alexa-488 fluorescence imaging of azide-FAM.*

The first control was to assess whether PG-alkyne was a substrate. The Lgt activity assay was completed as described previously with the substitution of POPG for PG-alkyne. As shown in (Figure 39C), PG-alkyne is a substrate when examined under the same conditions as POPG by Tricine-SDS-PAGE and Western blotting with a clear shift in migration of the peptide substrate.

The second control was to assess whether the 'click' reaction was successful in attaching azide-fluorophore (FAM) to the DAG-modified peptide. The click reaction was completed as described in the materials and methods section after the Lgt reaction was completed in the manner previously described [252, 254]. The resulting product was then migrated on a Tris-Tricine SDS-PAGE gel and imaged under Alexa 488 conditions and after Western-blotting. Figure 39D shows the faint but visible band representing the DAG modified peptide is fluorescent, indicating it has been clicked with the azide-fluorophore.

Further controls were conducted to assess whether the fluorescence of the binding control (biotin-fluorescein) and fluorophore (azide-FAM) could be measured in the TECAN plate reader (Figure 40A/B). Fluorescence read-out of the fluorescent reagent azido-FAM and biotin-fluorescein can be measured by fluorescence spectrometry and a concentration dependent linear correlation is observed. Binding of biotin-fluorescein to streptavidin-coated plates was shown efficient up to low concentrations (31 nM) and no binding was observed on non-coated plates but a high background level of fluorescence was also observed (Figure 40C). The washes were optimized to remove all non-specific binding. Three washes with PBS containing 0.1% Tween-20 was sufficient to obtain a specific fluorescent signal with biotin-fluorescein (Figure 40D).

However, when the full Lgt reaction was conducted, no fluorescent signal is observed in the reaction where DAG-peptide-biotin was formed (Figure 40E). The background fluorescence is also high (around 20,000 RFU for buffer alone) which may mask minor changes in fluorescence. This was not observed in the Lnt reaction which saw minimal background fluorescence [252]. A possible adjustment to the protocol could be the addition of BSA which will prevent non-specific binding and therefore may reduce background fluorescence.

The final step of the reaction which may be causing the issue is the binding of the peptide to the plate. Biotin-fluorescein bound successfully to the plate (Figure 40C) but the physical properties of the peptide may interfere with binding. To assess this, an assay was developed whereby the His_6 group on the peptide was utilised. The peptide was added to the plate in two concentrations (20 μM, 1 μM or PBS only) in wells with and without streptavidin. His-HRP was added to the wells and standard washing procedure was followed. One well containing His-HRP only was used as a positive control for detecting the chemiluminescence signal. Wells with buffers were used as negative controls (Figure 40F).

Western-blotting ECL chemiluminescence substrate was used to produce a readable signal. Figure 40F shows a signal in the well containing His-HRP only and not in the negative control wells. However, no signal was observed in the wells containing the peptide, with or without streptavidin present. It is possible that the quantity of peptide is too low and therefore α-His-HRP is below the detection limit. A further control of biotin-His would allow better interpretation of these results. However, it is possible that the peptide is not binding efficiently to the plate.

The chemical properties of the reaction, the hydrophobic peptide and high phospholipid concentration could all contribute to the lack of peptide binding. This adds to the complexity of the assay. Due to the lack of signal and the complexity of this *in vitro* system, it was decided to pursue other routes such as Mass Spectrometry, alongside optimising this protocol.

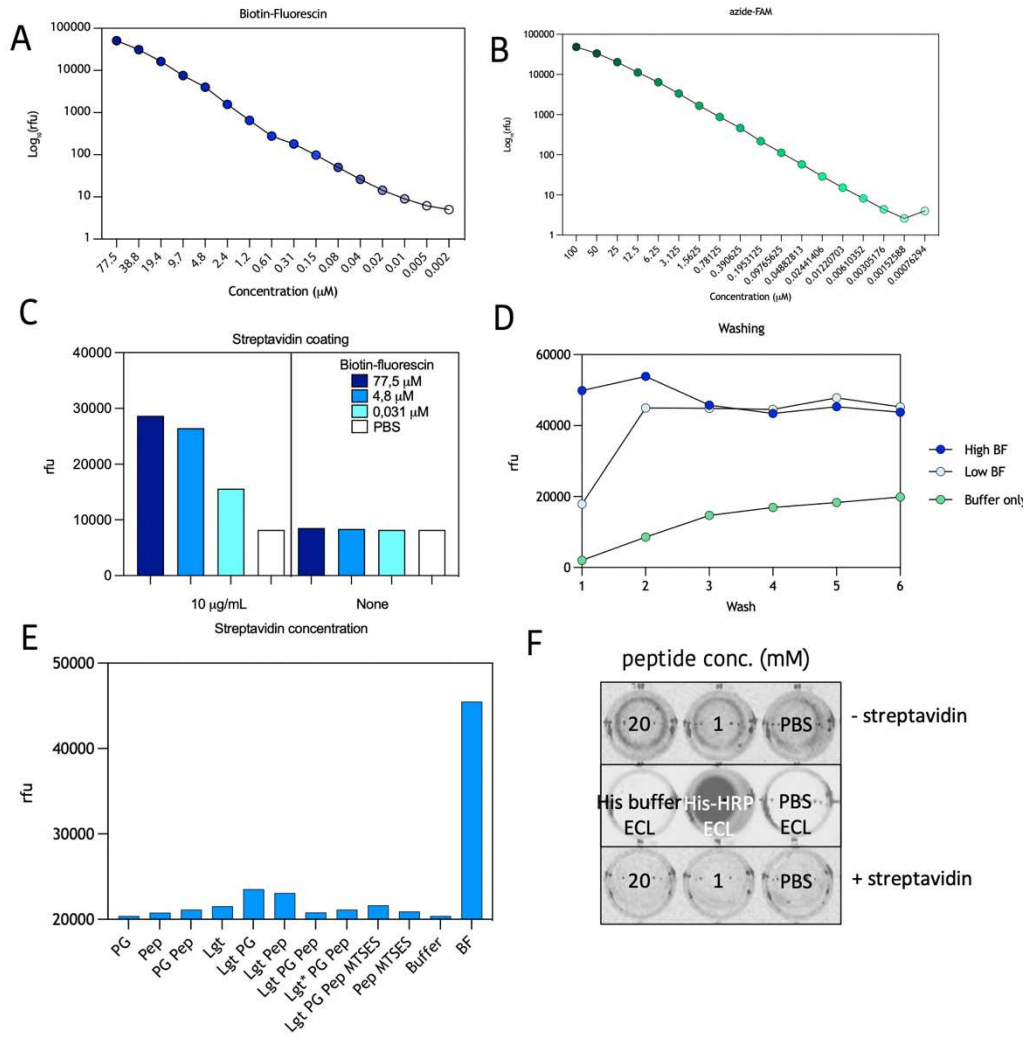


Figure 40. Click-chemistry Lgt assay. *A) and B) detection of fluorescence by TECAN plate reader. C) assessment of streptavidin coating of 96 well plates with biotin-fluorescein. D) effect of multiple washing steps on biotin-fluorescein detection. E) Lgt click-chemistry assay, BF = biotin-fluorescein, PG = PG-alkyne, Pep = peptide, * = heat-inactivated Lgt. F) Troubleshooting of peptide binding to 96 well plates, dark signal represents chemiluminescent signal in the well and therefore presence of peptide.*

Mass Spectrometry assay

Mass spectrometry was explored as a tool to study Lgt enzymatic activity. The advantages of Mass Spectrometry is the ability to use native substrates such as phospholipids and peptide substrates. Thus far, the assays have required the peptide to have a detectable tag, such as streptavidin or the phospholipids to be modified with a fluorescence or alkyne group. Each of these modifications could interfere with the optimal activity of the enzyme. However, for preliminary experiments and to maintain consistency with the click-based and gel-shift assays, we continued the use of the substrates described above.

The Lgt activity reaction as shown in Figure 37 was conducted and the reaction was applied to Liquid Chromatography-Mass Spectrometry (LC-MS) with a Q Exactive Plus Orbitrap Mass Spectrometer. Although the peptide was detected in each of the reactions where it was applied, no peptide containing a diacylglyceryl group was observed. Interestingly, the reaction containing both substrates and the active Lgt was seen to have a very low intensity compared to the other reactions. This suggests that less peptide may be reaching the Mass-Spectrometer, possibly due to the liquid-chromatography step preceding this. Due to the hydrophobic nature of the peptide with diacylglyceryl attached it may be that the modified peptide is failing to elute from the column. This method will be further adapted in light of the recently published protocol by Diao *et al.* (2021).

The LC step preceding the MS prevents this method being adaptable for high throughput applications and adds complexity to the assay. Matrix Assisted Laser Desorption Ionization - Time of Flight (MALDI-TOF) Mass Spectrometry provides a method of mass spectrometry that does not require a LC preparation step, but the reaction can be loaded directly onto a matrix and analysed with minimal sample preparation. This method has been shown to adaptable for high throughput applications such as drug-screening via *in vitro* activity assays [255]. This may provide a solution to study Lgt activity in a high throughput adaptable manner.

Conclusion

In order to study Lgt *in vitro* and to screen large libraries of molecules for enzymatic inhibition we sought to develop an assay that is adaptable for high throughput.

We successfully purified Lgt from *E. coli* in a similar manner to that previously described [81] though as DM failed to solubilised Lgt, we adapted the solubilisation procedure to include DDM. We found Lgt to be present in the membrane fraction and not in the soluble fraction as described previously [78, 116] and now contested in a number of studies [80, 81]. However, we did not perform sucrose gradients to test specifically its location but this has been conducted by Diao *et al.* (2021) and Pailler *et al.* (2012). During the purification procedure we noted a 'laddering' effect detected on α-His-HRP Western-blots. We hypothesised that this was an effect of the reduction in detergent concentration and therefore exchanged the buffer from DDM to LMNG or OG. We found that the exchange for LMNG removed the majority of the additional bands.

We tested the activity of the enzyme in an adapted gel-shift activity assay. This assay borrows principles from previous assays which employed radio-labelled substrates [76, 93] or a GFP-tagged peptide substrate [81]. Our assay employed a biotin-tagged peptide as a peptide substrate and PG as the lipid donor. After incubation with the enzyme the reaction was migrated on a Tricine-SDS-PAGE gel and analysed by Western-blot with streptavidin-HRP. From this assay, the incorporation of diacylglyceryl by Lgt is measured by a migration shift of the biotin-peptide. Via this assay, we determined that the purified enzyme was active. However, this detection method is not adaptable for high throughput.

We therefore sought to see whether a fluorescent phospholipid could be used in the assay. PG-NBD was not a suitable substrate so we moved to a click-based assay based upon an Lnt activity assay [252, 254]. This assay employs an alkyne-PG which can be attached to an azide-tagged fluorophore in a reaction known as click-chemistry after the Lgt reaction has been performed. The product can be bound to a streptavidin coated plate via the peptides biotin group and fluorescence measured as an indicator of Lgt activity. We show that PG-alkyne is a substrate for the reaction and the click reaction is successful. However, we met issues that are likely related to the peptide binding to the plate due to its hydrophobicity. This assay is being optimised further.

One issue with the click-chemistry based assay is its complexity due to hydrophobicity of the reaction mixture and the need for modified substrates such as custom-made PG-alkyne and synthetic peptides. We therefore sought to develop a method to study Lgt activity using mass spectrometry. A similar method has since been reported by Diao *et al.* (2021). A method using MALDI-TOF MS would entail minimal sample preparation and native substrates could still be used. There are also examples of high throughput MALDI-TOF mass spectrometry. We therefore believe this to be a suitable avenue to explore further.

Materials and Methods

Strains and vector design

pET28b was used as a cloning vector. *E. coli lgt* without its stop codon was amplified from pCHAP7546 (pBAD18-*lgt*) from previous studies of Lgt [80] with primers pr2E and pr2D which contained flanking regions of pET28b for Gibson cloning. pET28b was digested with NcoI and XhoI by incubation at 37°C for 4 hours. Amplified *lgt* and digested pET28b were ligated via Gibson Assembly (NEB) by incubation for 30 minutes at 50°C in a manner which maintained a C-terminal His_6-tag in the transcribed region. Gibson Assembly products were transformed into chemically competent *E. coli* DH5α cells (NEB) and incubated overnight on LBA-Kan50 (50 mg/mL kanamycin). Colonies were selected and streaked again on LBA-Kan50. Successful insertion of the *lgt* gene into pET28b was initially checked via PCR with primers pr3A and pr3B which amply the multiple cloning site (MCS) of the plasmid. Candidates with an insert of approximately 900 base pairs were sent for sequencing (Eurofins) to confirm correct formation of the plasmid (designated P09). P09 was transformed into chemically competent *E. coli* BL21 (DE3) containing the T7 RNA polymerase gene which is under control of an IPTG-inducible P_{lac} promoter (SLEC20).

Table 10. Strains, plasmids and primers for Section III

Strain	**Details**
BL21	*E. coli* str. B F^- *ompT gal dcm lon hsdS*$_B$($r_B^- m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K\text{-}12}$(λ^S)
SLEC20	BL21 + SLP09
Plasmids	
pCHAP7546	pBAD18-Lgt
SLP09	pET28b-Lgt-His_6
pET28b	OverExpress cloning vector
Primer	**Sequence**
pr2E	GGCCTGGTGCCGCGCGGCAGCCATAATGACCAGTAGCTATC
pr2D	GTGGTGGTGGTGCTCGAGGGAAACGTGTTGCTGTGGG
pr3A	cgaaattaatacgactcactataggg
pr3B	gctagttattgctcagcgg

SDS-PAGE and Western-blotting for protein purification

Where required samples were centrifuged at 13,200 x*g* in a table-top centrifuge and the supernatant was removed. The pellet was resuspended in sample buffer (10% glycerol, 2.5% SDS, 100 mM Tris pH8, 10 mM DTT, phenol red) to a ratio of 0.01 OD_{600} units / μL. All samples were heated at 100°C for 10 minutes before loading onto an SDS-PAGE Stain-Free 4-15% gel (BioRad) and migrated at 20 mA. Gels were either analysed by Coomasie Brilliant Blue (CBB) or Western-blotting. For CBB, gels were washed three times for 1 minute in water in incubated in CBB for 1 hour under gentle agitation. The gel was then washed extensively to remove excess dye and imaged on a BioRad GelDoc. For Western-blotting,

a BioRad TurboBlot instrument was used with nitrocellulose membranes. The blot was washed in water before incubation with PonceauS solution for 1 minute. The blot was washed with water and then incubated in blocking buffer (5% BSA, 1 x PBS, 0.5% Tween 20) for 1 hour at room temperature being replaced by an α-His-HRP (Sigma 1:10,000) antibody solution (1x PBS, 0.1% Tween 20, 5% BSA) and incubated for a further 1 hour at RT. The blot was washed 2 x for 5 minutes the 3x 10 minutes with 0.1% PBS-T. ECL Western blotting chemiluminescence substrate was applied as per manufacturer instructions (Pierce). The blot was imaged on a ChemiDoc imager (Bio-Rad).

Overproduction of Lgt-His6

One colony of SLEC20 was resuspended in 50 mL LB with 50 µg/mL kanamycin (Kan50) and grown at 180 rpm and 37°C overnight. The culture was diluted 1:100 in 4 L of Terrific Broth (TB-Kan50) and grown to an OD_{600} 0.6-0.8. Expression of *lgt* was induced by the addition of IPTG at 0.5 mM. Cells were grown overnight (> 20 h) at 16°C. Cells were harvested by centrifugation at 4,000 x *g* for 30 minutes at 4°C and resuspended in buffer A (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10% glycerol). Protease inhibitors (Complete Roche EDTA free) were added prior to cell fractionation.

Membrane Preparation and Solubilization of Lgt

All steps were carried out at 4 °C. Cells were broken by cell disruption (CellD or French Press) through three passages at 10,000 psi. DNAse and RNAse were added to a final concentration of 10 µg/mL together with 50 mM $MgSO_4$. Unbroken cells and debris were removed by centrifugation at 18,000 x g for 15 min at 4°C. Membranes were pelleted by ultracentrifugation at 100,000 x g for 60 min at 4°C and solubilized in buffer B (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol, 1% (w/v) DDM (or 0.5% DM, 1% LMNG or 1% TritonX-100), EDTA-Free protease inhibitor) for 2 hours at room temperature. All insoluble material was removed by centrifugation at 100,000 x g for 60 min at 4°C. The supernatant containing detergent-soluble proteins was retained and the pellet containing insoluble proteins was resuspended in buffer A.

Two-step purification of Lgt-His_6

The first step is by affinity chromatography on a His-Trap HP column coupled to an Äkta purification system. The column was equilibrated with 2 column volumes (CV) H_2O followed by 2 CV buffer C (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol, 1% (w/v) DDM). The supernatant was applied to the column. The column was washed with 5 CV buffer D (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol, 0.05% (w/v) DDM) and 5 CV buffer D (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol (v/v), 0.05% (w/v) DDM) + 40 mM imidazole and the protein was eluted with buffer D with a gradient of imidazole up to 300 mM. The presence of Lgt-His was in each fraction was determined by SDS-PAGE

and Western-blotting. The fractions containing the protein were pooled and concentrated in a Vivaspin20 10,000 MWCO column (Sartorius).

In the second step, Lgt was purified by size exclusion chromatography using a prepacked HiLoad Superdex 75 16/600 column equilibrated in buffer D (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol (v/v), 0.05% (w/v) DDM). The protein was added and eluted in buffer D. The presence of Lgt-His was in each fraction was determined by SDS-PAGE and Western-blotting. The fractions containing the protein were pooled and concentrated in a Vivaspin20 10,000 MWCO column (Sartorius).

Quality control

Quality control was conducted by the PFBMI platform at the Institut Pasteur. V-650 UV-VIS Spectrophotometric measurements were performed on a JASCO from 240-340 nm, bandwidth 2 nm. MALDI-TOF and DLS measurements were performed as per the platform's standard procedures.

Detergent removal

Detergent removal was performed on a sample of the purified protein by methanol-chloroform extraction. 4 volumes MeOH and 1 volume CHCl3 were added to the sample. The samples were vortexed for 30 seconds before 3 volumes of water was added and mixed. The samples were then centrifuged for 10 minutes at room temperature to separate two phases. The aqueous phase (supernatant) was removed, and 4 volumes methanol was added to precipitate the protein. The samples were then collected and analysed by SDS-PAGE and Western-blotting.

Buffer exchange

Two methods were employed to exchange the buffer and therefore the detergent present in the purified Lgt sample. The selected exchange buffers were (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol (v/v)) + either 0.6% OG, 1% OG, 0.05% LMNG, or 0.25% LMNG.

Zeba Spin Column

The Zeba Spin column was centrifuged at 1,000 x *g* for 2 minutes to remove the manufacturers storage solution and to compact the resin. 500 µL of exchange buffer was added to the column and was centrifuged at 1,000 xg for 2 minutes. This wash step was repeated 6 times. 175 µL purified Lgt was loaded (8.3 µg/mL) onto the column followed by a stacker 40 µL exchange buffer. The Zeba Spin column was centrifuged at 1,000 xg for 2 minutes to elute the protein. Samples were analysed via SDS-PAGE and Western blotting.

His-Select Nickel Affinity

The second method of buffer exchange took advantage of the His-affinity tag on the protein. 50 μL of His-Select Nickel Affinity gel was added to a microcentrifuge tube and centrifuged at 5,000 xg for 30 seconds. The supernatant containing storage solution was discarded. 200 μL of exchange buffer was added to the tube and mixed well and centrifuged at 5,000 x*g* for 30 seconds. The supernatant was discarded. 100 μL purified Lgt was loaded (8.3 μg/mL) into the tube, mixed and centrifuged at 5,000 x*g* for 30 seconds. The supernatant was stored at -20°C. The gel with the Lgt-His protein was washed by the addition of the exchange buffer and subsequent rounds of centrifugation three times. After the final wash step, 50 μL of exchange buffer + 250 mM imidazole was added to the gel. The sample was centrifuged at 5,000 xg for 30 seconds. This was repeated three times and each time the supernatant containing the eluted protein was pooled. Samples were analysed via SDS-PAGE and Western blotting.

Lgt reaction

The Lgt reaction was consistent for each of the activity assays unless stated otherwise. The peptide substrate was purchased from ProteoGenix based on the Lpp signal peptide but containing a His-tag and biotin group (MRSHHHHHHKATKLVLGAVILGSTLLAGCSSN-Lys(Biotin)) and was suspended in water. Various phospholipids were used; POPG (1-palmitoyl-2-Oleoyl-sn-Glycero-3-(Phospho-rac-(1-glycerol))); PG-NBD (1-myristoyl-2- (12-((7-nitro- 2-1,3-benzoxadiazol-4- yl)amino) dodecanoyl)-sn-glycero-3- (Phospho-rac-(1-glycerol)))); or PG-alkyne (1-hexadec-15-ynoyl-2-oleoyl-sn-glycero-3-phospho-(rac-(1-glycerol)). Phospholipids were suspended in TED buffer (20 mM Tris-HCl pH 8, 4 mM DTT, 5 mM EDTA) + 0.05% DDM (TED-DDM). Phospholipids and peptide were sonicated in a water bath for 3 minutes prior to their addition to the reaction. The reaction was performed in TED buffer with 100 μM phospholipid, 20 μM peptide and 1 μM purified Lgt. In the absence of either substrates or enzyme, their respective suspension buffers replaced them. Before Lgt was added the substrates were pre-warmed to 37°C. The reaction was incubated overnight at 37°C unless otherwise stated.

Tris-Tricine SDS-PAGE and Western-blotting

Samples from the Lgt reaction were heated at 100°C for 10 minutes in the presence of BioRad 3x Tricine Sample Buffer (200 mM Tris-HCl pH 6.8, 40% glycerol, 2% SDS, 0.04% Coomasie Blue). Samples were then loaded onto a self-cast or Precast 16% Tricine-PAGE (Invitrogen). The Self-Cast Tris-Tricine Gel contained three layers. The first layer (20%) contained 6 M urea, Gel Buffer (1 M Tris-HCl, 0.3% SDS, pH 8.45), Acrylamide 19:1 (33%) and water. The second layer (11%) contained 4 M urea, Gel Buffer, Acrylamide 19:1 (33%) and water. The final layer (6%) contained Gel Buffer, Acrylamide 19:1 (33%) and water. Each layer was poured individually after the addition of Ammonium Persulfate (APS) and 1,2-Bis(dimethylamino)ethane (TEMED).

Gels were run in the presence of Anode buffer (1 M Tris, 0.225 M HCl, pH 8.9) and Cathode buffer (1 M Tris, 1 M tricine, 1% SDS, pH 8.25) for self-cast gels or Cathode buffer only for pre-cast gels. After samples we loaded the gels we run at 100 V for a minimum of 4 hours.

If required, the gels were imaged for fluorescence on a BioRad ChemiDoc under Alexa488 settings to image either the products of the click reaction of PG-NBD.

The Western-blotting procedure was followed as described above. Briefly, after precipitation with PonceauS the blot was washed 3 x for 5 minutes in PBS-T (PBS + 0.1% Tween20). Streptavidin-HRP (1/15,000, PBS) was added and incubated 30 minutes. The blot was washed 2 x for 5 minutes then 3 x for 10 minutes with 0.1% PBS-T. ECL-Femto Western blotting chemiluminescence substrate was applied as per manufacturer instructions (Pierce). The blot was imaged on a ChemiDoc imager (Bio-Rad).

Click assay

Fluorophore fluorescence

To check whether the Azide-FAM or Biotin-Fluorescein could be detected on the fluorescent spectrophotometer (TECAN) ½ serial dilutions from 77.5 μM Biotin-Fluorescin and 100 μM azide-FAM were performed in clear-bottom black-sided 96-well plates. Fluorescence was measured with excitation at 485 nm and emission at 535 nm.

Streptavidin-coating

Streptavidin (20 μL) was added to a 96-well plate in three concentrations (2 mg/mL, 0.1 mg/mL and 0.01 mg/mL in water) to assess optimal coverage. Once added to the wells, the plates were dried at 37°C. Biotin-fluorescein diluted in PBS was added at 77.5 μM, 4.8 μM, 0.031 μM or PBS only to the streptavidin coated plate and incubated for 90 minutes. 30 μM of PBS was added to each well and the wells were washed 3 x in wash buffer (PBS 1 % Tween 20) and then 3 x in PBS by mixing for 5 minutes and then rinsing 3 x in fresh wash buffer and fluorescence measured with excitation at 485 nm and emission at 535 nm. Readings were taken after each wash step to assess the effect of washing on the measurements.

Click reaction

The Lgt reaction was performed as described above in the presence of PG-alkyne. The click-reaction was performed as previously described [252, 254] . 50 μM Azido-FAM in DMSO, 1 mM TCEP (Tris(2-carboxyethyl)phosphine hydrochloride prepared in water), 1 mM TBTA (Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine; in tert-butanol:DMSO (4:1)) were added to the reaction mixture. The

solution was vortexed and 1 mM CuSO4·5H2O added before being incubated in the dark at RT for 1 hour.

To check whether the click reaction was successful, the sample was prepared for Tricine-SDS-PAGE as described above and the gel and Western-blots were imaged with Alexa-488 settings on the a BioRad GelDoc.

To assess whether the reaction binds to the streptavidin-coated 96 well plates the reaction mixture was transferred to the plates and incubated for 1 hour at RT, 300 rpm in a tabletop thermocycler. The plates were washed as described above with an additional 3 x washes in PBS only. The fluorescence was measured with excitation at 485 nm and emission at 535 nm in a TECAN plate reader.

Mass Spectrometry

Lgt reaction samples were prepared as described above. LC-MS was performed by the Proteomics platform (Institut Pasteur) using Q Exactive Plus (Q1) Orbitrap MS and analysed with MaxQuant (1.6.6.0).

Section V: Conclusion and perspectives

This thesis set out to explore the hypothesis that Lgt is a good target for novel antibiotics. We have explored the conservation of Lgt, it's essentiality in *E. coli* and explored methods for an *in vitro* enzymatic study.

The need to develop novel antibiotics is clear. With high rates of AMR related deaths, the crisis of AMR is a global problem [8]. Various mechanisms are being employed to tackle this problem: from bacterial vaccines and microbiota alterations to increasing awareness and improving the delivery of existing antibiotics [12]. Another such option is to develop new antibiotics. These could be against existing targets, although resistance may develop rapidly against these inhibitors, or against novel targets. Reports suggest the ideal antibiotic will have multiple targets or target pathways as this reduces the occurrence of resistance.

Of the pathogens which cause the most mortality due to AMR, a majority are diderm bacteria and fall into the category of Gram-negative bacteria. There are a number of essential processes in all bacteria such as DNA synthesis and replication or protein synthesis. Some uniquely essential pathways exist in Gram-negative bacteria due to the importance of their outer membrane. These essential pathways include the Bam system which inserts outer membrane proteins into the membrane, the LPS translocation pathway which transports and inserts the outer-leaflet lipopolysaccharide molecule to its destination and finally the lipoprotein modification pathway which is required to modify lipoproteins involved in the Bam and LPS pathways. None of these three pathways are the target of antibiotics. As the lipoprotein modification pathway is required for the other two essential pathways, we have explored whether it could be a good target for novel antibiotics. This pathway is made up of three enzymes, Lgt, Lsp and Lnt followed the Lol translocation machinery [38]. As the first enzyme in the pathway and an extremely well conserved protein, we chose to study Lgt in greater detail.

To challenge the hypothesis that Lgt is a good target for novel antibiotics we employed three axes of study. The first was to study the conservation of Lgt. The second was to understand its essentiality in greater detail and finally we sought to develop an *in vitro* assay to screen for inhibitors.

To hone our efforts, we selected key AMR pathogens to study in more detail. We found that Lgt is well conserved at a sequence level with the exception of the previously described key HGGL motif. This motif contains the essential residue H^{103} which has been speculated a key residue in the catalytic activity of Lgt. However, in concurrence with previous reports, H^{103} is not widely conserved with tyrosine or tryptophan observed in the selected pathogens [78, 79, 81]. This raises the possibility that H^{103}, while clearly essential for activity, may not be the proposed catalytic base for the reaction. The location of H^{103} in the side cleft, a possible exit channel for substrates may play a more structural role as a gate, allowing or blocking substrate entry. Further study is required before we can conclude that

the previous models of Lgt enzymatic activity are correct. The computational models use a shortened lipobox peptide and not the full-length lipoprotein [79]. Therefore, the role of the extended, periplasmic region of the lipoprotein may not be assessed. With the high predicted protein-protein binding of L6-7, arm-1, arm-2 and some head domains, further modelling with full length lipoprotein may shed light on possible lipoprotein-Lgt interactions.

A newly described conserved sequence, termed the neck motif found below the enzymes head domain is highly conserved within firmicutes and within proteobacteria with its proximity near the Lgt signature motif in the proposed catalytic core of the enzyme. The role of the neck motif should be explored further and may provide species specificity for the enzyme.

We explored whether Lgt from *A. baummannii*, *P. aeruginosa* and *H. pylori* could complement two depletion strains, Δlgt^{P} and Δlgt^{C}. Δlgt^{P} has wild type *lgt* present on a multicopy plasmid and Δlgt^{C} has it present as a single copy on the chromosome. In both strains, wild type *lgt* is under the control of P_{ara}. Complementing *lgt* genes are present on a pAM238 plasmid under the control of P_{lac}.

We observed that while growth was restored fully in Δlgt^{P}, this may be due to residual expression of the wild type *lgt* under the control of P_{ara} on a high copy number plasmid and therefore the ability for the complementing Lgts to restore growth could be interpreted more akin to provided support for the residual wild type Lgt and that the cells are not fully reliant on the complementing enzyme. In the Δlgt^{C} model, the issue of residual Lgt expression is largely removed due to the single copy of wild type Lgt present under the control of P_{ara}. In this strain we observe that only Lgt from *H. pylori* can fully complement growth and viability. This is surprising as *H. pylori* is the more evolutionary distant of the three enzymes. Lgt from a wider range of species, especially clinically relevant Gram-positive bacteria such as *Streptococcus spp*. should be included in further complementation studies. To date, all successful complementation studies in *E. coli* have had Lgt enzymes which have H^{103} and not a variation of this key residue.

The advent for high confidence structural prediction software such as AlphaFold2 has enabled us to analyse differences in the structure of Lgt from the key pathogenic species. While they are predicted to be highly similar across the majority of the enzyme, the periplasmic exposed head domain showed high variability. Enterobacteriales have a large head domain with two alpha-helices that protrudes high above the membrane plane and *M. tuberculosis* has a larger head domain still. The firmicutes and more distance proteobacterial pathogens had a smaller head domain that may rest closer to the membrane plane. We hypothesized that the head domain may have a role in localizing the components of the lipoprotein modification pathway into proximity of each other. However, ColabFold prediction deemed interactions between Lgt, Lsp and Lnt unlikely. The head domain of *E.*

coli Lgt had low predicted protein-protein binding but the head domains a the distant proteobacteria *H. pylori*, the firmicute *S. aureus* and *M. tuberculosis* had higher predicted protein-protein binding. This suggests the periplasmic head domain may have functional significance. To explore this experimentally, we cloned the head groups of *H. pylori, S. aureus* and *M. tuberculosis* into the *E. coli lgt* gene and expressed it in the Lgt depletion strains Δlgt^{P} and Δlgt^{C}. We found that only the head domain from *H. pylori* could complement growth of the depletion strains. This further suggests the head domains may play a functional role. However, although the proteins were produced and protein production did not appear to correlate with viability and growth, we have not yet ascertained the localization of the proteins or their stability. As the core of the enzyme is well conserved the head domains may provide a site of inhibition that is more narrow spectrum whereas the core of the enzyme may provide broader spectrum inhibition.

The second axis of the study was to explore the essentiality of Lgt. Inhibitory molecules of the lipoprotein modification pathway such as those which inhibit Lsp or the Lol machinery lose their inhibitory capabilities in *E. coli* when the abundant lipoprotein Lpp is removed. It has been shown that the ability for Lpp to crosslink to the peptidoglycan while mislocalised to the inner membrane is the cause of cell death in these incidents. We therefore sought to ascertain whether this was true for inhibition of Lgt. As we are yet to have an inhibitor of Lgt we used two depletion models to study the loss of Lgt, Δlgt^{P} and Δlgt^{C} [75, 80]. We observed that in Δlgt^{P}, Lgt production from P_{ara} under repressed conditions, below detectable levels, was likely sufficient to restore growth after a long lag phase. These growth revertants were explored further and we observed that although the revertant growth phenotype allowed permanent wild type-type growth in repressed conditions, there were no genetic changes on the chromosome but the plasmid could not be cured. It is likely that the restoration of growth is not simple from residual Lgt production from P_{ara} as the extended lag-phase is not observed when Δlgt^{P} is grown in low levels of P_{ara} inducer, L-arabinose. Deletion of *lpp* in conditions when *lgt* was not actively expressed was also possible in Δlgt^{P}.

However, in Δlgt^{C}, where there is a single copy of *lgt* under the control of P_{ara}, growth and viability were not restored under restrictive conditions and deletion of Lpp did not rescue growth. This suggested that an alternative mechanism of cell death was occurring. It should be noted that Lpp rescued growth and viability at low levels of *lgt* expression in this background showing that although Lpp is not the sole cause of cell death, the presence of the abundant lipoprotein is a factor. During this study we observed severe morphological changes to the cell under *lgt* depletion conditions which differed in the presence and absence of Lpp. When Lgt is depleted, cells become more rounded at the poles, multiple poles are observed and their area increases before eventually lysing. In the absence of

Lpp, the cells elongate and do not round up at the poles but still eventually lyse. We concluded that although the Lpp is a factor in cell death in the absence of lipoprotein modification, other essential or key lipoproteins have a larger role. In the absence of Lpp, the residual *lgt* expression from P_{ara} in Δlgt^{P} is sufficient to allow the modification of enough key lipoproteins, such as LolA, or BamB-E for example, to maintain survival of the cells. However, when this residual expression is reduced in Δlgt^{C}, even in the absence of Lpp, there is insufficient processing of lipoproteins to maintain survival [250].

Therefore, in agreement with a previous report [75] which showed that Lpp deletion did not reduce inhibition of Lgt by a peptide inhibitor, we observed too that Lpp is not the main cause of death under Lgt depletion conditions.

What this study has brought out is the important nature of the pathway for maintaining viability but also cell shape. The morphological changes to the cell when Lgt is reduced are severe and a previous report has shown that an Lgt depletion strain is more susceptible to serum killing. These data suggest that incomplete inhibition of Lgt could cause a sufficient effect on the cell that total inhibition may not be required in an infection environment. As discussed in Section I, partial inhibition may be sufficient to have an effect on fitness and therefore support the immune system, making it easier to clear the infection. Lgt depletion has previously been shown to reduce the MIC to a range of antibiotics [75] and our observation that Lgt has a servere effect on cell morphology could point to Lgt being a good target as part of a combination therapy. With the advent of screening assays for Lnt or the Lol machinery [74, 252] and some inhibitors being described for Lsp and LolCDE [71, 72, 74, 256] it raises the possibility of a combination therapy with multiple targets in the same pathway.

The full lipoproteome of *E. coli* is yet to be comprehensively defined and dozens of lipoproteins in *E. coli* have no assigned function. This is true for many other species of bacteria and therefore the full impact of inhibiting the lipoprotein modification pathway is not known. Some studies have sought to study the lipoproteomes of bacteria by applying proteomics and lipidomics by MALDI-TOF MS and by these methods discovered novel modifications [255]. This work should be pursued further. A method for the determination and definition of lipoproteomes of bacteria would hopefully refine and improve the prediction tools which are heavily relied upon. By identifying the lipoproteins we may then have a better understanding of the roles they perform. It may also shed light on the possible species specificity or preference of lipoproteins for Lgt. As we observed differential complementation ability of Lgt enzymes from different species in *E. coli*, we hypothesise that this may be due to lipoprotein-Lgt specificity. For example, the lipoproteins of *E. coli* can be supported better by Lgt from *H. pylori* than Lgt from *A. baummannii*.

Greater *in vivo* studies of Lgt depletion or inhibition would also give us a deeper understanding of the roles of lipoproteins in infection models and may highlight other targets for novel antibiotics.

Our results confirm that Lgt is indeed essential in *E. coli*, and that *lgt* is present as a single copy in 22 key pathogenic bacteria. As discussed, single, essential gene products as antibiotic targets are likely to have increase rates of resistance development as the selection pressure is high [21, 22]. This is certainly an issue with Lgt where no other protein has been discovered to conduct the same role. Since we observed a revertant growth phenotype when Δlgt^{P} is grown in restrictive conditions that is likely not caused solely by basal expression from P_{ara}, it raises the possibility that other cellular mechanisms are supporting cell growth and viability in the absence of sufficient Lgt. Diao *et al.* (2021) state that they do not observe resistance development to their cyclic peptide inhibitor under laboratory conditions which is a promising sign but this itself raises the question as to whether Lgt is the sole target of their inhibitor or if it has multiple effects on the cell.

Finally, we sought to develop an *in vitro* assay to study Lgt and to screen for inhibitors. We successfully purified Lgt after alterations to previously report purification protocols were made and found that the enzyme was active in our adapted gel-shift assay. We were unsuccessful in developing a quantitative fluorescence-based assay but this avenue should be pursued further. We speculate that the use of mass-spectrometry, particularly MALDI-TOF may provide a better option for studying Lgt. This method can be adapted for high throughput and requires minimal sample preparation. It has been shown in principle to work [75] but is yet to used for screening purposes.

There are other areas which require further study. Some bacteria have multiple Lgt enzymes but the full scale of this is not yet known. A full-scale phylogenetic analysis should be conducted to ascertain which organisms have multiple Lgts and what the physiological benefit of this may be. The regulation mechanisms of LMP is not known and as it has been demonstrated that different forms of lipoproteins may be preferred under certain conditions, the mechanism by which this is controlled would be an interesting area of study.

Lgt is essential and present in all key AMR pathogens, it has a conserved central cavity thought to be the catalytic core which may provide a site of broad spectrum inhibition but also has a variable head domain that could be targeted with more narrow spectrum inhibitors. The function of the head group is still unknown. A more efficient screen needs to be developed for Lgt activity in order to find targeted inhibitors. Without inhibitors, it is hard to assess their potential *in vivo*, G2824 provides an insight but alone is not sufficient to answer the hypothesis. However, with the urgent need for new antibiotics, targeting Lgt as a novel target shows some promise.

Section VI: Bibliography

1. Armbruster, K.M. and T.C. Meredith, *Identification of the Lyso-Form N-Acyl Intramolecular Transferase in Low-GC Firmicutes.* J Bacteriol, 2017. **199**(11): p. e00099-17.
2. Vernon, G., *Syphilis and Salvarsan.* Br J Gen Pract, 2019. **69**(682): p. 246.
3. Hutchings, M.I., A.W. Truman, and B. Wilkinson, *Antibiotics: past, present and future.* Curr Opin Microbiol, 2019. **51**: p. 72-80.
4. Blair, J.M., et al., *Molecular mechanisms of antibiotic resistance.* Nat Rev Microbiol, 2015. **13**(1): p. 42-51.
5. Langford, B.J. and A.M. Morris, *Is it time to stop counselling patients to "finish the course of antibiotics"?* Can Pharm J (Ott), 2017. **150**(6): p. 349-350.
6. Llewelyn, M.J., et al., *The antibiotic course has had its day.* BMJ, 2017. **358**: p. j3418.
7. O'Neill, J., *Tackling drug-resistant infections globally: Final report and recommendations*. 2016: UK Government.
8. Antimicrobial Resistance, C., *Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis.* Lancet, 2022. **399**(10325): p. 629-655.
9. Organisation, W.H., *Prioritization of pathogens to guide discovery, research and development of new antibiotics for drug-resistant bacterial infections, including tuberculosis*. 2017.
10. Rice, L.B., *Federal funding for the study of antimicrobial resistance in nosocomial pathogens: no ESKAPE.* J Infect Dis, 2008. **197**(8): p. 1079-81.
11. CDC, *Antibiotic resistance threats in the United States*. 2019.
12. Organisation, W.H., *Global action plan on antimicrobial resistance.* 2016.
13. Holm, M., et al., *Measuring the Link Between Vaccines and Antimicrobial Resistance in Low Resource Settings – Limitations and Opportunities in Direct and Indirect Assessments and Implications for Impact Studies.* Frontiers in Tropical Diseases, 2022. **3**.
14. Jansen, K.U., et al., *The impact of human vaccines on bacterial antimicrobial resistance. A review.* Environ Chem Lett, 2021. **19**(6): p. 4031-4062.
15. Langford, B.J., et al., *Antibiotic prescribing in patients with COVID-19: rapid review and meta-analysis.* Clin Microbiol Infect, 2021. **27**(4): p. 520-531.
16. Kwong, J.C., et al., *The effect of universal influenza immunization on antibiotic prescriptions: an ecological study.* Clin Infect Dis, 2009. **49**(5): p. 750-6.
17. Van Boeckel, T.P., et al., *Global trends in antimicrobial use in food animals.* Proc Natl Acad Sci U S A, 2015. **112**(18): p. 5649-54.
18. Micoli, F., et al., *The role of vaccines in combatting antimicrobial resistance.* Nat Rev Microbiol, 2021. **19**(5): p. 287-302.
19. Organisation, W.H., *2020 antibacterial agents in clinical and preclinical development: an overview and analysis.* 2021.
20. Ribeiro da Cunha, B., L.P. Fonseca, and C.R.C. Calado, *Antibiotic Discovery: Where Have We Come from, Where Do We Go?* Antibiotics (Basel), 2019. **8**(2).
21. Silver, L.L., *Appropriate Targets for Antibacterial Drugs.* Cold Spring Harb Perspect Med, 2016. **6**(12).
22. Silver, L.L., *Multi-targeting by monotherapeutic antibacterials.* Nat Rev Drug Discov, 2007. **6**(1): p. 41-55.
23. Baba, T., et al., *Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection.* Molecular Systems Biology, 2006. **2**.
24. Goodall, E.C.A., et al., *The Essential Genome of Escherichia coli K-12.* mBio, 2018. **9**(1).
25. Rousset, F., et al., *Genome-wide CRISPR-dCas9 screens in E. coli identify essential genes and phage host factors.* PLoS Genet, 2018. **14**(11): p. e1007749.
26. Zhang, R., H.Y. Ou, and C.T. Zhang, *DEG: a database of essential genes.* Nucleic Acids Res, 2004. **32**(Database issue): p. D271-2.

27. Xu, L., Z. Guo, and X. Liu, *Prediction of essential genes in prokaryote based on artificial neural network.* Genes Genomics, 2020. **42**(1): p. 97-106.
28. Paradis-Bleau, C., et al., *Lipoprotein cofactors located in the outer membrane activate bacterial cell wall polymerases.* Cell, 2010. **143**(7): p. 1110-20.
29. Monserrat-Martinez, A., Y. Gambin, and E. Sierecki, *Thinking Outside the Bug: Molecular Targets and Strategies to Overcome Antibiotic Resistance.* Int J Mol Sci, 2019. **20**(6).
30. Tyers, M. and G.D. Wright, *Drug combinations: a strategy to extend the life of antibiotics in the 21st century.* Nat Rev Microbiol, 2019. **17**(3): p. 141-155.
31. Sohlenkamp, C. and O. Geiger, *Bacterial membrane lipids: diversity in structures and pathways.* FEMS Microbiol Rev, 2016. **40**(1): p. 133-59.
32. Elias-Wolff, F., et al., *Curvature sensing by cardiolipin in simulated buckled membranes.* Soft Matter, 2019. **15**(4): p. 792-802.
33. Oliver, P.M., et al., *Localization of Anionic Phospholipids in Escherichia coli Cells.* Journal of Bacteriology, 2014. **196**(19): p. 3386-3398.
34. Sauvage, E., et al., *The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis.* FEMS Microbiol Rev, 2008. **32**(2): p. 234-58.
35. Kracke, F., I. Vassilev, and J.O. Kromer, *Microbial electron transport and energy conservation - the foundation for optimizing bioelectrochemical systems.* Front Microbiol, 2015. **6**: p. 575.
36. Mitchell, A.M. and T.J. Silhavy, *Envelope stress responses: balancing damage repair and toxicity.* Nat Rev Microbiol, 2019. **17**(7): p. 417-428.
37. May, K.L., et al., *A Stress Response Monitoring Lipoprotein Trafficking to the Outer Membrane.* mBio, 2019. **10**(3).
38. Legood, S., I.G. Boneca, and N. Buddelmeijer, *Mode of action of lipoprotein modification enzymes-Novel antibacterial targets.* Mol Microbiol, 2021. **115**(3): p. 356-365.
39. Dalbey, R.E. and A. Kuhn, *Protein traffic in Gram-negative bacteria--how exported and secreted proteins find their way.* FEMS Microbiol Rev, 2012. **36**(6): p. 1023-45.
40. Denks, K., et al., *The Sec translocon mediated protein transport in prokaryotes and eukaryotes.* Mol Membr Biol, 2014. **31**(2-3): p. 58-84.
41. Ruiz, N., *Lipid Flippases for Bacterial Peptidoglycan Biosynthesis.* Lipid Insights, 2015. **8**(Suppl 1): p. 21-31.
42. Workman, S.D. and N.C.J. Strynadka, *A Slippery Scaffold: Synthesis and Recycling of the Bacterial Cell Wall Carrier Lipid.* J Mol Biol, 2020. **432**(18): p. 4964-4982.
43. Harvat, E.M., et al., *Lysophospholipid flipping across the Escherichia coli inner membrane catalyzed by a transporter (LplT) belonging to the major facilitator superfamily.* J Biol Chem, 2005. **280**(12): p. 12028-34.
44. Egan, A.J.F., J. Errington, and W. Vollmer, *Regulation of peptidoglycan synthesis and remodelling.* Nat Rev Microbiol, 2020. **18**(8): p. 446-460.
45. Sarkar, P., et al., *A review on cell wall synthesis inhibitors with an emphasis on glycopeptide antibiotics.* Medchemcomm, 2017. **8**(3): p. 516-533.
46. Sperandeo, P., A.M. Martorana, and A. Polissi, *The lipopolysaccharide transport (Lpt) machinery: A nonconventional transporter for lipopolysaccharide assembly at the outer membrane of Gram-negative bacteria.* J Biol Chem, 2017. **292**(44): p. 17981-17990.
47. Mas, G., et al., *Regulation of chaperone function by coupled folding and oligomerization.* Sci Adv, 2020. **6**(43).
48. Wu, R., et al., *The big BAM theory: An open and closed case?* Biochim Biophys Acta Biomembr, 2020. **1862**(1): p. 183062.
49. Han, L., et al., *Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins.* Nat Struct Mol Biol, 2016. **23**(3): p. 192-6.
50. Sharma, S., et al., *Mechanism of LolCDE as a molecular extruder of bacterial triacylated lipoproteins.* Nat Commun, 2021. **12**(1): p. 4687.

51. Grabowicz, M. and T.J. Silhavy, *Redefining the essential trafficking pathway for outer membrane lipoproteins.* Proc Natl Acad Sci U S A, 2017. **114**(18): p. 4769-4774.
52. Grabowicz, M., *Lipoprotein Transport: Greasing the Machines of Outer Membrane Biogenesis: Re-Examining Lipoprotein Transport Mechanisms Among Diverse Gram-Negative Bacteria While Exploring New Discoveries and Questions.* Bioessays, 2018. **40**(4): p. e1700187.
53. Yamaguchi, K., F. Yu, and M. Inouye, *A single amino acid determinant of the membrane localization of lipoproteins in E. coli.* Cell, 1988. **53**(3): p. 423-32.
54. Lewenza, S., M.M. Mhlanga, and A.P. Pugsley, *Novel inner membrane retention signals in Pseudomonas aeruginosa lipoproteins.* J Bacteriol, 2008. **190**(18): p. 6119-25.
55. Lewenza, S., D. Vidal-Ingigliardi, and A.P. Pugsley, *Direct visualization of red fluorescent lipoproteins indicates conservation of the membrane sorting rules in the family Enterobacteriaceae.* J Bacteriol, 2006. **188**(10): p. 3516-24.
56. Simpson, B.W. and M.S. Trent, *Pushing the envelope: LPS modifications and their consequences.* Nat Rev Microbiol, 2019. **17**(7): p. 403-416.
57. Mazgaeen, L. and P. Gurung, *Recent Advances in Lipopolysaccharide Recognition Systems.* Int J Mol Sci, 2020. **21**(2).
58. Rao, S., et al., *Characterizing Membrane Association and Periplasmic Transfer of Bacterial Lipoproteins through Molecular Dynamics Simulations.* Structure, 2020. **28**(4): p. 475-487 e3.
59. Hooda, Y. and T.F. Moraes, *Translocation of lipoproteins to the surface of gram negative bacteria.* Curr Opin Struct Biol, 2018. **51**: p. 73-79.
60. Huynh, M.S., et al., *Reconstitution of surface lipoprotein translocation through the Slam translocon.* Elife, 2022. **11**.
61. Prajapati, J.D., U. Kleinekathofer, and M. Winterhalter, *How to Enter a Bacterium: Bacterial Porins and the Permeation of Antibiotics.* Chem Rev, 2021. **121**(9): p. 5158-5192.
62. Mathelie-Guinlet, M., et al., *Lipoprotein Lpp regulates the mechanical properties of the E. coli cell envelope.* Nat Commun, 2020. **11**(1): p. 1789.
63. Magnet, S., et al., *Identification of the L,D-transpeptidases responsible for attachment of the Braun lipoprotein to Escherichia coli peptidoglycan.* J Bacteriol, 2007. **189**(10): p. 3927-31.
64. McLeod, S.M., et al., *Small-molecule inhibitors of gram-negative lipoprotein trafficking discovered by phenotypic screening.* J Bacteriol, 2015. **197**(6): p. 1075-82.
65. Yakushi, T., et al., *Lethality of the covalent linkage between mislocalized major outer membrane lipoprotein and the peptidoglycan of Escherichia coli.* J. Bacteriol., 1997. **179**: p. 2857-2862.
66. Xiao, Y. and D. Wall, *Genetic redundancy, proximity, and functionality of lspA, the target of antibiotic TA, in the Myxococcus xanthus producer strain.* J Bacteriol, 2014. **196**(6): p. 1174-83.
67. Hussain, M., S. Ichihara, and S. Mizushima, *Accumulation of glyceride-containing precursor of the outer membrane lipoprotein in the cytoplasmic membrane of Escherichia coli treated with globomycin.* J Biol Chem, 1980. **255**(8): p. 3707-12.
68. Stone, K.J. and J.L. Strominger, *Mechanism of action of bacitracin: complexation with metal ion and C 55 -isoprenyl pyrophosphate.* Proc Natl Acad Sci U S A, 1971. **68**(12): p. 3223-7.
69. Bush, K. and P.A. Bradford, *beta-Lactams and beta-Lactamase Inhibitors: An Overview.* Cold Spring Harb Perspect Med, 2016. **6**(8).
70. Mohapatra, S.S., S.K. Dwibedy, and I. Padhy, *Polymyxins, the last-resort antibiotics: Mode of action, resistance emergence, and potential solutions.* J Biosci, 2021. **46**.
71. Inukai, M., et al., *Globomycin, a new peptide antibiotic with spheroplast-forming activity. I. Taxonomy of producing organisms and fermentation.* J Antibiot (Tokyo), 1978. **31**(5): p. 410-20.
72. Xiao, Y., et al., *Myxobacterium-produced antibiotic TA (myxovirescin) inhibits type II signal peptidase.* Antimicrob Agents Chemother, 2012. **56**(4): p. 2014-21.

73. Pantua, H., et al., *Unstable Mechanisms of Resistance to Inhibitors of Escherichia coli Lipoprotein Signal Peptidase.* mBio, 2020. **11**(5).
74. Nickerson, N.N., et al., *A Novel Inhibitor of the LolCDE ABC Transporter Essential for Lipoprotein Trafficking in Gram-Negative Bacteria.* Antimicrob Agents Chemother, 2018. **62**(4).
75. Diao, J., et al., *Inhibition of Escherichia coli Lipoprotein Diacylglyceryl Transferase Is Insensitive to Resistance Caused by Deletion of Braun's Lipoprotein.* J Bacteriol, 2021. **203**(13): p. e0014921.
76. Gan, K., et al., *The umpA gene of Escherichia coli encodes phosphatidylglycerol:prolipoprotein diacylglyceryl transferase (lgt) and regulates thymidylate synthase levels through translational coupling.* J Bacteriol, 1995. **177**(7): p. 1879-82.
77. Qi, H.Y., et al., *Structure-function relationship of bacterial prolipoprotein diacylglyceryl transferase: functionally significant conserved regions.* J Bacteriol, 1995. **177**(23): p. 6820-4.
78. Banerjee, S. and K. Sankaran, *First ever isolation of bacterial prolipoprotein diacylglyceryl transferase in single step from Lactococcus lactis.* Protein Expr Purif, 2013. **87**(2): p. 120-8.
79. Singh, W., et al., *Mechanism of Phosphatidylglycerol Activation Catalyzed by Prolipoprotein Diacylglyceryl Transferase.* J Phys Chem B, 2019. **123**(33): p. 7092-7102.
80. Pailler, J., et al., *Phosphatidylglycerol::prolipoprotein diacylglyceryl transferase (Lgt) of Escherichia coli has seven transmembrane segments, and its essential residues are embedded in the membrane.* J Bacteriol, 2012. **194**(9): p. 2142-51.
81. Mao, G., et al., *Crystal structure of E. coli lipoprotein diacylglyceryl transferase.* Nat Commun, 2016. **7**: p. 10198.
82. Sankaran, K., et al., *Roles of histidine-103 and tyrosine-235 in the function of the prolipoprotein diacylglyceryl transferase of Escherichia coli.* J Bacteriol, 1997. **179**(9): p. 2944-8.
83. Tschumi, A., et al., *Functional analyses of mycobacterial lipoprotein diacylglyceryl transferase and comparative secretome analysis of a mycobacterial lgt mutant.* J Bacteriol, 2012. **194**(15): p. 3938-49.
84. Thompson, B.J., et al., *Investigating lipoprotein biogenesis and function in the model Gram-positive bacterium Streptomyces coelicolor.* Mol Microbiol, 2010. **77**(4): p. 943-57.
85. Widdick, D.A., et al., *Dissecting the complete lipoprotein biogenesis pathway in Streptomyces scabies.* Mol Microbiol, 2011. **80**(5): p. 1395-412.
86. Chung, S.T. and G.R. Greenberg, *Loss of an essential function of Escherichia coli by deletions in the thyA region.* J Bacteriol, 1973. **116**(3): p. 1145-9.
87. Bell-Pedersen, D., J.L. Galloway Salvo, and M. Belfort, *A transcription terminator in the thymidylate synthase (thyA) structural gene of Escherichia coli and construction of a viable thyA::Kmr deletion.* J Bacteriol, 1991. **173**(3): p. 1193-200.
88. Williams, M.G., et al., *Identification and genetic mapping of the structural gene for an essential Escherichia coli membrane protein.* J Bacteriol, 1989. **171**(1): p. 565-8.
89. Gan, K., et al., *Isolation and characterization of a temperature-sensitive mutant of Salmonella typhimurium defective in prolipoprotein modification.* J. Biol. Chem., 1993. **268**: p. 16544-16550.
90. Reffuveille, F., et al., *The prolipoprotein diacylglyceryl transferase (Lgt) of Enterococcus faecalis contributes to virulence.* Microbiology (Reading), 2012. **158**(Pt 3): p. 816-825.
91. Stoll, H., et al., *Staphylococcus aureus deficient in lipidation of prelipoproteins is attenuated in growth and immune activation.* Infect Immun, 2005. **73**(4): p. 2411-23.
92. Babu, M.M., et al., *A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins.* J Bacteriol, 2006. **188**(8): p. 2761-73.
93. Selvan, A.T. and K. Sankaran, *Localization and characterization of prolipoprotein diacylglyceryl transferase (Lgt) critical in bacterial lipoprotein biosynthesis.* Biochimie, 2008. **90**(11-12): p. 1647-55.

94. Baumgartner, M., et al., *Inactivation of Lgt allows systematic characterization of lipoproteins from Listeria monocytogenes.* J Bacteriol, 2007. **189**(2): p. 313-24.
95. Inouye, S., et al., *Prolipoprotein signal peptidase of Escherichia coli requires a cysteine residue at the cleavage site.* EMBO J, 1983. **2**(1): p. 87-91.
96. Lin, J.J., et al., *An Escherichia coli mutant with an amino acid alteration within the signal sequence of outer membrane prolipoprotein.* Proc. Natl Acad. Sci. USA., 1978. **75**: p. 489-495.
97. Wu, H.C., et al., *Biochemical characterization of a mutant lipoprotein of Escherichia coli.* Proc Natl Acad Sci U S A, 1977. **74**(4): p. 1388-92.
98. Wu, H.C. and J.J. Lin, *Escherichia coli mutants altered in murein lipoprotein.* J Bacteriol, 1976. **126**(1): p. 147-56.
99. Inouye, S., et al., *Effects of mutations at glycine residues in the hydrophobic region of the Escherichia coli prolipoprotein signal peptide on the secretion across the membrane.* J Biol Chem, 1984. **259**(6): p. 3729-33.
100. Lee, N., H. Yamagata, and M. Inouye, *Inhibition of secretion of a mutant lipoprotein across the cytoplasmic membrane by the wild-type lipoprotein of the Escherichia coli outer membrane.* J Bacteriol, 1983. **155**(1): p. 407-11.
101. Sankaran, K. and H.C. Wu, *Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol.* J Biol Chem, 1994. **269**(31): p. 19701-6.
102. Rowlett, V.W., et al., *Impact of Membrane Phospholipid Alterations in Escherichia coli on Cellular Function and Bacterial Stress Adaptation.* J Bacteriol, 2017. **199**(13).
103. Brulle, J.K., et al., *Cloning, expression and characterization of Mycobacterium tuberculosis lipoprotein LprF.* Biochem Biophys Res Commun, 2010. **391**(1): p. 679-84.
104. Brulle, J.K., A. Tschumi, and P. Sander, *Lipoproteins of slow-growing Mycobacteria carry three fatty acids and are N-acylated by apolipoprotein N-acyltransferase BCG_2070c.* BMC Microbiol, 2013. **13**: p. 223.
105. Jackson, M., *The mycobacterial cell envelope-lipids.* Cold Spring Harb Perspect Med, 2014. **4**(10).
106. Arimoto, T. and T. Igarashi, *Role of prolipoprotein diacylglyceryl transferase (Lgt) and lipoprotein-specific signal peptidase II (LspA) in localization and physiological function of lipoprotein MsmE in Streptococcus mutans.* Oral Microbiol Immunol, 2008. **23**(6): p. 515-9.
107. Denham, E.L., P.N. Ward, and J.A. Leigh, *In the absence of Lgt, lipoproteins are shed from Streptococcus uberis independently of Lsp.* Microbiology (Reading), 2009. **155**(Pt 1): p. 134-141.
108. Henneke, P., et al., *Lipoproteins are critical TLR2 activating toxins in group B streptococcal sepsis.* J Immunol, 2008. **180**(9): p. 6149-58.
109. Chimalapati, S., et al., *Effects of deletion of the Streptococcus pneumoniae lipoprotein diacylglyceryl transferase gene lgt on ABC transporter function and on growth in vivo.* PLoS One, 2012. **7**(7): p. e41393.
110. Petit, C.M., et al., *Lipid modification of prelipoproteins is dispensable for growth in vitro but essential for virulence in Streptococcus pneumoniae.* FEMS Microbiol Lett, 2001. **200**(2): p. 229-33.
111. Dautin, N., et al., *Role of the unique, non-essential phosphatidylglycerol::prolipoprotein diacylglyceryl transferase (Lgt) in Corynebacterium glutamicum.* Microbiology (Reading), 2020. **166**(8): p. 759-776.
112. Braun, V. and V. Bosch, *Sequence of the murein-lipoprotein and the attachment site of the lipid.* Eur J Biochem, 1972. **28**(1): p. 51-69.
113. Braun, V. and K. Rehn, *Chemical characterization, spatial distribution and function of a lipoprotein (murein-lipoprotein) of the E. coli cell wall.* Eur J Biochem, 1969. **10**: p. 426-438.
114. Braun, V. and U. Sieglin, *The covalent murein-lipoprotein structure of the Escherichia coli cell wall. The attachment site of the lipoprotein on the murein.* Eur. J. Biochem., 1970. **13**(2): p. 336-46.

115. Lloyd, A.L., D.A. Rasko, and H.L. Mobley, *Defining genomic islands and uropathogen-specific genes in uropathogenic Escherichia coli.* J Bacteriol, 2007. **189**(9): p. 3532-46.
116. Sangith, N., S. Kumar, and K. Sankaran, *Evidence to Suggest Bacterial Lipoprotein Diacylglyceryl Transferase (Lgt) is a Weakly Associated Inner Membrane Protein.* J Membr Biol, 2019. **252**(6): p. 563-575.
117. Sundaram, S., S. Banerjee, and K. Sankaran, *The first nonradioactive fluorescence assay for phosphatidylglycerol:prolipoprotein diacylglyceryl transferase that initiates bacterial lipoprotein biosynthesis.* Anal Biochem, 2012. **423**(1): p. 163-70.
118. Rath, A., et al., *Detergent binding explains anomalous SDS-PAGE migration of membrane proteins.* Proc Natl Acad Sci U S A, 2009. **106**(6): p. 1760-5.
119. Machata, S., et al., *Lipoproteins of Listeria monocytogenes are critical for virulence and TLR2-mediated immune activation.* J Immunol, 2008. **181**(3): p. 2028-35.
120. Pribyl, T., et al., *Influence of impaired lipoprotein biogenesis on surface and exoproteome of Streptococcus pneumoniae.* J Proteome Res, 2014. **13**(2): p. 650-67.
121. Hamilton, A., et al., *Mutation of the maturase lipoprotein attenuates the virulence of Streptococcus equi to a greater extent than does loss of general lipoprotein lipidation.* Infect Immun, 2006. **74**(12): p. 6907-19.
122. Mohammad, M., et al., *The role of Staphylococcus aureus lipoproteins in hematogenous septic arthritis.* Sci Rep, 2020. **10**(1): p. 7936.
123. Tokunaga, M., H. Tokunaga, and H.C. Wu, *Post-translational modification and processing of Escherichia coli prolipoprotein in vitro.* Proc Natl Acad Sci U S A, 1982. **79**(7): p. 2255-9.
124. Kurokawa, K., et al., *Environment-mediated accumulation of diacyl lipoproteins over their triacyl counterparts in Staphylococcus aureus.* J Bacteriol, 2012. **194**(13): p. 3299-306.
125. LoVullo, E.D., et al., *Revisiting the Gram-negative lipoprotein paradigm.* J Bacteriol, 2015. **197**(10): p. 1705-15.
126. da Silva, R.A.G., et al., *The role of apolipoprotein N-acyl transferase, Lnt, in the lipidation of factor H binding protein of Neisseria meningitidis strain MC58 and its potential as a drug target.* Br J Pharmacol, 2017. **174**(14): p. 2247-2260.
127. Gwin, C.M., et al., *The apolipoprotein N-acyl transferase Lnt is dispensable for growth in Acinetobacter species.* Microbiology, 2018. **164**(12): p. 1547-1556.
128. McClain, M.S., B.J. Voss, and T.L. Cover, *Lipoprotein Processing and Sorting in Helicobacter pylori.* mBio, 2020. **11**(3).
129. Foster, J., et al., *The Wolbachia genome of Brugia malayi: endosymbiont evolution within a human pathogenic nematode.* PLoS Biol, 2005. **3**(4): p. e121.
130. Shigenobu, S., et al., *Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS.* Nature, 2000. **407**(6800): p. 81-6.
131. Mohiman, N., et al., *The ppm operon is essential for acylation and glycosylation of lipoproteins in Corynebacterium glutamicum.* PLoS One, 2012. **7**(9): p. e46225.
132. Tschumi, A., et al., *Identification of apolipoprotein N-acyltransferase (Lnt) in mycobacteria.* J Biol Chem, 2009. **284**(40): p. 27146-56.
133. Hutchings, M.I., et al., *Lipoprotein biogenesis in Gram-positive bacteria: knowing when to hold 'em, knowing when to fold 'em.* Trends Microbiol, 2009. **17**(1): p. 13-21.
134. Kovacs-Simon, A., R.W. Titball, and S.L. Michell, *Lipoproteins of bacterial pathogens.* Infect. Immun., 2010. **79**(2): p. 548-561.
135. Kurokawa, K., et al., *Novel bacterial lipoprotein structures conserved in low-GC content gram-positive bacteria are recognized by Toll-like receptor 2.* J Biol Chem, 2012. **287**(16): p. 13170-81.
136. Asanuma, M., et al., *Structural evidence of alpha-aminoacylated lipoproteins of Staphylococcus aureus.* FEBS J, 2011. **278**(5): p. 716-28.

137. Kurokawa, K., et al., *The Triacylated ATP Binding Cluster Transporter Substrate-binding Lipoprotein of Staphylococcus aureus Functions as a Native Ligand for Toll-like Receptor 2.* J Biol Chem, 2009. **284**(13): p. 8406-11.
138. Hayashi, S. and H.C. Wu, *Accumulation of prolipoprotein in Escherichia coli mutants defective in protein secretion.* J Bacteriol, 1985. **161**(3): p. 949-54.
139. Chambaud, I., H. Wroblewski, and A. Blanchard, *Interactions between mycoplasma lipoproteins and the host immune system.* Trends Microbiol, 1999. **7**(12): p. 493-9.
140. Serebryakova, M.V., et al., *The acylation state of surface lipoproteins of mollicute Acholeplasma laidlawii.* J Biol Chem, 2011. **286**(26): p. 22769-76.
141. Gardiner, J.H., et al., *Lipoprotein N-acylation in Staphylococcus aureus is catalyzed by a two-component acyl transferase system.* mBio, 2020. **11**(4): p. e01619-20.
142. Nakanishi, H., et al., *Separation and quantification of sn-1 and sn-2 fatty acid positional isomers in phosphatidylcholine by RPLC-ESIMS/MS.* J Biochem, 2010. **147**(2): p. 245-56.
143. Armbruster, K.M., G. Komazin, and T.C. Meredith, *Bacterial lyso-form lipoproteins are synthesized via an intramolecular acyl chain migration.* J Biol Chem, 2020. **295**: p. 10195-10211.
144. Weyer, K.A., et al., *Amino acid sequence of the cytochrome subunit of the photosynthetic reaction centre from the purple bacterium Rhodopseudomonas viridis.* EMBO J, 1987. **6**(8): p. 2197-202.
145. Kulathila, R., et al., *Crystal structure of Escherichia coli CusC, the outer membrane component of a heavy metal efflux pump.* PLoS One, 2011. **6**(1): p. e15610.
146. Sun, C., et al., *Structure of the alternative complex III in a supercomplex with cytochrome oxidase.* Nature, 2018. **557**(7703): p. 123-126.
147. Ehrmann, M. and American Society for Microbiology., *The periplasm*. 2007, Washington, D.C.: ASM Press. xviii, 415 p.
148. Rangan, K.J., et al., *Rapid visualization and large-scale profiling of bacterial lipoproteins with chemical reporters.* J Am Chem Soc, 2010. **132**(31): p. 10628-9.
149. Liu, J., W. Cao, and M. Lu, *Core side-chain packing and backbone conformation in Lpp-56 coiled-coil mutants.* J Mol Biol, 2002. **318**(3): p. 877-88.
150. Shu, W., et al., *Core structure of the outer membrane lipoprotein from Escherichia coli at 1.9 A resolution.* J Mol Biol, 2000. **299**(4): p. 1101-12.
151. Li, G.W., et al., *Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources.* Cell, 2014. **157**(3): p. 624-35.
152. Inouye, M., J. Shaw, and C. Shen, *The assembly of a structural lipoprotein in the envelope of Escherichia coli.* J Biol Chem, 1972. **247**(24): p. 8154-9.
153. Cowles, C.E., et al., *The free and bound forms of Lpp occupy distinct subcellular locations in Escherichia coli.* Mol Microbiol, 2011. **79**(5): p. 1168-81.
154. Witwinowski, J., et al., *An ancient divide in outer membrane tethering systems in bacteria suggests a mechanism for the diderm-to-monoderm transition.* Nat Microbiol, 2022. **7**(3): p. 411-422.
155. Asmar, A.T., et al., *Communication across the bacterial cell envelope depends on the size of the periplasm.* PLoS Biol, 2017. **15**(12): p. e2004303.
156. Suzuki, H., et al., *Murein-lipoprotein of Escherichia coli: a protein involved in the stabilization of bacterial cell envelope.* Mol Gen Genet, 1978. **167**(1): p. 1-9.
157. Diao, J., et al., *Peptidoglycan Association of Murein Lipoprotein Is Required for KpsD-Dependent Group 2 Capsular Polysaccharide Expression and Serum Resistance in a Uropathogenic Escherichia coli Isolate.* mBio, 2017. **8**(3).
158. Sha, J., et al., *The two murein lipoproteins of Salmonella enterica serovar Typhimurium contribute to the virulence of the organism.* Infect Immun, 2004. **72**(7): p. 3987-4003.
159. Wu, T., et al., *Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli.* Cell, 2005. **121**(2): p. 235-45.

160. Malinverni, J.C., et al., *YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in Escherichia coli.* Mol. Microbiol., 2006. **61**(1): p. 151-64.
161. Chng, S.S., et al., *Characterization of the two-protein complex in Escherichia coli responsible for lipopolysaccharide assembly at the outer membrane.* Proc Natl Acad Sci U S A, 2010. **107**(12): p. 5363-8.
162. Wu, T., et al., *Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of Escherichia coli.* Proc. Nat. Acad. Sci. USA, 2006. **103**(31): p. 11754-11759.
163. Matsuyama, S., N. Yokota, and H. Tokuda, *A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of Escherichia coli.* EMBO J., 1997. **16**: p. 6947-6955.
164. Typas, A., et al., *Regulation of peptidoglycan synthesis by outer-membrane proteins.* Cell, 2010. **143**: p. 1097-1109.
165. Lupoli, T.J., et al., *Lipoprotein activators stimulate Escherichia coli penicillin-binding proteins by different mechanisms.* J Am Chem Soc, 2014. **136**(1): p. 52-5.
166. Banzhaf, M., et al., *Outer membrane lipoprotein NlpI scaffolds peptidoglycan hydrolases within multi-enzyme complexes in Escherichia coli.* EMBO J, 2020. **39**(5): p. e102246.
167. Tao, J., et al., *Heat shock proteins IbpA and IbpB are required for NlpI-participated cell division in Escherichia coli.* Front Microbiol, 2015. **6**: p. 51.
168. Schwechheimer, C., D.L. Rodriguez, and M.J. Kuehn, *NlpI-mediated modulation of outer membrane vesicle production through peptidoglycan dynamics in Escherichia coli.* Microbiologyopen, 2015. **4**(3): p. 375-89.
169. Singh, S.K., et al., *Regulated proteolysis of a cross-link-specific peptidoglycan hydrolase contributes to bacterial morphogenesis.* Proc Natl Acad Sci U S A, 2015. **112**(35): p. 10956-61.
170. Anantharaman, V. and L. Aravind, *Evolutionary history, structural features and biochemical diversity of the NlpC/P60 superfamily of enzymes.* Genome Biol, 2003. **4**(2): p. R11.
171. Lommatzsch, J., et al., *Outer membrane localization of murein hydrolases: MltA, a third lipoprotein lytic transglycosylase in Escherichia coli.* J Bacteriol, 1997. **179**: p. 5465-5470.
172. Dijkstra, A.J. and W. Keck, *Identification of new members of the lytic transglycosylase family in Haemophilus influenzae and Escherichia coli.* Microb Drug Resist, 1996. **2**: p. 141-145.
173. Fibriansah, G., F.I. Gliubich, and A.M. Thunnissen, *On the mechanism of peptidoglycan binding and cleavage by the endo-specific lytic transglycosylase MltE from Escherichia coli.* Biochemistry, 2012. **51**(45): p. 9164-77.
174. Ehlert, K., J.V. Holtje, and M.F. Templin, *Cloning and expression of a murein hydrolase lipoprotein from Escherichia coli.* Mol Microbiol, 1995. **16**(4): p. 761-8.
175. Gerding, M.A., et al., *The trans-envelope Tol-Pal complex is part of the cell division machinery and required for proper outer-membrane invagination during cell constriction in E. coli.* Mol Microbiol, 2007. **63**(4): p. 1008-25.
176. Petiti, M., et al., *Tol Energy-Driven Localization of Pal and Anchoring to the Peptidoglycan Promote Outer-Membrane Constriction.* J Mol Biol, 2019. **431**(17): p. 3275-3288.
177. Yakhnina, A.A. and T.G. Bernhardt, *The Tol-Pal system is required for peptidoglycan-cleaving enzymes to complete bacterial cell division.* Proc Natl Acad Sci U S A, 2020. **117**(12): p. 6777-6783.
178. Uehara, T., T. Dinh, and T.G. Bernhardt, *LytM-domain factors are required for daughter cell separation and rapid ampicillin-induced lysis in Escherichia coli.* Journal of Bacteriology, 2009. **191**(16): p. 5094-5107.
179. Garcia, D.L. and J.P. Dillard, *AmiC functions as an N-acetylmuramyl-l-alanine amidase necessary for cell separation and can promote autolysis in Neisseria gonorrhoeae.* J Bacteriol, 2006. **188**(20): p. 7211-21.
180. Heidrich, C., et al., *Involvement of N-acetylmuramyl-L-alanine amidases in cell separation and antibiotic-induced autolysis of Escherichia coli.* Mol Microbiol, 2001. **41**(1): p. 167-78.

181. Kerff, F., et al., *Specific structural features of the N-acetylmuramoyl-L-alanine amidase AmiD from Escherichia coli and mechanistic implications for enzymes of this family.* J Mol Biol, 2010. **397**(1): p. 249-59.
182. Lange, R. and R. Hengge-Aronis, *The nlpD gene is located in an operon with rpoS on the Escherichia coli chromosome and encodes a novel lipoprotein with a potential function in cell wall formation.* Mol Microbiol, 1994. **13**(4): p. 733-43.
183. Berezuk, A.M., et al., *Outer membrane lipoprotein RlpA is a novel periplasmic interaction partner of the cell division protein FtsK in Escherichia coli.* Sci Rep, 2018. **8**(1): p. 12933.
184. Jorgenson, M.A., et al., *The bacterial septal ring protein RlpA is a lytic transglycosylase that contributes to rod shape and daughter cell separation in Pseudomonas aeruginosa.* Mol Microbiol, 2014. **93**(1): p. 113-28.
185. Arends, S.J., et al., *Discovery and characterization of three new Escherichia coli septal ring proteins that contain a SPOR domain: DamX, DedD, and RlpA.* J Bacteriol, 2010. **192**(1): p. 242-55.
186. Plesa, M., et al., *The SlyB outer membrane lipoprotein of Burkholderia multivorans contributes to membrane integrity.* Res Microbiol, 2006. **157**(6): p. 582-92.
187. Queiroz, P.A., et al., *Proteomic profiling of Klebsiella pneumoniae carbapenemase (KPC)-producer Klebsiella pneumoniae after induced polymyxin resistance.* Future Microbiol, 2021. **16**: p. 1195-1207.
188. Yeow, J., et al., *The architecture of the OmpC-MlaA complex sheds light on the maintenance of outer membrane lipid asymmetry in Escherichia coli.* J Biol Chem, 2018. **293**(29): p. 11325-11340.
189. Hughes, G.W., et al., *Evidence for phospholipid export from the bacterial inner membrane by the Mla ABC transport system.* Nat Microbiol, 2019. **4**(10): p. 1692-1705.
190. Nakayama, T. and Q.M. Zhang-Akiyama, *pqiABC and yebST, Putative mce Operons of Escherichia coli, Encode Transport Pathways and Contribute to Membrane Integrity.* J Bacteriol, 2017. **199**(1).
191. Goyal, P., et al., *Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG.* Nature, 2014. **516**(7530): p. 250-3.
192. Zhang, M., et al., *Cryo-EM structure of the nonameric CsgG-CsgF complex and its implications for controlling curli biogenesis in Enterobacteriaceae.* PLoS Biol, 2020. **18**(6): p. e3000748.
193. Loferer, H., M. Hammar, and S. Normark, *Availability of the fibre subunit CsgA and the nucleator protein CsgB during assembly of fibronectin-binding curli is limited by the intracellular concentration of the novel lipoprotein CsgG.* Mol Microbiol, 1997. **26**(1): p. 11-23.
194. Rouf, S.F., et al., *Opposing contributions of polynucleotide phosphorylase and the membrane protein NlpI to biofilm formation by Salmonella enterica serovar Typhimurium.* J Bacteriol, 2011. **193**(2): p. 580-2.
195. Beis, K., et al., *Crystallization and preliminary X-ray diffraction analysis of Wza outer-membrane lipoprotein from Escherichia coli serotype O9a:K30.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 3): p. 558-60.
196. Dong, C., et al., *Wza the translocon for E. coli capsular polysaccharides defines a new class of membrane protein.* Nature, 2006. **444**(7116): p. 226-9.
197. Larson, M.R., et al., *Escherichia coli O127 group 4 capsule proteins assemble at the outer membrane.* PLoS One, 2021. **16**(11): p. e0259900.
198. Peleg, A., et al., *Identification of an Escherichia coli operon required for formation of the O-antigen capsule.* J Bacteriol, 2005. **187**(15): p. 5259-66.
199. Granato, L.M., et al., *The ecnA Antitoxin Is Important Not Only for Human Pathogens: Evidence of Its Role in the Plant Pathogen Xanthomonas citri subsp. citri.* J Bacteriol, 2019. **201**(20).
200. Weber, M.M., et al., *A previously uncharacterized gene, yjfO (bsmA), influences Escherichia coli biofilm formation and stress response.* Microbiology (Reading), 2010. **156**(Pt 1): p. 139-147.

201. Matsumura, K., et al., *Roles of multidrug efflux pumps on the biofilm formation of Escherichia coli K-12.* Biocontrol Sci, 2011. **16**(2): p. 69-72.
202. Schoenhals, G.J. and R.M. Macnab, *Physiological and biochemical analyses of FlgH, a lipoprotein forming the outer membrane L ring of the flagellar basal body of Salmonella typhimurium.* J Bacteriol, 1996. **178**(14): p. 4200-7.
203. Yamaguchi, T., et al., *Structure of the molecular bushing of the bacterial flagellar motor.* Nat Commun, 2021. **12**(1): p. 4469.
204. Dailey, F.E. and R.M. Macnab, *Effects of lipoprotein biogenesis mutations on flagellar assembly in Salmonella.* J Bacteriol, 2002. **184**(3): p. 771-6.
205. Colclough, A.L., et al., *RND efflux pumps in Gram-negative bacteria; regulation, structure and role in antibiotic resistance.* Future Microbiol, 2020. **15**: p. 143-157.
206. Smith, H.E. and J.M. Blair, *Redundancy in the periplasmic adaptor proteins AcrA and AcrE provides resilience and an ability to export substrates of multidrug efflux.* J Antimicrob Chemother, 2014. **69**(4): p. 982-7.
207. Jeong, H., et al., *Pseudoatomic Structure of the Tripartite Multidrug Efflux Pump AcrAB-TolC Reveals the Intermeshing Cogwheel-like Interaction between AcrA and TolC.* Structure, 2016. **24**(2): p. 272-6.
208. Gumbart, J.C., et al., *Lpp positions peptidoglycan at the AcrA-TolC interface in the AcrAB-TolC multidrug efflux pump.* Biophys J, 2021. **120**(18): p. 3973-3982.
209. Alav, I., V.N. Bavro, and J.M.A. Blair, *A role for the periplasmic adaptor protein AcrA in vetting substrate access to the RND efflux transporter AcrB.* Sci Rep, 2022. **12**(1): p. 4752.
210. Lei, H.T., et al., *Crystal structures of CusC review conformational changes accompanying folding and transmembrane channel formation.* J Mol Biol, 2014. **426**(2): p. 403-11.
211. Semchenko, E.A., C.J. Day, and K.L. Seib, *MetQ of Neisseria gonorrhoeae Is a Surface-Expressed Antigen That Elicits Bactericidal and Functional Blocking Antibodies.* Infect Immun, 2017. **85**(2).
212. Sharaf, N.G., et al., *Characterization of the ABC methionine transporter from Neisseria meningitidis reveals that lipidated MetQ is required for interaction.* Elife, 2021. **10**.
213. Pei, X., et al., *Quantitative proteomics revealed modulation of macrophages by MetQ gene of Streptococcus suis serotype 2.* AMB Express, 2020. **10**(1): p. 195.
214. Fedorchuk, C., I.T. Kudva, and S. Kariyawasam, *The Escherichia coli O157:H7 carbon starvation-inducible lipoprotein Slp contributes to initial adherence in vitro via the human polymeric immunoglobulin receptor.* PLoS One, 2019. **14**(6): p. e0216791.
215. Manning, P.A., L. Beutin, and M. Achtman, *Outer membrane of Escherichia coli: properties of the F sex factor traT protein which is involved in surface exclusion.* J Bacteriol, 1980. **142**(1): p. 285-94.
216. Perumal, N.B. and E.G. Minkley, Jr., *The product of the F sex factor traT surface exclusion gene is a lipoprotein.* J Biol Chem, 1984. **259**(9): p. 5357-60.
217. Aguero, M.E., et al., *A plasmid-encoded outer membrane protein, TraT, enhances resistance of Escherichia coli to phagocytosis.* Infect Immun, 1984. **46**(3): p. 740-6.
218. Binns, M.M., J. Mayden, and R.P. Levine, *Further characterization of complement resistance conferred on Escherichia coli by the plasmid genes traT of R100 and iss of ColV,I-K94.* Infect Immun, 1982. **35**(2): p. 654-9.
219. Moll, A., P.A. Manning, and K.N. Timmis, *Plasmid-determined resistance to serum bactericidal activity: a major outer membrane protein, the traT gene product, is responsible for plasmid-specified serum resistance in Escherichia coli.* Infect Immun, 1980. **28**(2): p. 359-67.
220. Tseng, Y.T., et al., *NlpI facilitates deposition of C4bp on Escherichia coli by blocking classical complement-mediated killing, which results in high-level bacteremia.* Infect Immun, 2012. **80**(10): p. 3669-78.
221. Nasu, H., et al., *Knockout of mlaA increases Escherichia coli virulence in a silkworm infection model.* PLoS One, 2022. **17**(7): p. e0270166.

222. Callewaert, L., et al., *A new family of lysozyme inhibitors contributing to lysozyme tolerance in gram-negative bacteria.* PLoS Pathog, 2008. **4**(3): p. e1000019.
223. Yum, S., et al., *Structural basis for the recognition of lysozyme by MliC, a periplasmic lysozyme inhibitor in Gram-negative bacteria.* Biochem Biophys Res Commun, 2009. **378**(2): p. 244-8.
224. Merlin, C., et al., *The Escherichia coli metD locus encodes an ABC transporter which includes Abc (MetN), YaeE (MetI), and YaeC (MetQ).* J Bacteriol, 2002. **184**(19): p. 5513-7.
225. Rahman, M.M., M.A. Machuca, and A. Roujeinikova, *Bioinformatics analysis and biochemical characterisation of ABC transporter-associated periplasmic substrate-binding proteins ModA and MetQ from Helicobacter pylori strain SS1.* Biophys Chem, 2021. **272**: p. 106577.
226. Beck, B.J. and D.M. Downs, *The apbE gene encodes a lipoprotein involved in thiamine synthesis in Salmonella typhimurium.* J Bacteriol, 1998. **180**(4): p. 885-91.
227. Beck, B.J., et al., *Evidence that rseC, a gene in the rpoE cluster, has a role in thiamine synthesis in Salmonella typhimurium.* J Bacteriol, 1997. **179**(20): p. 6504-8.
228. Bertsova, Y.V., et al., *Mutational analysis of the flavinylation and binding motifs in two protein targets of the flavin transferase ApbE.* FEMS Microbiol Lett, 2019. **366**(22).
229. Meheust, R., et al., *Post-translational flavinylation is associated with diverse extracytosolic redox functionalities throughout bacterial life.* Elife, 2021. **10**.
230. Cho, S.H., et al., *Detecting envelope stress by monitoring beta-barrel assembly.* Cell, 2014. **159**(7): p. 1652-64.
231. Laloux, G. and J.F. Collet, *Major Tom to Ground Control: How Lipoproteins Communicate Extracytoplasmic Stress to the Decision Center of the Cell.* J Bacteriol, 2017. **199**(21).
232. Snyder, W.B., et al., *Overproduction of NlpE, a new outer membrane lipoprotein, suppresses the toxicity of periplasmic LacZ by activation of the Cpx signal transduction pathway.* J Bacteriol, 1995. **177**(15): p. 4216-23.
233. Delhaye, A., G. Laloux, and J.F. Collet, *The Lipoprotein NlpE Is a Cpx Sensor That Serves as a Sentinel for Protein Sorting and Folding Defects in the Escherichia coli Envelope.* J Bacteriol, 2019. **201**(10).
234. Alexander, D.M. and A.C. St John, *Characterization of the carbon starvation-inducible and stationary phase-inducible gene slp encoding an outer membrane lipoprotein in Escherichia coli.* Mol Microbiol, 1994. **11**(6): p. 1059-71.
235. Schiefner, A., et al., *Structural and biochemical analyses reveal a monomeric state of the bacterial lipocalin Blc.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 12): p. 1308-15.
236. Campanacci, V., et al., *The membrane bound bacterial lipocalin Blc is a functional dimer with binding preference for lysophospholipids.* FEBS Lett, 2006. **580**(20): p. 4877-83.
237. Lutticke, C., et al., *E. coli LoiP (YggG), a metalloprotease hydrolyzing Phe-Phe bonds.* Mol Biosyst, 2012. **8**(6): p. 1775-82.
238. Boulanger, A., et al., *Multistress regulation in Escherichia coli: expression of osmB involves two independent promoters responding either to sigmaS or to the RcsCDB His-Asp phosphorelay.* J Bacteriol, 2005. **187**(9): p. 3282-6.
239. Gutierrez, C., S. Gordia, and S. Bonnassie, *Characterization of the osmotically inducible gene osmE of Escherichia coli K-12.* Mol Microbiol, 1995. **16**(3): p. 553-63.
240. Oberto, J., *SyntTax: a web server linking synteny to prokaryotic taxonomy.* BMC Bioinformatics, 2013. **14**: p. 4.
241. Janda, J.M. and S.L. Abbott, *16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls.* J Clin Microbiol, 2007. **45**(9): p. 2761-4.
242. Yang, Z. and B. Rannala, *Molecular phylogenetics: principles and practice.* Nat Rev Genet, 2012. **13**(5): p. 303-14.
243. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.
244. Mirdita, M., et al., *ColabFold: making protein folding accessible to all.* Nat Methods, 2022. **19**(6): p. 679-682.

245. Smithers, L., S. Olatunji, and M. Caffrey, *Bacterial Lipoprotein Posttranslational Modifications. New Insights and Opportunities for Antibiotic and Vaccine Development.* Front Microbiol, 2021. **12**: p. 788445.
246. Tubiana, J., D. Schneidman-Duhovny, and H.J. Wolfson, *ScanNet: an interpretable geometric deep learning model for structure-based protein binding site prediction.* Nat Methods, 2022. **19**(6): p. 730-739.
247. Romero, P.R. and P.D. Karp, *Using functional and organizational information to improve genome-wide computational prediction of transcription units on pathway-genome databases.* Bioinformatics, 2004. **20**(5): p. 709-17.
248. Price, M.N., et al., *A novel method for accurate operon predictions in all sequenced prokaryotes.* Nucleic Acids Res, 2005. **33**(3): p. 880-92.
249. Seemann, T., *Prokka: rapid prokaryotic genome annotation.* Bioinformatics, 2014. **30**(14): p. 2068-9.
250. Legood, S., et al., *A Defect in Lipoprotein Modification by Lgt Leads to Abnormal Morphology and Cell Death in Escherichia coli That Is Independent of Major Lipoprotein Lpp.* J Bacteriol, 2022: p. e0016422.
251. Binet, R. and C. Wandersman, *Protein secretion by hybrid bacterial ABC-transporters: specific functions of the membrane ATPase and the membrane fusion protein.* EMBO J., 1995. **14**: p. 2298-2306.
252. Nozeret, K., et al., *A sensitive fluorescence-based assay to monitor enzymatic activity of the essential integral membrane protein Apolipoprotein N-acyltransferase (Lnt).* Sci Rep, 2019. **9**(1): p. 15978.
253. Qin, X., et al., *Polarity-based fluorescence probes: properties and applications.* RSC Med Chem, 2021. **12**(11): p. 1826-1838.
254. Nozeret, K., A. Pernin, and N. Buddelmeijer, *Click-chemistry based fluorometric assay for Apolipoprotein N-acyltransferase from enzyme characterization to high-throughput screening.* JoVE, 2020. **159**: p. e61146.
255. De Cesare, V., et al., *High-throughput matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry-based deubiquitylating enzyme assay for drug discovery.* Nat Protoc, 2020. **15**(12): p. 4034-4057.
256. Nakajima, M., et al., *Globomycin, a new peptide antibiotic with spheroplast-forming activity. III. Structural determination of globomycin.* J Antibiot (Tokyo), 1978. **31**(5): p. 426-32.

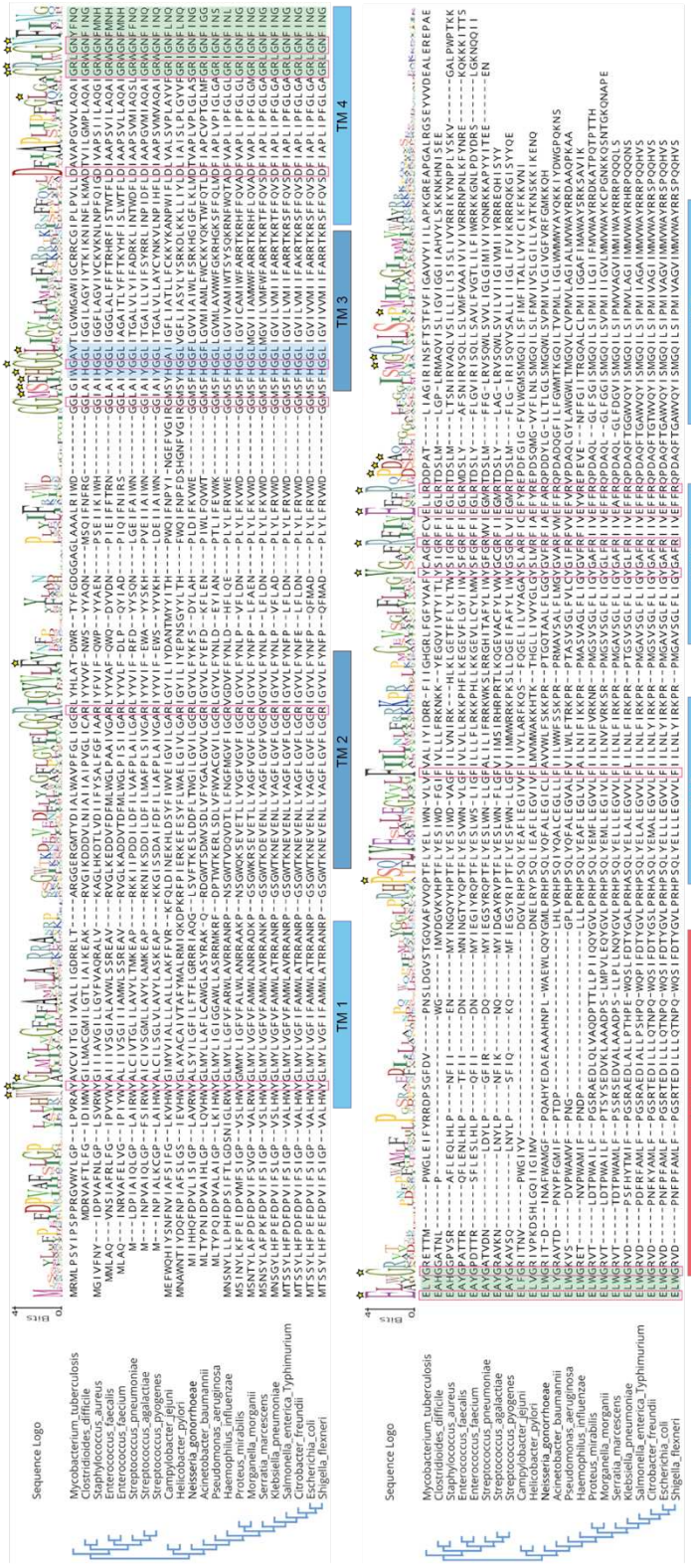
Appendices

Appendix I: Strains and their taxonomy

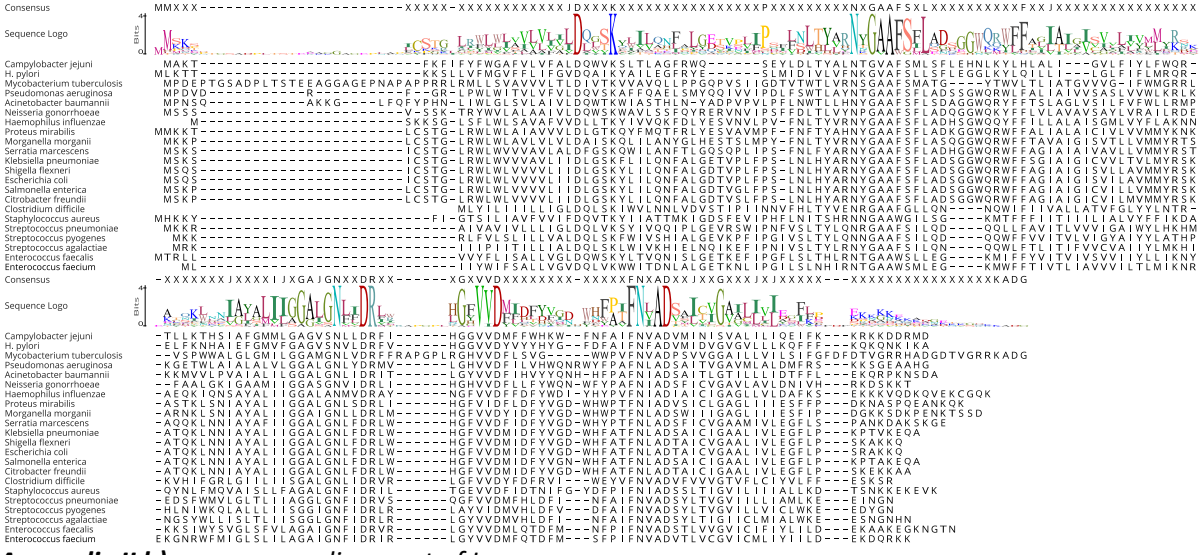
Kingdom	Phylum	Class	Order	Family	Genus/species	Short name
Bacteria	Actinobacteria	Actinomycetia	Corynebacteriales	Corynebacteriaceae	*Corynebacterium glutamicum*	*C. glutamicum*
				Mycobacteriaceae	*Mycobacterium smegmatis*	*M. smegmatis*
					Mycobacterium tuberculosis	*M. tuberculosis*
					Mycobacterium tuberculosis v. bovis	*M. bovis*
			Streptomycetales	Streptomycetaceae	*Streptomyces coelicolor*	*S. coelicolor*
					Streptomyces scabies	*S. scabies*
	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Bacillus cereus*	*B. cereus*
					Bacillus subtlis	*B. subtlis*
				Listeriaceae	*Listeria monocytogenes*	*L. monocytogenes*
				Staphylococcus aureus	*Staphylococcus aureus*	*S. aureus*
			Lactobacillales	Aerococcaceae	*Aerococcus spp.*	
				Enterococcaceae	*Enterococcus faecalis*	*E. faecalis*
					Enterococcus faecium	*E. faecium*
				Lactobacillaceae	*Lactobacillus spp.*	
					Pediococcus spp.	
				Streptococcaceae	*Lactococcus lactis*	*L. lactis*
					Streptococcus agalactiae	*S. agalactiae*
					Streptococcus equi	*S. equi*
					Streptococcus mutans	*S. mutans*
					Streptococcus pneumoniae	*S. pneumoniae*
					Streptococcus pyogenes	*S. pyogenes*
					Streptococcus uberis	*S. uberis*
			Lactobacilli	Carnobacteriaceae	*Carnobacterium spp.*	
		Clostridia	Eubacteriales	Clostridiaceae	*Clostridium perfringens*	*C. perfringens*
				Peptostreptococcaceae	*Clostridioides difficile*	*C. difficile*
	Proteobacteria	Alphaproteobacteria	Hyphomicrobiales	Blastochloridaceae	*Blastochloris viridis*	*B. viridis*
			Rickettsiales	Anaplasmataceae	*Wolbachia spp.*	
		betaproteobacteria	Burkholderiales	Alcaligenaceae	*Bordetella pertussis*	*B. pertussis*
				Burkholderiaceae	*Burkholderia multivorans*	*B. multivorans*
			Neisseriales	Neisseriaceae	*Neisseria gonorrhoeae*	*N. gonorrhoeae*
					Neisseria menigitidis	*N. menigitidis*
		Deltaproteobacteria	Myxococcales	Myxococcaceae	*Myxococcus xanthus*	*M. xanthus*
		Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	*Campylobacter jejuni*	*C. jejuni*
				Helicobacteraceae	*Helicobacter pylori*	*H. pylori*
		gammaproteobacteria	Enterobacterales	Enterobacteriaceae	*Citrobacter freundii*	*C. freundii*
					Escherichia coli	*E. coli*
					Klebsiella pneumoniae	*K. pneumoniae*
					Salmonella enterica s. Typhimurium	*S.enterica serovar Typhimurium*
					Shigella flexneri	*S. flexneri*
				Erwiniaceae	*Buchnera spp.*	
				Morganellaceae	*Morganella morganii*	*M. morganii*
					Proteus mirabilis	*P. mirabilis*
				Yersiniaceae	*Serratia marcescens*	*S. marcescens*
			Moraxellales	Moraxellaceae	*Acinetobacter baumannii*	*A. baumannii*
			Pasteurellales	Pasteurellaceae	*Haemophilus influenzae*	*H. influenzae*
			Pseudomonadales	Pseudomonadaceae	*Pseudomonas aeruginosa*	*P. aeruginosa*
			Thiotrichales	Francisellaceae	*Francisella tularensis*	*F. tularensis*
	Tenericutes	Mollicutes	Acholeplasmatales	Acholeplasmataceae	*Acholeplasma laidlawii*	*A. laidlawii*
			Mycoplasmatales	Mycoplasmataceae	*Mycoplasma gentalium*	*M. gentalium*
	Thermotogae	Thermotogae	Thermotogales	Thermotogaceae	*Thermotoga maritima*	*T. maritima*
Eukaryota	Cercozoa	Imbricatea	Euglyphida	Paulinellidae	*Paulinella chromatophora*	*P. chromattophora*

Appendix I *– Strains and taxonomy of bacteria mentioned in this thesis*

Appendix II: Sequence alignments of Lgt and Lsp

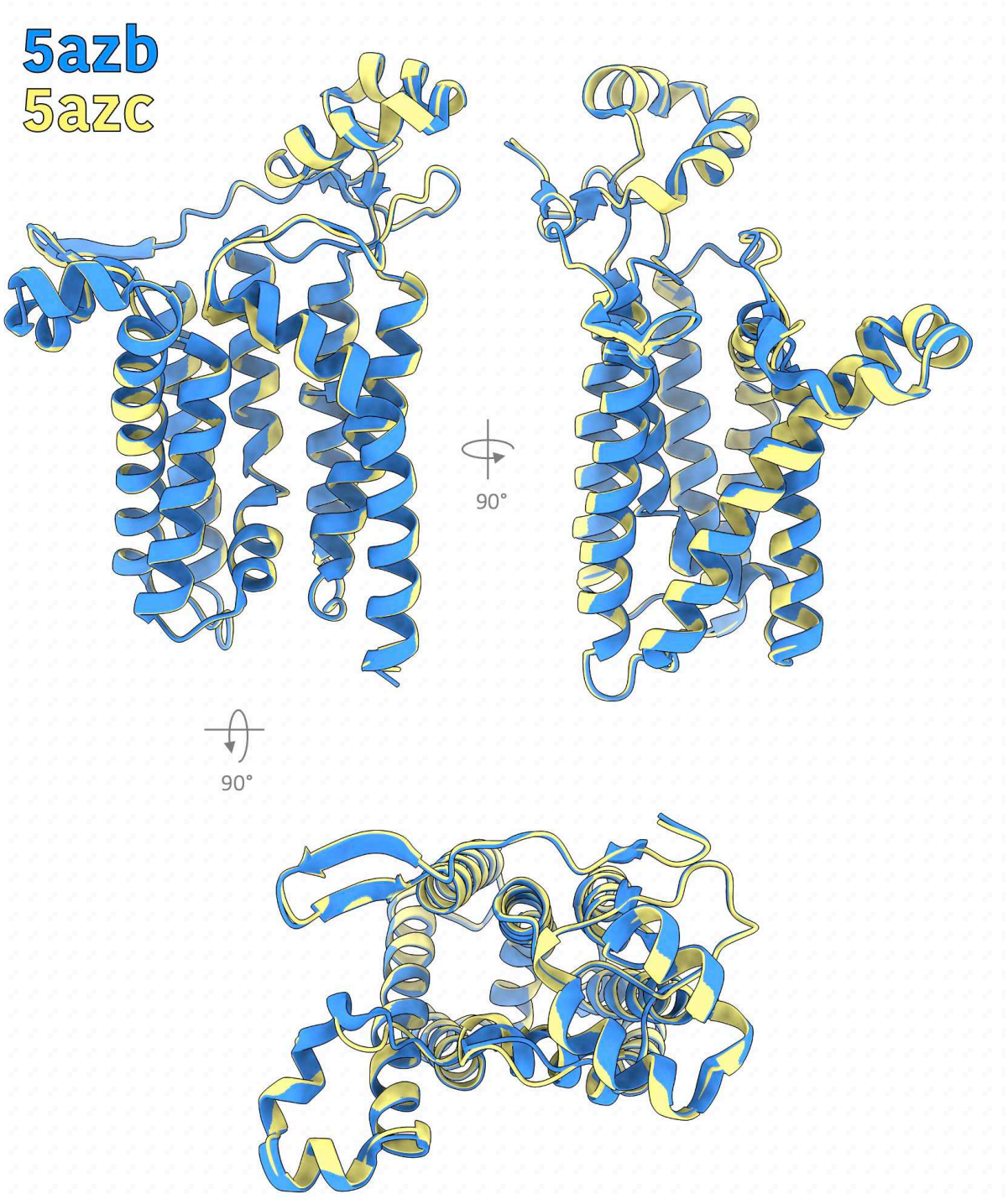


Appendix II a) – *sequence alignment of Lgt. Stars indicated residues described as essential (Table 5). Blue highlighting = HGGL motif, green highlighting = Lgt signature motif, Pink squares = fully conserved residues. TM = transmembrane helix.*



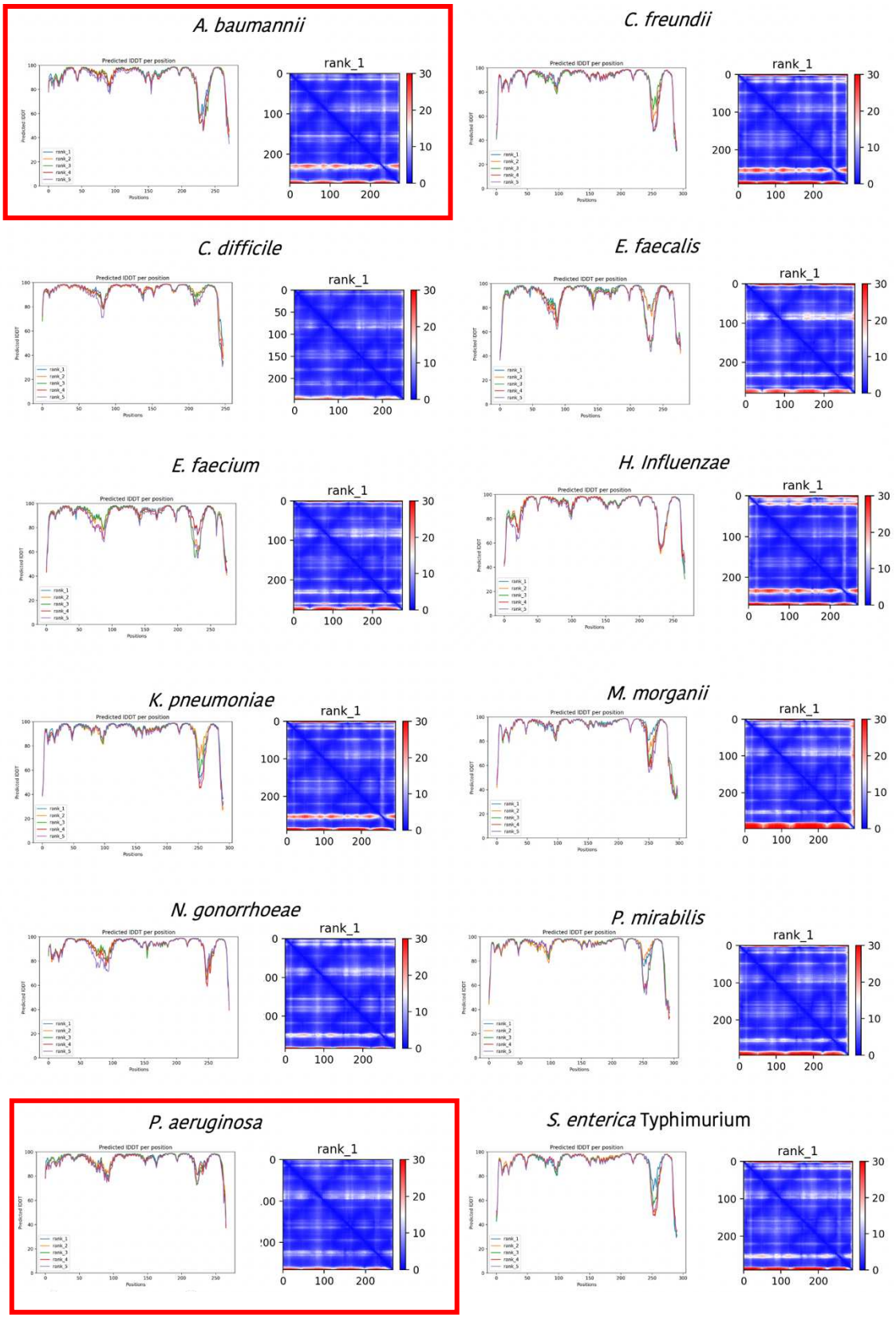
Appendix II b) *– sequence alignment of Lsp*

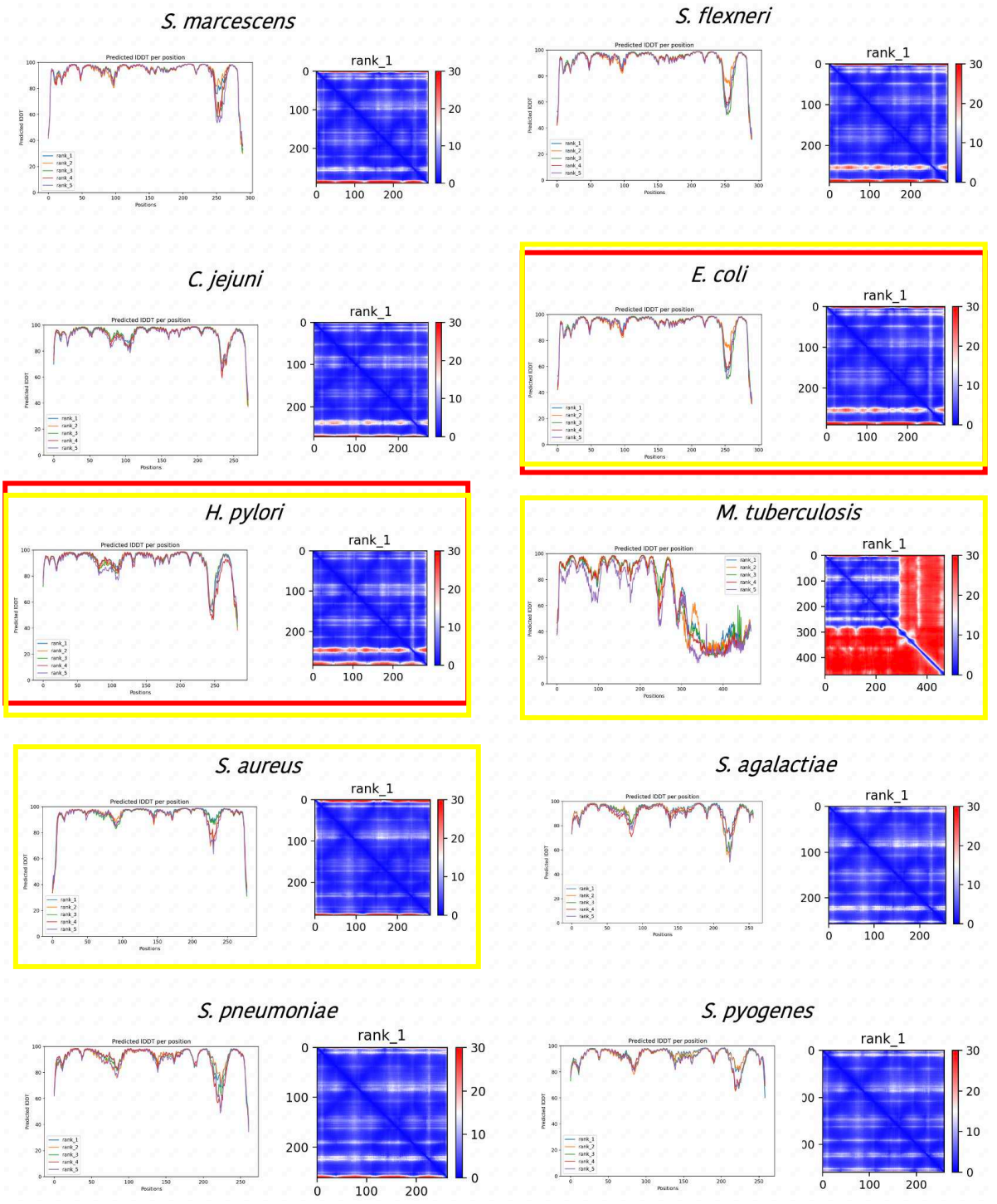
Appendix III: Lgt X-ray crystal structures



Appendix III *Comparison of two X-ray crystal structures of Lgt. Form-1 = 5azb, form-2 = 5azc [81]*

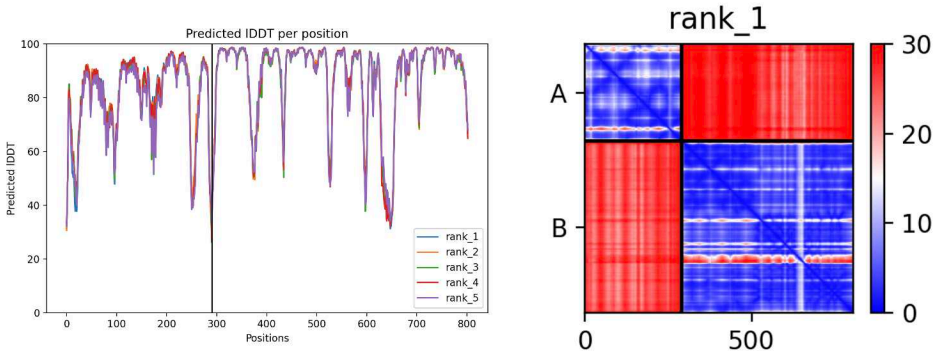
Appendix IV: AlphaFold2 confidence metrics



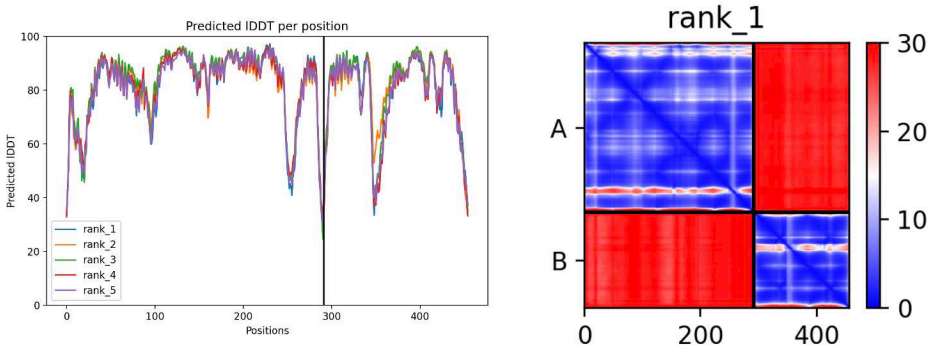


Appendix IV a) *- AlphaFold2 confidence metrics for Lgt predicted structures. Left-hand panel = plDDTs, right-hand panel = PAE. Red squares = strains used in complementation analysis, yellow squares = strains used in head domain swapping*

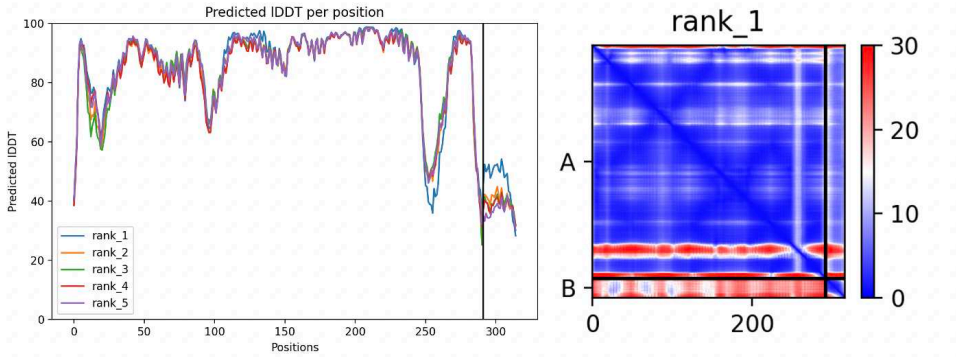
Lgt + Lnt



Lgt + Lsp

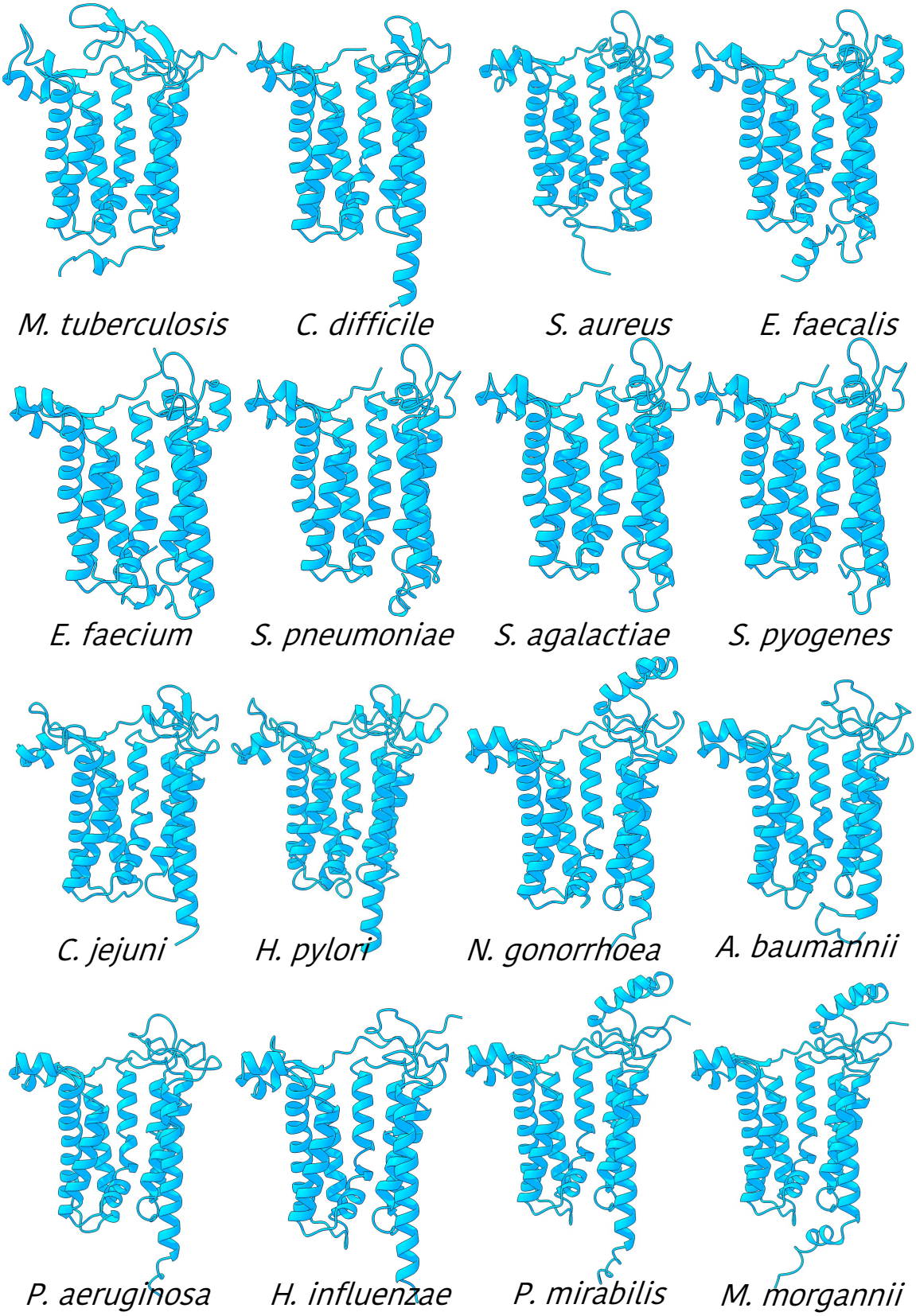


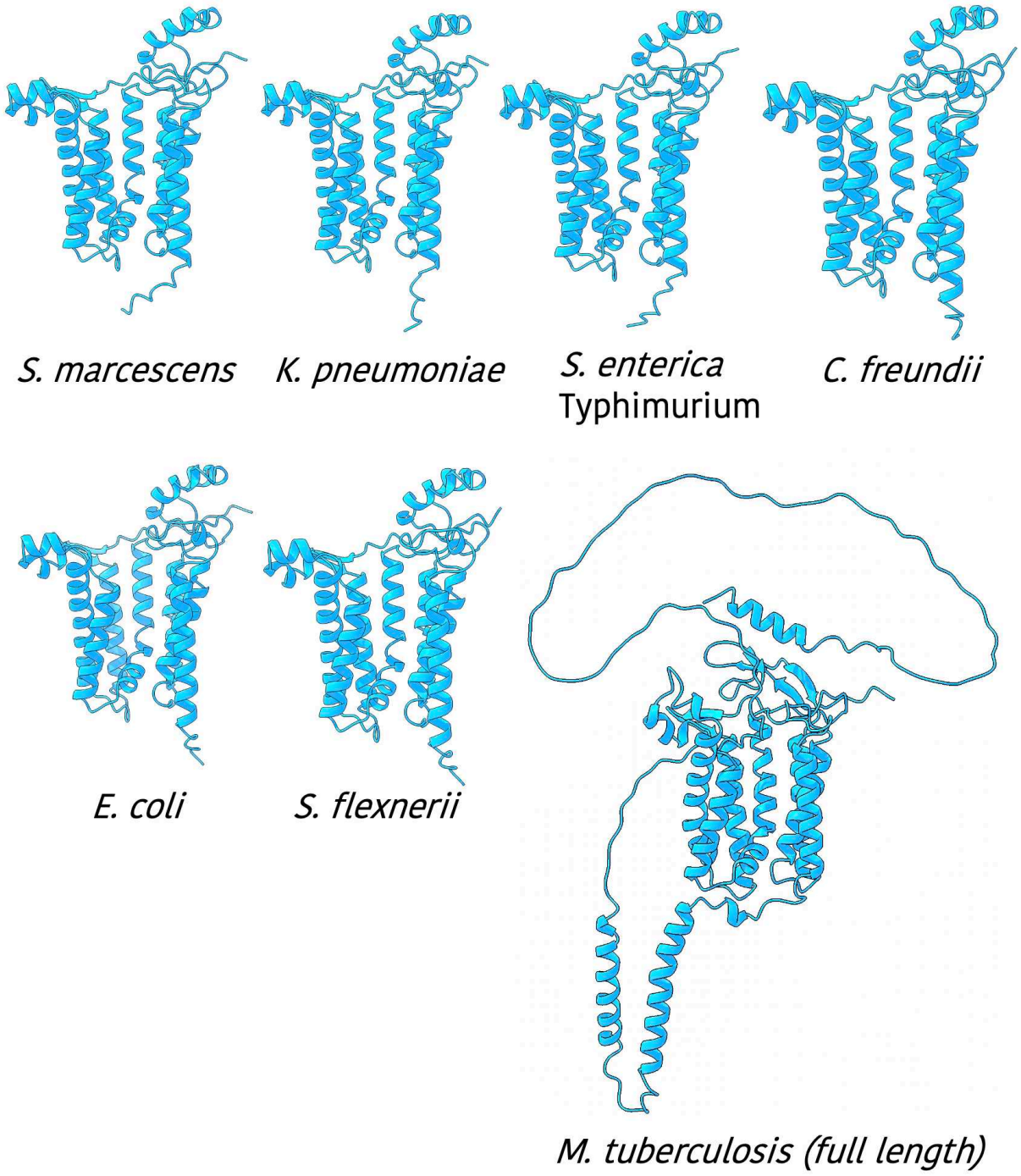
Appendix IV b) *- ColabFold and AlphaFold2 confidence metrics for Lgt and Lsp/Lnt protein interactions Left-hand panel = plDDTs, right-hand panel = PAE*



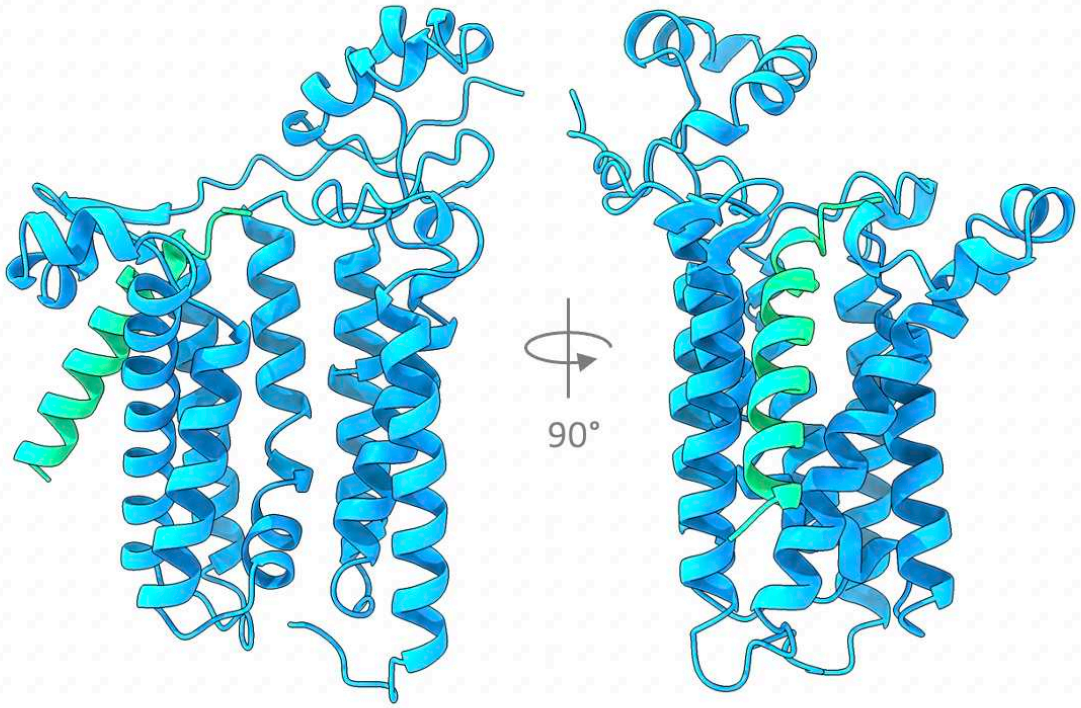
Appendix IV c) *- ColabFold and AlphaFold2 confidence metrics for Lgt and Lpp protein interactions Left-hand panel = plDDTs, right-hand panel = PAE*

Appendix V: AlphaFold2 structures of Lgt



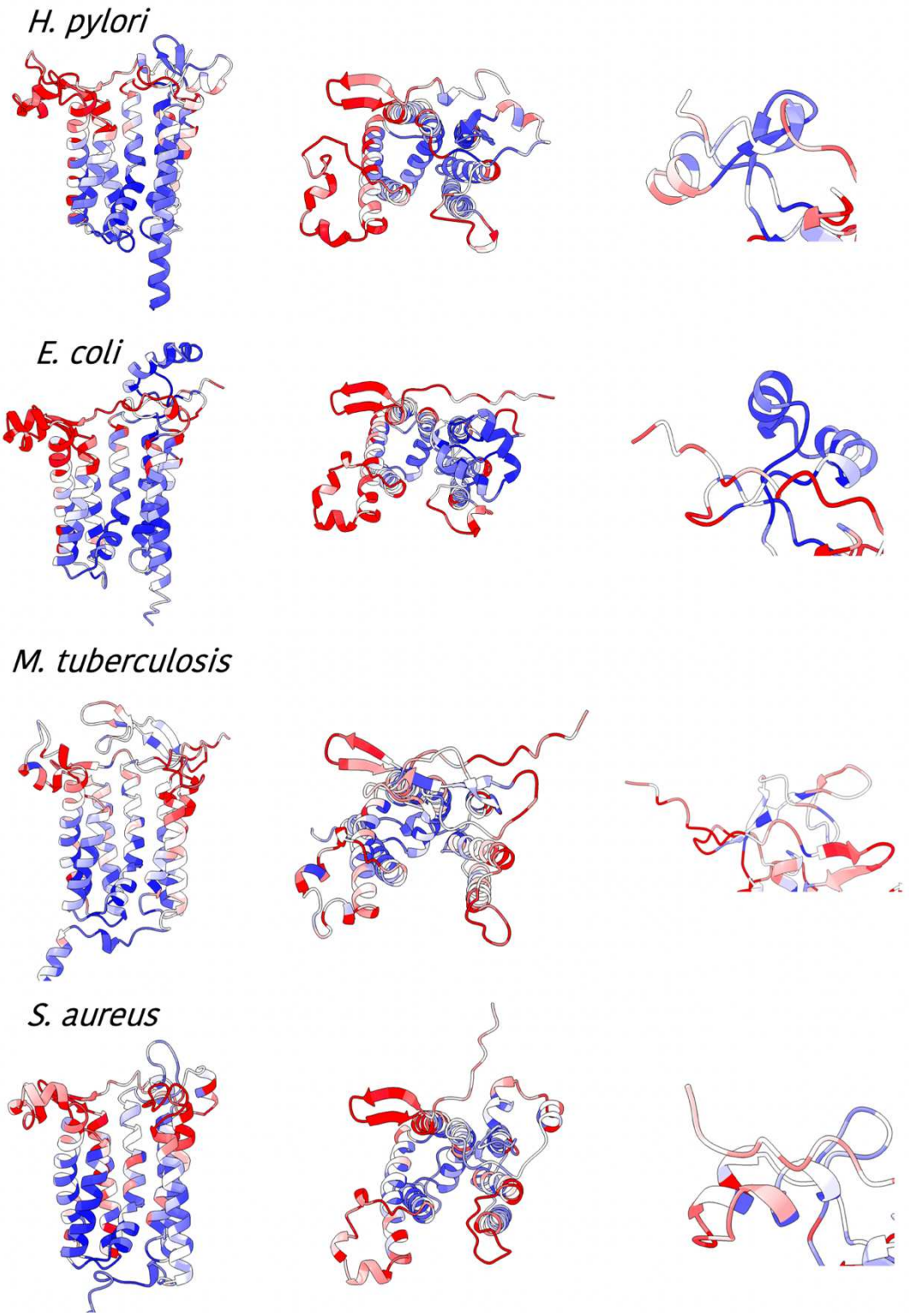


Appendix V a) **-** *AlphaFold2 structures of Lgt*



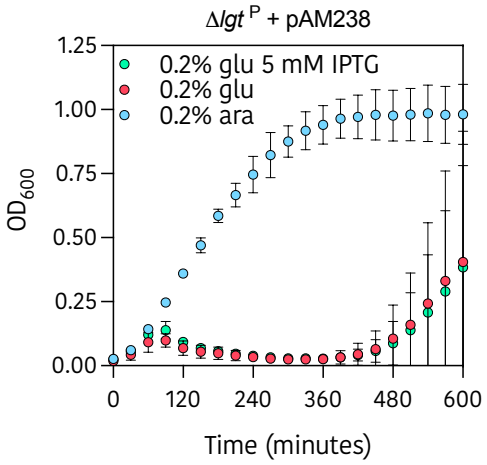
Appendix V b) **-** *ColabFold structures and interaction of Lgt (blue) and the signal peptide of Lpp (green)*

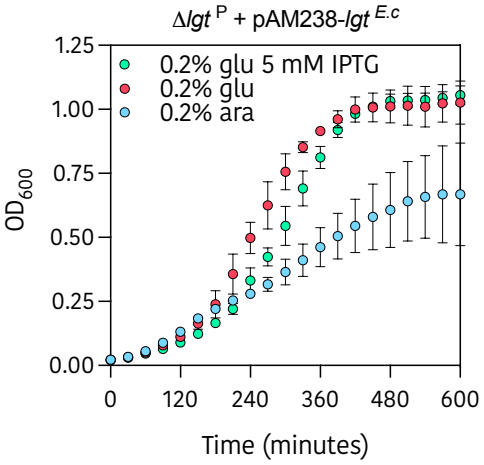
Appendix VI: ScanNet preidicted protein-protein binding sites

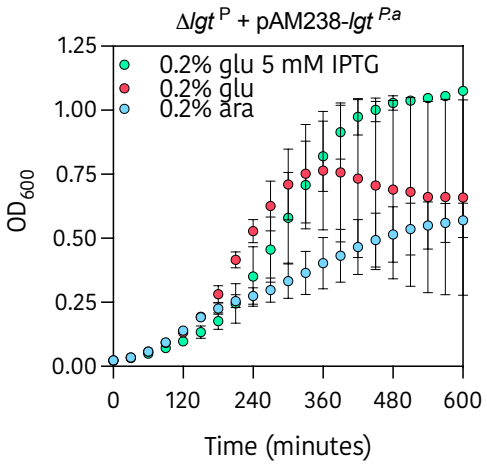


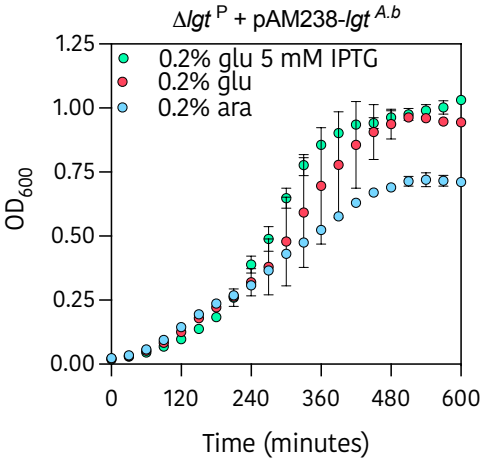
Appendix VI - *ScanNet protein-protein binding probability (blue = low probability, red = high probability), left-hand panel = side view, central panel = top view, and right-hand panel = head domains with N-terminal sequence visible*

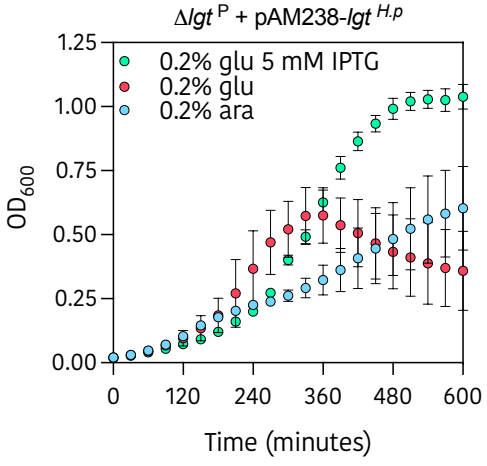
Appendix VII: Growth kinetics

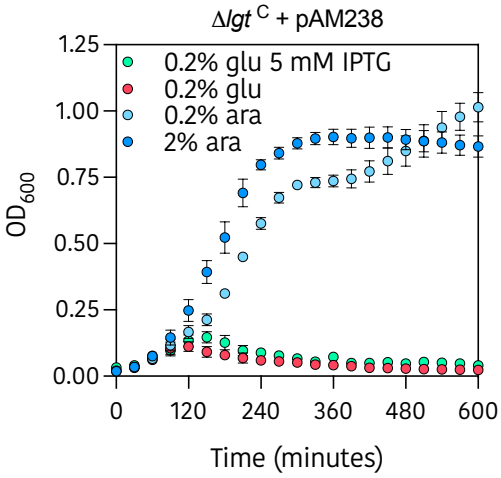


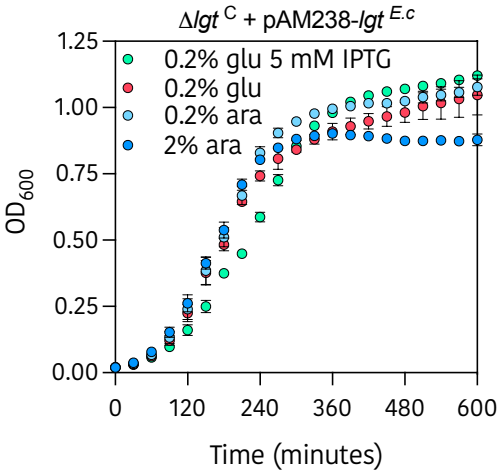


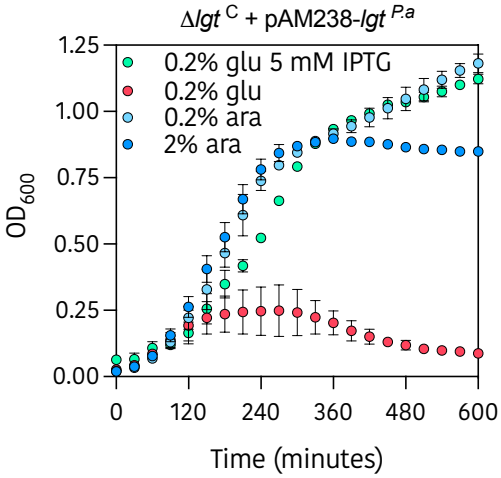


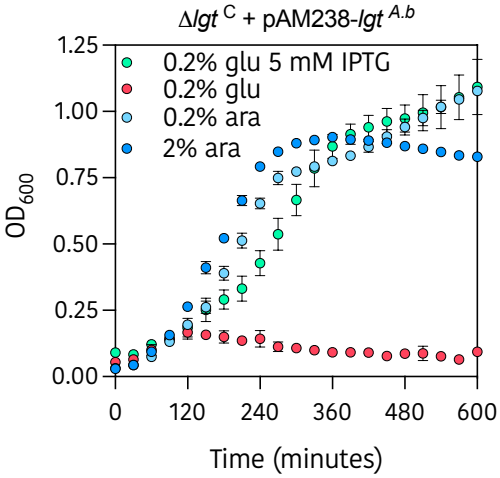


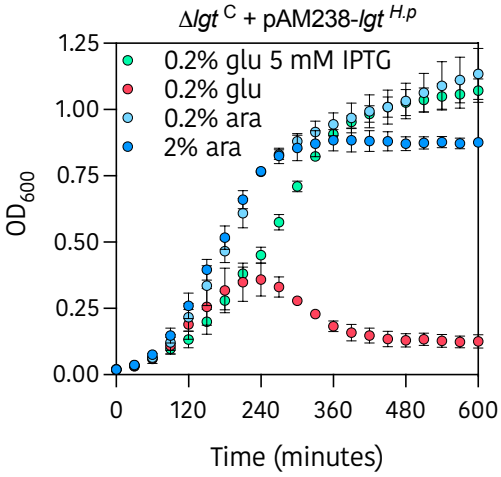


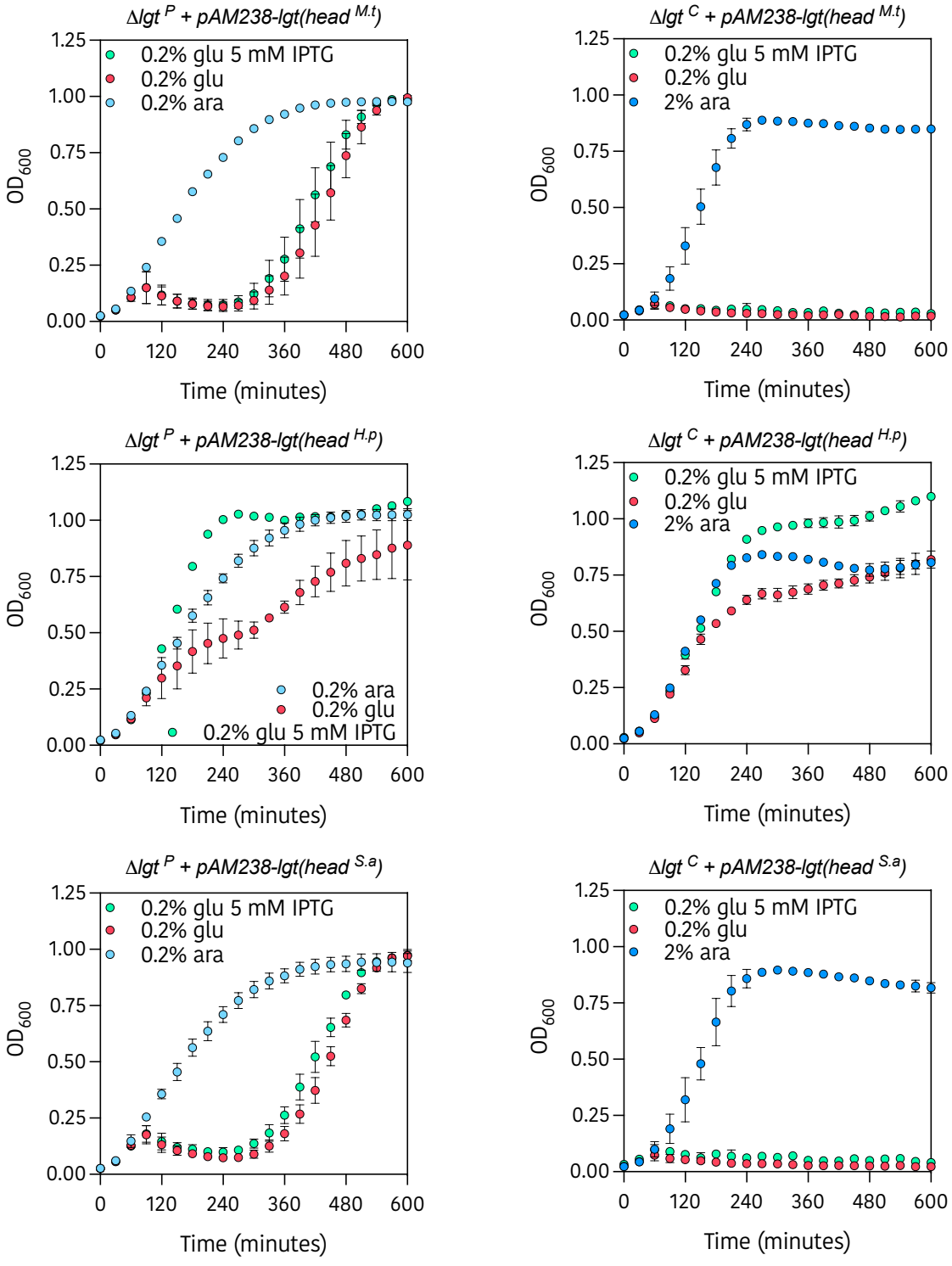

Δlgt C + pAM238
0.2% glu 5 mM IPTG
0.2% glu
0.2% ara
2% ara
OD600
Time (minutes)


Δlgt C + pAM238-lgt E.c
0.2% glu 5 mM IPTG
0.2% glu
0.2% ara
2% ara
OD600
Time (minutes)

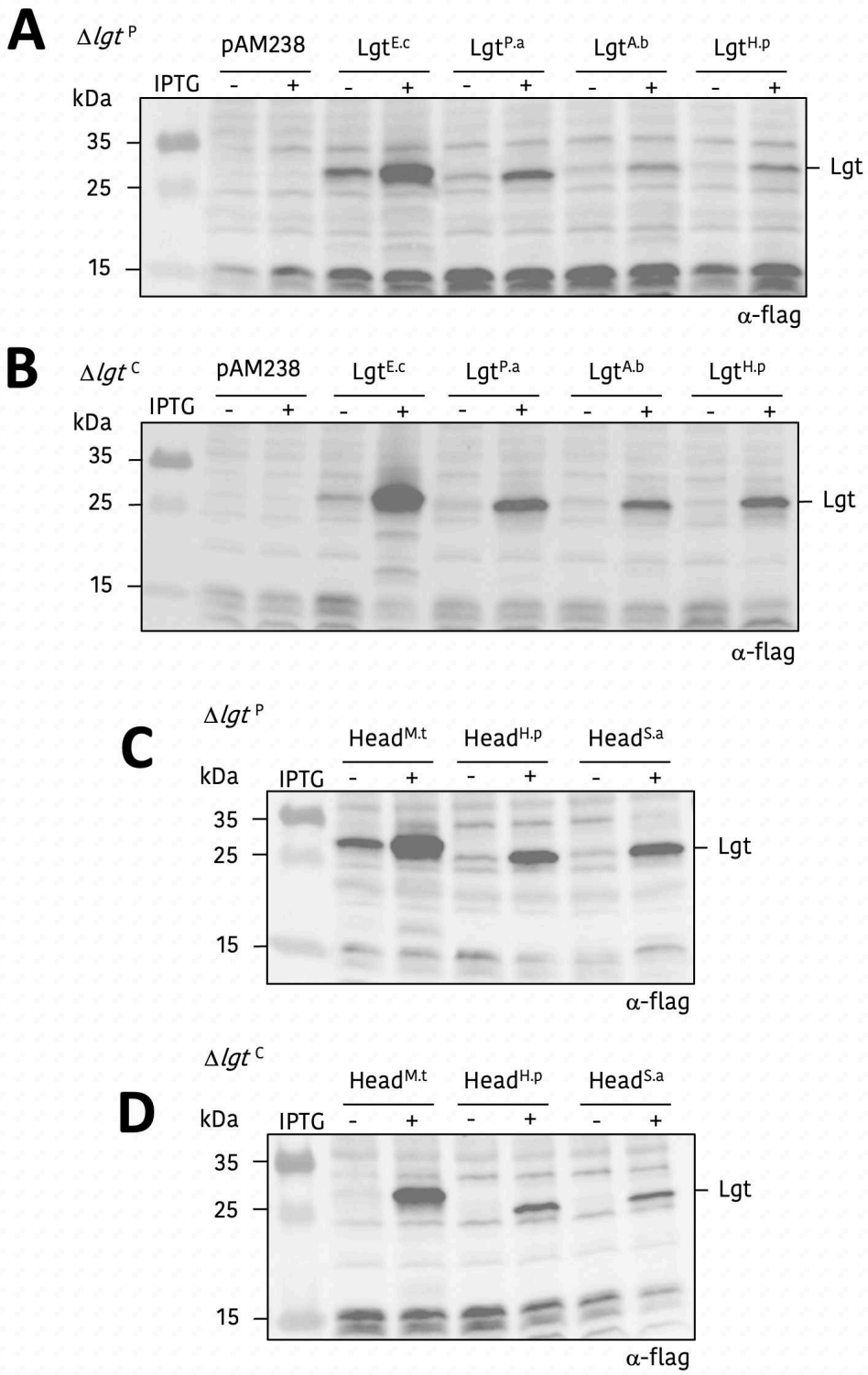

Δlgt C + pAM238-lgt P.a
0.2% glu 5 mM IPTG
0.2% glu
0.2% ara
2% ara
OD600
Time (minutes)


Δlgt C + pAM238-lgt A.b
0.2% glu 5 mM IPTG
0.2% glu
0.2% ara
2% ara
OD600
Time (minutes)


Δlgt C + pAM238-lgt H.p
0.2% glu 5 mM IPTG
0.2% glu
0.2% ara
2% ara
OD600
Time (minutes)

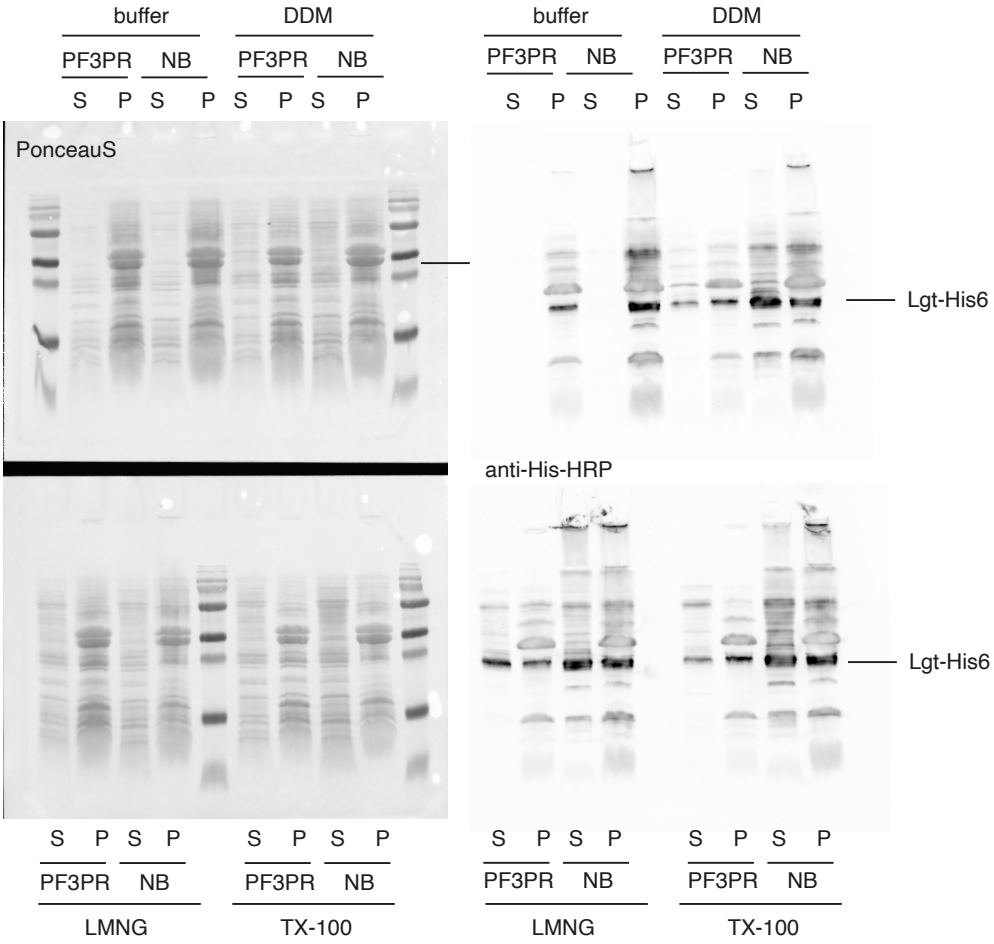


Appendix VII a **-** *Growth kinetics of Lgt depletion strains with complementing pAM238 plasmids grown in all conditions*



Appendix VII b - *Western blot from analysis of Lgt-flag expression.*

Appendix VIII: Solubilization of Lgt



Appendix VIII- *Left-hand panel represents PonceauS staining and right-hand panels represent α-His-HRP Western-blots. S = soluble fraction, P= pellet, DDM; LMNG and TX-100 refer to detergents. PF3PR and NB refer to the solubilisation procedure conducted from flask cell cultures (NB) fermenter cell cultures (PF3PR).*