UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Câmpus de São José do Rio Preto

Carolina Gismene

Determinantes estruturais da atividade necrótica e epidermolítica de toxinas esfoliativas produzidas por bactérias dos gêneros Mammaliicoccus e Staphylococcus

Campus de São José do Rio Preto
2024

Carolina Gismene

Determinantes estruturais da atividade necrótica e epidermolítica de toxinas esfoliativas produzidas por bactérias dos gêneros Mammaliicoccus e Staphylococcus

Tese apresentada como parte dos requisitos para obtenção do título de Doutor em Microbiologia, junto ao Programa de Pós-Graduação em Microbiologia, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Câmpus de São José do Rio Preto, IBILCE/UNESP, SP.

Financiadora: FAPESP (Processos: 2019/10230-0; 2020/13921-0 e 2022/10941-6) e CAPES

Orientador: Prof. Dr. Raghuvir Krishnaswamy Arni.

Coorientador: Dr. Ricardo Barros Mariutti

São José do Rio Preto
2024

G534d

Gismene, Carolina

Determinantes estruturais da atividade necrótica e epidermolítica de toxinas esfoliativas produzidas por bactérias dos gêneros Mammaliicoccus e Staphylococcus / Carolina Gismene. -- São José do Rio Preto, 2024

155 p.

Tese (doutorado) - Universidade Estadual Paulista (UNESP), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto

Orientador: Raghuvir Krishnaswamy Arni

1. toxina esfoliativa C. 2. toxina esfoliativa E. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Universidade Estadual Paulista (UNESP), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Carolina Gismene

Determinantes estruturais da atividade necrótica e epidermolítica de toxinas esfoliativas produzidas por bactérias dos gêneros Mammaliicoccus e Staphylococcus

Tese apresentada como parte dos requisitos para obtenção do título de Doutor em Microbiologia, junto ao Programa de Pós-Graduação em Microbiologia, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de São José do Rio Preto, IBILCE/UNESP, SP.

COMISSÃO EXAMINADORA

Prof. Dr. Raghuvir Krishnaswamy Arni
UNESP - Campus de São José do Rio Preto (IBILCE)
Orientador

Prof. Dr. Vasco Ariston de Carvalho Azevedo
UFMG – Universidade Federal de Minas Gerais

Prof. Dr. Patrick Jack Spencer
IPEN – Instituto de Pesquisas Energéticas e Nucleares

Profª. Drª. Flávia Cristina Rodrigues Lisoni
UNESP - Campus de São José do Rio Preto (IBILCE)

Profª. Drª. Aparecida Sadae Tanaka
UNIFESP – Universidade Federal de São Paulo

Data da Defesa 29 de abril de 2024

São José do Rio Preto
2024

DEDICATÓRIA

Dedico essa conquista à minha mãe, Maria Heloísa, meu maior exemplo de integridade e determinação.

"All my life through, the new sights of Nature made me rejoice like a child."

– Marie Curie

AGRADECIMENTOS

A todos os envolvidos nesse projeto, que foram muitos, em especial:

Aos meus orientadores, Prof. Dr. Raghuvir Krishnaswamy Arni e Dr. Ricardo Barros Mariutti, por terem me acolhido como aluna, pelos ensinamentos para a realização deste trabalho e aconselhamentos para a futura carreira como pesquisadora. Serei eternamente grata por todo empenho, paciência e dedicação para a minha formação acadêmica e pessoal. Não há livros, nem artigos que me proporcionariam maiores conhecimentos que nossas conversas em laboratório. Vocês me ensinaram ciência na prática. Eu os admiro e tenho muito orgulho de ter sido orientada por profissionais como vocês.

A Prof[a]. Dr[a]. Paula Rahal e Prof[a] Dr[a]. Rejane Maira Góes, por me receberem no laboratório de Estudos Genômicos e laboratório de Ciências Biológicas, respectivamente. Também agradeço a Dr[a]. Marília Calmon pela paciência e dedicação no ensino de novos conhecimentos que contribuíram para o meu crescimento como pesquisadora.

Ao Prof. Dr. Dieter Willbold, pela calorosa recepção no Institute of Biological Information Processing (IBI) Structural Biochemistry (IBI-7) - Forschungszentrum Jülich (Alemanha). Agradeço também pela orientação e empenho no desenvolvimento da minha pesquisa científica, e por me fazer sentir acolhida mesmo quilômetros de distância do meu país.

À minha mãe Maria Heloisa, que esteve sempre ao meu lado e nunca poupou esforços e trabalho para que meus sonhos fossem realizados. Obrigada por todo apoio no início e continuidade da minha formação. Você é o maior exemplo de força e superação que eu poderia ter e graças a você, conclui este trabalho com muito amor e dedicação.

À minha irmã, Beatriz, por ser minha melhor companhia do dia-dia e pelo apoio e ajuda que sempre me deu.

Ao meu físico preferido, Vítor, que durante a minha jornada profissional e pessoal nunca permitiu que eu me sentisse sozinha. Obrigada por compartilhar sonhos comigo e por transformar simples momentos em memórias inesquecíveis. Sua paciência, dedicação e companheirismo tornaram minha vida mais alegre e completa.

"In the vastness of space and the immensity of time, it is my joy to share a planet and an epoch with you."

Aos meus amigos Alana, Pulkit, Thainá, Vinícius, Ana Paula, Mariana e Taísa que compartilharam comigo momentos difíceis, mas também os momentos mais alegres e inesquecíveis vivenciados ao longo da minha pós-graduação. Obrigada por sempre estarem presentes de forma tão especial em minha vida.

Aos meus colegas de trabalho e colaboradores do Centro Multiusuário de Inovação Biomolecular e do laboratório de Estudos Genômicos: Jorge, Giovanna, Fernanda, Ângela, Barbosa, Bruno, Dayanne, Fábio, Luís, Vinícius, Kelly, Rafaela, Tamara, Dayla, Anelize e Felipe. Agradeço o companheirismo, amizade, sugestões e por tornarem o trabalho diário muito mais divertido e prazeroso.

A todos os professores que tive durante a graduação e pós-graduação. O conhecimento é o bem mais valioso do ser humano e vocês fizeram parte da construção do meu.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo apoio financeiro, concedido por meio dos Processos nº 2019/10230-0; 2020/13921-0 e 2022/10941-6. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Aos membros da banca examinadora por aceitarem contribuir com o presente trabalho. A todas estas pessoas, muito obrigada. A realização deste trabalho só foi possível com a participação de cada um.

RESUMO

Mammaliicoccus sciuri é uma bactéria comensal e patogênica de relevância clínica e veterinária. Recentemente uma cepa de *M. sciuri* foi descrita como agente etiológico da epidermite exsudativa (EE) em suínos e o principal fator de virulência envolvido nesta manifestação clínica foi a toxina esfoliativa C (ExhC). Essa enzima pertence ao grupo de toxinas esfoliativas (ETs), comumente descritas como glutamil endopeptidases semelhantes à quimotripsinas. No entanto, ao contrário da maioria dos membros desta família, as ETs são inativas contra uma ampla gama de substratos, apresentando alta especificidade à desmogleína-1 (Dsg-1), uma proteína de adesão que é crucial para manter a integridade da pele. A clivagem da Dsg-1 causa esfoliação da epiderme de mamíferos e aumenta a eficiência da invasão e infecção do hospedeiro por bactérias patogênicas. Esta especificidade ao substrato sugere que, além da tríade catalítica *H57-D102-S195*, outras regiões desconhecidas da superfície das ETs e Dsgs-1 podem ser essenciais para desencadeamento da atividade epidermolítica. A ExhC tem importância adicional pois, além de causar EE em suínos e esfoliação de pele em camundongos, é capaz de induzir necrose em múltiplas linhagens celulares de mamíferos e de inibir a atividade fagocítica de macrófagos, propriedades não observadas para outras ETs. Experimentos anteriores indicaram que interações específicas com o fragmento $ExhC_{44-92}$ são responsáveis pela atividade necrótica observada. A partir dessa observação, estudou-se a mutagênese sítio-dirigida de resíduos específicos não conservados dentro deste domínio, possibilitando a construção de uma variante de ExhC com mutação em quatro resíduos ($ExhC_{mut4}$) sem potencial necrótico, mas mantendo níveis de atividade enzimática similares da ExhC selvagem ($ExhC_{wt}$). Ademais, a determinação das estruturas cristalinas de $ExhC_{wt}$ e $ExhC_{mut4}$ garantiu a identificação da região associado à necrose. Apesar dos avanços na compreensão de como essa toxina atua, ainda não havia a descrição de inibidores para as múltiplas funções exercidas pela ExhC, assim como seus potenciais alvos moleculares durante os processos de necrose celular e inibição da atividade fagocítica de macrófagos. Para responder a essas questões, utilizamos a técnica de phage display e identificamos a *Prenylcysteine oxidase-1-like protein* de *M. musculus* como um potencial alvo da ExhC. A mesma técnica possibilitou a descoberta de peptídeos que atuam como inibidores e ativadores que modulam a atividade enzimática desta toxina. Os resultados obtidos tendo como alvo a ExhC, forneceram *insights* para investigações com à toxina esfoliativa E (ETE), a qual forma dímeros em solução. Essa informação, em conjunto com o fato de a ExhC ser encontrada como dímero na natureza, tornaram-se indicativos de que a homodimerização pode ser uma propriedade generalista entre as ETs, desempenhando papel funcional. No geral, os resultados aqui descritos representam avanços a respeito dos mecanismos de ação exercidos pela ExhC, contribuindo para uma compreensão mais profunda das ETs em um contexto amplo.

Palavras-chave: toxina esfoliativa C; toxina esfoliativa E; *Mammaliicoccus sciuri*; *Staphylococcus aureus*; atividade necrótica; atividade enzimática; região associada à necrose; *phage display*; peptídeos; inibidores; potencializadores; toxina esfoliativa E; homodimerização.

ABSTRACT

Mammaliicoccus sciuri is a commensal and pathogenic bacterium of clinical and veterinary relevance. Recently, a strain of *M. sciuri* was described as the etiologic agent of exudative epidermitis (EE) in pigs and the main virulence factor involved in this clinical manifestation was the exfoliative protein C (ExhC). This enzyme belongs to the group of exfoliative toxins (ETs), which are commonly described as glutamyl-endopeptidases belonging to the chymotrypsin family. However, unlike most members of this family, ETs are inactive against a wide range of substrates, showing high specificity for desmoglein-1 (Dsg-1), a cadherin-like adhesion protein that is crucial for maintaining skin integrity. Cleavage of Dsg-1 causes exfoliation of the mammalian epidermis and increases the efficiency of host invasion and infection by pathogenic bacteria. The previous substrate specificity suggests that, in addition to the *H57-D102-S195* catalytic triad, other unknown regions on the surface of ETs and Dsgs-1 may be essential for triggering exfoliative activity. ExhC plays an important role because, in addition, to causing EE in pigs and skin exfoliation in mice, it is capable of inducing necrosis in multiple mammalian cell lines and of inhibiting the phagocytic activity of macrophages, properties not observed for other ETs. Previous experiments have indicated that specific interactions with the $ExhC_{44-92}$ fragment are responsible for the necrotic activity. Based on this observation, site-directed mutagenesis of specific non-conserved residues within this fragment were studied, enabling the construction of an ExhC variant with a mutation in four residues ($ExhC_{mut4}$) lacking necrotic potential but retaining nearly wild-type ($ExhC_{wt}$) levels of enzymatic activity. Moreover, the determination of the crystal structures of $ExhC_{wt}$ and $ExhC_{mut4}$ enabled the identification of the necrosis-linked region. Despite some progress in understanding how this toxin acts, there has been no comprehensive description of inhibitors for the different functions carried out by ExhC, as well as its potential molecular targets involved in cell necrosis and the inhibition of macrophage phagocytic activity. To answer these questions, we used the phage display technique, and identified the Prenylcysteine oxidase-1-like protein of *M. musculus* as a potential target of ExhC. The same technique enabled the discovery of peptides acting as inhibitors and activators that modulate the enzymatic activity of this toxin. The results obtained using ExhC as a target give insights for research into exfoliative toxin E (ETE), which forms dimers in solution. This information, together with the fact that ExhC is found as a dimer in nature gives an indication that homodimerization may be a general property among ETs, playing a functional role. Overall, the results outlined here mark a noteworthy advancement in uncovering the mechanisms of action of ExhC, contributing to a deeper understanding of ETs within a broader context.

Keywords: exfoliative toxin C; exfoliative toxin E; *Mammaliicoccus sciuri*; *Staphylococcus aureus*; necrotic activity; enzymatic activity; necrosis-linked region; phage display; peptides; inhibitors; potentiators; exfoliative toxin E; homodimerization.

SUMÁRIO

PREFÁCIO……16
CAPÍTULO I. Considerações gerais 18
1.Introdução 19
1.1 *Mammaliicoccus sciuri* 19
1.2 Adesão celular e toxinas esfoliativas do gênero Staphylocci e Mammaliicocci 20
1.3 Classificação e análises estruturais das toxinas esfoliativas 22
1.4 Diversidade e patogenicidade das toxinas esfoliativas 24
1.5 Toxinas esfoliativas e suas propriedades multifuncionais 26
2. Objetivos 27
3. Materiais e métodos 27
3.1 Alinhamento de sequências primárias e modelos estruturais 27
3.2 Síntese de vetores de expressão em *Escherichia coli* 28
3.3 Transformação bacteriana 30
3.4 *Western blot* 30
3.5 Expressão das proteínas alvos 30
3.6 Purificação das proteínas alvos 31
3.6.1 Cromatografia de afinidade por Ni-Sepharose 31
3.6.2 Cromatografia de exclusão molecular 32
3.7 Diálise das toxinas alvos 32
3.8 Análise de pureza e estado oligomérico das toxinas por gel de poliacrilamida 33
3.9 Espalhamento dinâmico de luz 33
3.10 Dicroísmo circular 34
3.11 Linhagens celulares e condições de cultura 34
3.12 Ensaios de viabilidade celular 35
3.13 Ensaios de apoptose 35
3.14 Análises estatísticas 36
3.15 Ensaios enzimáticos 36
3.16 Phage display 37
3.16.1 Biblioteca TriCO-16 37
3.16.2 Amplificação dos fagos 38
3.16.3 Titulação dos fagos 38
3.16.4 Ensaio de imunoabsorção enzimática (ELISA) 39
3.16.5 Extração de DNA e PCR 39
3.16.5 Análises das sequências peptídicas 40

3.17 Ensaios de interação com peptídeos 41
3.18 Análises por *Basic Local Alignment Search Tool* (BLAST) 42
3.19 Cristalização, coleta de dados, processamento e determinação de estruturas 42
3.20 Espalhamento de raios-X em baixos ângulos 43
3.21 Análises computacionais das toxinas esfoliativas 44
3.21.1 *Proteins, Interfaces, Structures and Assemblies* (PDBePISA) e CavityPlus 44
3.21.2 Predição homodímeros 44
3.21.3 Cálculo do potencial eletrostático superficial e hidrofobicidade da ExhC e toxinas esfoliativas homólogas 45
3.21.4 Previsão de modos de ligação e simulações de Dinâmica Molecular 46
4. Referências 46
CAPÍTULO II. Necrotic activity of Exhc from *Mammaliicoccus sciuri* is mediated by specific amino acid residues 52
1. Introduction 54
2. Materials and Methods 57
2.1 Sequence alignment and structural models 57
2.2. Protein expression and purification 57
2.3. Dynamic light scattering 58
2.4. Circular dichroism spectroscopy 58
2.5. Cell lines and culture conditions 58
2.6. Cell viability assay 59
2.7. Apoptosis assay 59
2.8. Statistical analysis 59
2.9. Enzymatic Assays 60
2.10. Crystallization, data collection, processing and structure determination 60
2.11. Calculation of the surface electrostatic potential and hydrophobicity for ExhC and homologous exfoliative toxins 60
3. Results 61
3.1. Sequence and structural analyses of ExhC 61
3.2. Mutated variants of ExhC are dimers in solution 62
3.3. *In vitro* inhibition of ExhC necrotic activity in cell line BHK-21 by point mutations 64
3.4. Esterolytic activity of $ExhC_{wt}$ and $ExhC_{mut4}$ 66
3.5. The crystallographic structures of $ExhC_{wt}$ and $ExhC_{mut4}$ 67
3.6. Analysis of physico-chemical surface necrosis-liked region of ExhC and its homologous regions in other ETs 69
4. Discussion 71

5. Conclusion **73**
6. Acknowledgements **73**
7. References **73**
CAPÍTULO III. *Staphylococcus aureus* Exfoliative Toxin E, oligomeric state and flip of P186: implications for its action mechanism **77**
1. Introduction **79**
2. Results **83**
2.1. Refinement of the ETE Crystal Structure 83
2.2. Oxyanion Hole and P186 Flip 84
2.3. Esterolytic Activity of ETE Wild-Type and ETE/P186G Mutant 87
2.4. Oligomeric State of ETE in Solution 88
2.5. Prediction of the ETE Homodimer Structure in Solution 90
3. Discussion **93**
4. Materials and Methods **96**
4.1. Protein Expression and Purification 96
4.2. SDS-PAGE and BN-PAGE 15% 96
4.3. Enzymatic Assays 96
4.4. Crystallization, Data Collection, Processing and Structure Determination 97
4.5. Small-Angle X-ray Scattering Experiments 97
4.6. Computational Approaches to Predict the Conformation of ETE Homodimers 98
4.7. Molecular Dynamics Simulations 99
4.8. MM-GBSA Free Energy Calculations 100
4.9. Umbrella Sampling Free Energy Calculations 100
4.10. Trajectory Analyses 101
5. Conclusions **101**
6. Author Contributions **102**
7. Acknowledgments **102**
8. References **102**
Capítulo IV. Identification of potential partner protein of exfoliative toxin C from *Mammaliicoccus sciuri* and inhibitors/activators of its enzymatic activity **109**
1. Introdution **110**
2. Materials and methods **112**
2.1 Protein expression and purification 112
2.2 Phage display with TriCo-16 library 113
2.3 Amplification of phages 114
2.4 Phage titration 115

2.5 ELISA 115
2.6 DNA extraction and PCR 116
2.7 Analysis of peptide sequences 117
2.8 Enzimatic assays 117
2.9 Peptide-ExhC interaction assays 117
2.10 BLAST analysis 118
2.11 Prediction of binding modes and MD simulations 118
3. Results **119**
3.1 Phage display identification of ExhC-interacting peptides 119
3.2 Peptide-ExhC interaction assays 120
3.3 Search for ExhC partner proteins 122
4. Discussion 127
5. Conclusion 128
6. References 128
Capítulo V. Conclusões gerais 130
Referências 133
Apêndices 135
Appendix A - Supplementary material for the scientific article entitled Necrotic activity of ExhC from *Mammaliicoccus sciuri* is mediated by specific amino acid residues 136
Table S1. 136
Figure S1. 137
Figure S2. 137
Figure S3.. 138
Table S2. 138
Table S3. 139
Figure S4.. 139
References 140
Appendix B - Supplementary material for the scientific article entitled Staphylococcus aureus Exfoliative Toxin E, Oligomeric State and Flip of P186: Implications for Its Action Mechanism 141
Figure S1. Representation of the two chains in the asymmetric units in the available crystal structures of ETE. 141
Table S1. Analysis of the stability of putative ETE homodimers in the available crystal structures. 142
Figure S2. Fluctuation of P186 N-CA-C-O (ξ) and Y158 N-CA-CB-CG (χ1) dihedral angles throughout 1 μs replicate MD simulations of ETE monomers in water. 142
Figure S3. PMF for the rotation of L-alanyl-L-proline peptide bond. 143

Text S1. Size-exclusion chromatograms for ETE and ExhC.....143
Figure S4. Comparison between ETE of *S. aureus* and ExhC of *M. sciuri*.....144
Figure S5. Comparison between SEC chromatograms of ETE and ETE/P186G of *S. aureus*. 145
Figure S6. Workflow for the prediction of the most stable ETE dimers in solution. 146
Figure S7. Examples of SAXS intensity profile fits using three different homodimer models. 147
Figure S8. Models of ETE homodimer using ColabFold's AlphaFold-multimer version. 147
Table S2. Analyses conducted for the selected docking poses.....148
Table S3. Bioinformatic analysis of the stability of the predicted ETE homodimer. 150
Table S4. Comparison of homodimer interfaces in the available crystal structures and the proposed model. 150
Figure S9. Structural alignment of different serine proteases shown at the level of their primary sequences.....151
Figure S10. Structural alignment of different serine proteases.....152
Figure S11. Model of ETE in complex with Dsg1 EC3-EC4 linker and superposition with the predicted structure of ETE homodimer.....153
Text S2. Preparation of *N*-Tertbutyioxycarbonyl glutamic acid *O*-phenyl ester (Boc-L-Glu-OPh) 153
Text S3. Details of the Umbrella sampling free energy calculation protocols 154
References.....155

PREFÁCIO

A presente tese descreve o trabalho realizado pela aluna Carolina Gismene durante o desenvolvimento de sua pesquisa para obtenção do título de Doutor. O projeto teve como objetivo principal compreender estereoquimicamente os diferentes mecanismos de ação da toxina esfoliativa C (ExhC) de *Mammaliicoccus sciuri* e toxina esfoliativa E (ETE) de *Staphylococcus aureus*.

A tese foi dividida em cinco capítulos. O primeiro capítulo é composto por revisão literária detalhada em conjunto com os objetivos e materiais e métodos utilizados para a realização da pesquisa científica. Os capítulos dois e três são referentes a artigos publicados nas revistas internacionais "*International Journal of Biological Macromolecules*" (fator de impacto em 2024: 8.2) e "*International Journal of Molecular Sciences*" (fator de impacto em 2022: 5.6), respectivamente. O capítulo quatro, o qual ainda não foi publicado uma revista periódica, é referente aos resultados alcançados durante o estágio de pesquisa no exterior para o *Institute of Biological Information Processing* (IBI-7: *Structural Biochemistry*), localizado no Forschungszentrum Jülich (Alemanha). Por fim, o capítulo cinco, decreve as considerações gerais sobre as conquistas científicas obtidas durante o desenvolvimento deste trabalho.

CAPÍTULO I. Considerações gerais

1. Introdução

1.1 *Mammaliicoccus sciuri*

Mammaliicoccus sciuri é um grupo do gênero Mammaliicoccus formado por bactérias esféricas, cocos gram-positivas, imóveis, com diâmetros entre 0,7 a 1,5 µm, podendo ser encontradas de forma isolada ou em pares e aglomerados irregulares (Kloos; Ballard; Webster; Hubner et al., 1997; Lienen; Schnitt; Hammerl; Maurischat et al., 2022). São oxidase-positivas, anaeróbicas facultativas, mesófilas, coagulase-negativas e resistentes à novobiocina (Boone; Castenholz; Garrity; Stanley, 2001). Apresentam capacidade de adaptação a diferentes ambientes e alta capacidade de dispersão, além de serem resistentes a dessecação, sobrevivendo em ambientes não fisiológicos (Dakic; Morrison; Vukovic; Savic *et al.*, 2005; Kloos; Ballard; Webster; Hubner et al., 1997).

Devido a análises filogenéticas, esse grupo de micro-organismos foi recentemente reclassificado do gênero Staphylococcus para o gênero Mammaliicoccus (Lienen; Schnitt; Hammerl; Maurischat et al., 2022; Madhaiyan; Wirth; Saravanan, 2020) que é composto por cinco espécies (*M. sciuri, M. lentus, M. vitulinus, M. fleurettii e M. stepanovicii*) as quais frequentemente são apresentadas como bactérias comensais associadas a animais (Stepanovic; Dakic; Djukic; Lozuk et al., 2002). A espécie *M. sciuri* já foi encontrada em uma ampla gama de animais de sangue quente, incluindo marsupiais, roedores, macacos, cetáceos, gado, ovelhas, cavalos, cães e humanos (Adegoke, 1986; Kloos; Schleifer; Smith, 1976; Stepanović; Vuković; Ježek; Pavlović et al., 2001).

M. sciuri também já foi isolada de cabras doentes que sofreram de peste bovina e ovina (Adegoke, 1986) e, mais recentemente, em casos de dermatite canina (Hauschild; Wojcik, 2007) e em infecções patogênicas de leitões com epidermite exsudativa (EE) (Chen; Wang; Chen; Yang et al., 2007), comumente causada por *S. hyicus*, que produz toxinas esfoliativas essenciais para o estabelecimento e proliferação da infecção, além de serem os principais fatores que causam as lesões na pele do animal (Wegener; Andresen; Bille-hansen, 1993).

Em suínos a EE é uma doença de pele contagiosa que é vista principalmente como uma infecção aguda em recém-desmamados, no entanto pode ocorrer como uma infecção crônica em adultos (Schwarz; Loncaric; Brunthaler; Knecht et al., 2021; Wegener; Andresen; Bille-Hansen, 1993). As manifestações clínicas são caracterizadas por esfoliação acompanhada de separação celular epidérmica, eritema e exsudação serosa (Ahrens; Andresen, 2004). O suíno com epiderme exsudativa generalizada sofre de desidratação que pode levar à morte dentro de

poucos dias (Schwarz; Loncaric; Brunthaler; Knecht et al., 2021; Wegener; Andresen; Bille-Hansen, 1993).

Diversas ocorrências clínicas envolvendo esse micro-organismo sinalizam sua importância médica e veterinária (Dakic; Morrison; Vukovic; Savic et al., 2005; Nemeghaire; Vanderhaeghen; Argudin; Haesebrouck et al., 2014) uma vez que embora *M. sciuri* seja isolada principalmente em espécies de animais, também já foi o micro-organismo responsável por diversas manifestações clínicas em humanos, como endocardite, peritonite, choque séptico, infecção do trato urinário, doença inflamatória pélvica e infecções de feridas (Dakic; Morrison; Vukovic; Savic et al., 2005; Hedin; Widerstrom, 1998; Shittu; Lin; Morrison; Kolawole, 2004; Stepanovic; Dakic; Djukic; Lozuk et al., 2002; Wallet; Stuit; Boulanger; Roussel-Delvallez et al., 2000).

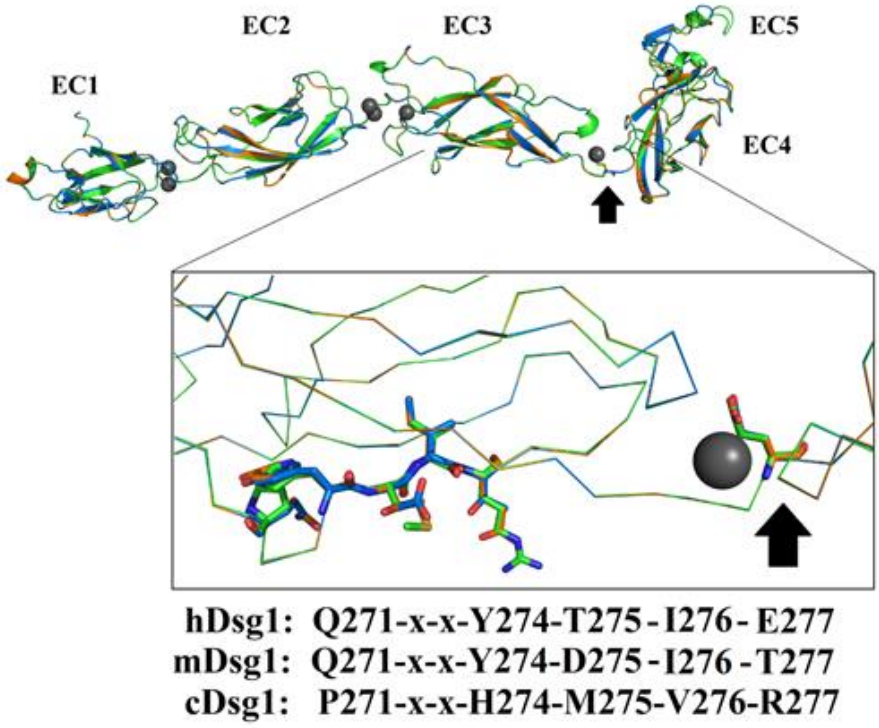
1.2 Adesão celular e toxinas esfoliativas do gênero Staphylocci e Mammaliicocci

As toxinas esfoliativas (ETs) são fatores de virulência expressos por certas cepas de estafilococos e mammaliicocos infecciosos, que clivam especificamente a caderina desmossômica desmogleína 1 (Dsg-1), proteína transmembrana formada por cinco ectodomínios (ECs) (Figura 1) (Mariutti; Tartaglia; Seyffert; De Paula Castro et al., 2017). Esse processo resulta em danos à epiderme (Ladhani; Joannou; Lochrie; Evans et al., 1999; Mariutti; Tartaglia; Seyffert; De Paula Castro et al., 2017).

A epiderme, que corresponde à camada mais superficial da pele, é um tecido epitelial multiestratificado com células justapostas, as quais funcionam como uma barreira de proteção contra a invasão de micro-organismos (Junqueira; Carneiro, 2018). A manutenção da estrutura tecidual por meio da adesão célula-célula é realizada por proteínas denominadas de desmogleínas (Dsgs) e desmocolinas (Dscs) (Junqueira; Carneiro, 2018). As Dsgs cruciais para a integridade da epiderme (Stanley; Amagai, 2006; Yin; Green, 2004) possuem 4 isoformas (Dsg-1 a Dsg-4) descritas e apenas a Dsg-1 é seletivamente reconhecida e hidrolisada pelas ETs, enquanto Dsg-3 não é clivada.

Como Dsg-1 é expressa em todos os estratos da epiderme e a Dsg-3 é expressa apenas nos estratos espinhoso e basal (Getsios; Huen; Green, 2004), nas camadas mais profundas da epiderme, a destruição de Dsg-1 por ETs é compensada pela presença de Dsg-3, ou seja, a esfoliação causada pelas ETs ocorre apenas no estrato granuloso, onde Dsg-3 não está presente (Figura 2) (Bukowski; Wladyka; Dubin, 2010).

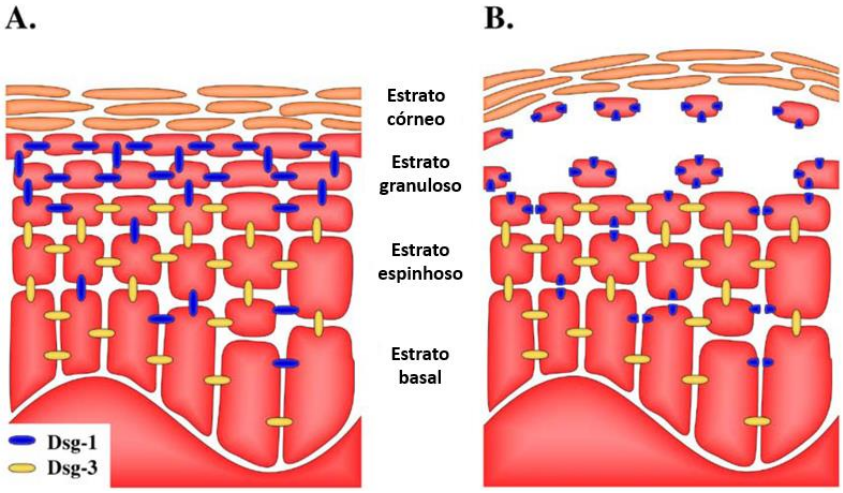
Figura 1. Sobreposição dos modelos estruturais indicando os 5 ectodomínios (ECs) de Dsg-1 humana (laranja; hDsg-1), Dsg-1 canina (verde; cDsg-1) e Dsg-1 murina (azul; mDsg-1). As esferas cinzas representam íons de cálcio e o ácido glutâmico (Glu381) do sítio de clivagem é indicado por uma seta preta. As sequências de aminoácidos da Dsg-1 humana, murina e canina, 110 resíduos *upstream* ao sítio de clivagem, estão destacadas. A letra x indica resíduos de aminoácidos das Dsgs-1 considerados não cruciais para o reconhecimento e clivagem das Dsgs-1 pela ETA.



Fonte: MARIUTTI et al., 2017.

Embora exista alta homologia estrutural entre as Dsgs-1 de diferentes mamíferos (Figura 1), há também uma alta especificidade das ETs no reconhecimento e clivagem dos seus substratos (Bukowski; Wladyka; Dubin, 2010; Mariutti; Tartaglia; Seyffert; De Paula Castro et al., 2017), o que ainda é pouco estudado e compreendido. A toxina esfoliativa A (ETA) de *S. aureus* reconhece especificamente cinco resíduos de aminoácidos *upstream* ao sítio de clivagem das Dsgs-1 humana, murina e canina (Figura 1) e esses resíduos são cruciais para o acoplamento da enzima ao seu substrato, entretanto a hidrólise enzimática, entre os resíduos E381 e G382, ocorre apenas nas Dsgs-1 humana e murina, que possuem quatro resíduos de aminoácidos idênticos (Hanakawa; Schechter; Lin; Nishifuji et al., 2004).

Figura 2. A distribuição diferencial das isoformas de desmogleína (Dsg-1 e Dsg-3) na epiderme explica a ação induzida por toxinas esfoliativas no estrato granuloso. Representação esquemática da distribuição de desmogleína em **(A)** pele saudável e **(B)** pele exposta às ETs. Em todos os estratos, exceto o estrato granuloso, a hidrólise da Dsg-1 mediada por toxinas esfoliativas é compensada pela Dsg-3. Dsg-3 está ausente no estrato granuloso, o que explica o desprendimento celular e a divisão das camadas epidérmicas após a hidrólise de Dsg-1.



Fonte: BUKOWSKI et al., 2010.

1.3 Classificação e análises estruturais das toxinas esfoliativas

As ETs são glutamil endopeptidases (GEPs) pertencentes à família das serino proteinases semelhantes à quimotripsina (CLSP) (Demidyuk; Chukhontseva; Kostrov, 2017) que clivam ligações peptídicas seguindo o grupo α-carboxila de resíduos Glu (Demidyuk; Chukhontseva; Kostrov, 2017).

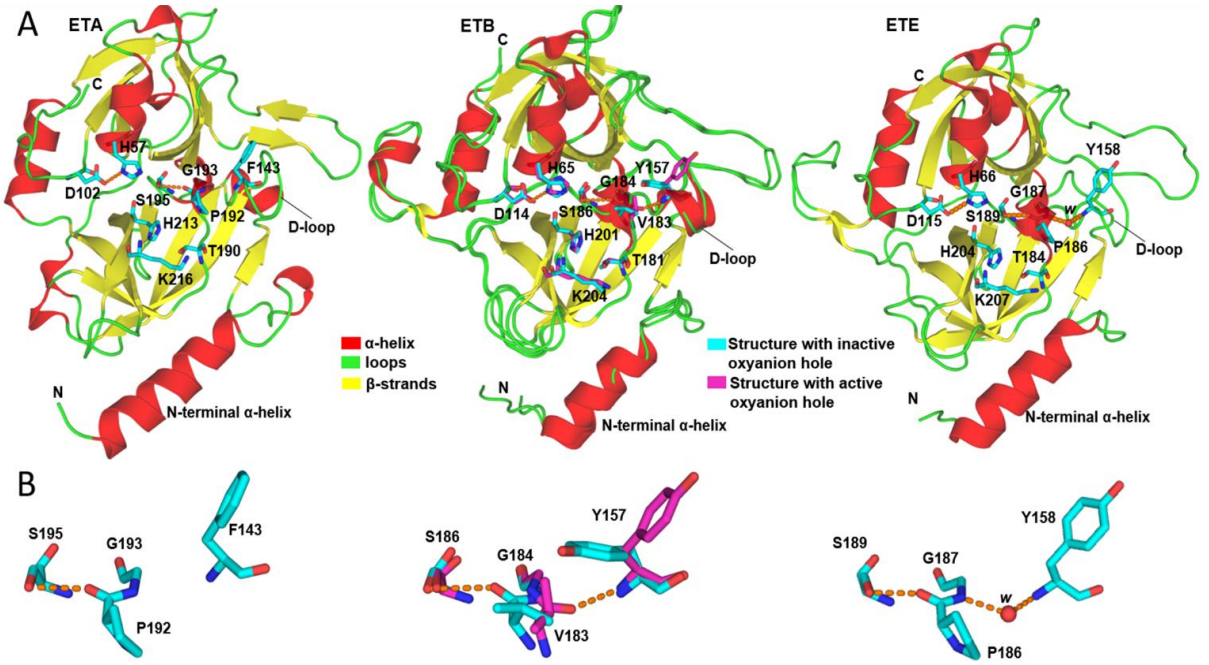
Estruturalmente, essas toxinas apresentam dobramento típico das CLSPs, composto por dois domínios, cada um deles contendo seis cadeias-β antiparalelas que formam um barril-β (Demidyuk; Chukhontseva; Kostrov, 2017; Vath; Earhart; Monie; Iandolo et al., 1999) e carregam tríade catalítica composta pelos resíduos *H57-D102-S195* (os resíduos em itálico são baseados na numeração da quimotripsina) (Figura 3.A) (Demidyuk; Chukhontseva; Kostrov, 2017).

Além disso, o *pocket* S1 das ETs contém dois elementros estruturais presentes em todas as GEPs: His e Thr/Ser (Thr nas ETs) nas posições *213* e *190* (Figura 3.A), os quais desempenham papel fundamental na preferência das GEPs por Glu (Demidyuk; Chukhontseva; Kostrov, 2017). Em contrapartida, a maioria das ETs descritas possuem o resíduo Lys na posição *216* no lugar do resíduo Ser encontrado nas GEPs (Figura 3.A), exceto ExhA de *S. hyicus* , que carrega Arg na posição equivalente (Demidyuk; Chukhontseva; Kostrov, 2017).

As estruturas cristalinas disponíveis das ETs de *S. aureus* (PDBs: 1EXF, 1AGJ, 1DUE e 1DUA para ETA; 1DT2 e 1QTF para ETB; e 5C2Z para ETE) revelaram duas características estruturais únicas para as ETs, que diferem do observado para as CLSPs: (i) a presença de uma α-hélice N-terminal contendo vários aminoácidos carregados e (ii) uma conformação incomum dos resíduos que formam o *oxyanion hole* (Figura 3.A, B) (Cavarelli; Prevost; Bourguet; Moulinier et al., 1997; Demidyuk; Chukhontseva; Kostrov, 2017; Papageorgiou; Plano; Collins; Acharya, 2000; Vath; Earhart; Rago; Kim et al., 1997). Sabe-se que a deleção da extensão da α-hélice N-terminal da ETA anula atividade catalítica dessa enzima, destacando a importância funcional deste elemento estrutural (Cavarelli; Prevost; Bourguet; Moulinier et al., 1997).

Em referência a conformação incomum do *oxyanion hole* das ETs, observou-se pela primeira vez na ETA que a ligação peptídica *P192-G193* é invertida 180° em relação à sua posição típica em outras CLSPs (Vath; Earhart; Rago; Kim et al., 1997). Devido a essa conformação incomum, denominada I (I=inativa), o oxigênio carbonílico da *P192* (*P192*(O)) direciona-se para o grupo -OH da *S195*, formando uma ligação de hidrogênio (Figura 3.B), o que aparentemente dificultaria o acesso dos substratos ao *oxyanion hole* (Vath; Earhart; Rago; Kim et al., 1997). Assim, foi proposto que a ativação catalítica da ETA exigiria que o P192(O) adotasse a conformação geralmente observada em CLSP, aqui denominada A (A = ativa) (Cavarelli; Prevost; Bourguet; Moulinier et al., 1997; Vath; Earhart; Rago; Kim et al., 1997). A partir de análises estruturais, Vath et al (1997) propôs a hipótese de que a mudança de *P192*(O) da conformação I para A poderia ser promovida por um ativador desconhecido que induziria uma mudança conformacional no loop D ($loop_{143-153}$, Figura 3.A) interagindo com a α-hélice N-terminal (Vath; Earhart; Rago; Kim et al., 1997). A mesma conformação I das estruturas cristalinas da ETA (PDB: 1AGJ, 1DUE e 1DUA), também foram observadas para V183(O) e P186(O) de ETB (PDB: 1QTF) e ETE (PDB: 5C2Z), respectivamente (Figura 3.B).

Figura 3. (A) Representações 3D das estruturas cristalinas da ETA (PDB: 1EXF), ETB (PDB 1DT2 *sticks* em magenta e 1QTF *sticks* em ciano) e ETE (PDB: 5C2Z), respectivamente. As principais características estruturais dessas proteínas estão representadas: tríade catalítica D102/114/115-H75/65/66-S195/186/189, a hélice α N-terminal, a presença no *pokcket* S1 dos resíduos T190 /181/184, H213/201/204 e K216/204/207, conformação invertida atípica do átomo P192/V183/P186(O) e o loop D, proposto como fundamental para o controle do giro de 180 º da *P192*. **(B)** *Sticks* com conformações alternativas P192/V183/P186(O) observadas nas estruturas cristalinas disponíveis de ETA, ETB e ETE. Em todos os casos, as linhas tracejadas em laranja representam ligações H. Os resíduos de ETA são rotulados de acordo com a numeração de quimotripsina para ETA, como no PDB: 1EXF. Como os PDBs disponíveis de ETB e ETE não seguem a numeração da quimotripsina, os resíduos dessas toxinas são rotulados em ordem sequencial. Os rótulos numéricos separados por / correspondem às posições dos resíduos equivalentes nas sequências ETA, ETB e ETE, nesta ordem. Digno de nota, uma molécula de água, nomeada w, forma uma ponte de água (esfera vermelha) entre G187(N) e Y158(N) na estrutura cristalina de ETE (PDB: 5C2Z).



Fonte: GISMENE et al., 2022.

1.4 Diversidade e patogenicidade das toxinas esfoliativas

Quatro ETs nomeadas como ETA (Lee; Schmidt; Johnson-Winegar; Spero et al., 1987; O'toole; Foster, 1987) ETB (Lee; Schmidt; Johnson-Winegar; Spero et al., 1987), ETD (Yamaguchi; Nishifuji; Sasaki; Fudaba et al., 2002) e ETE (Le Marechal; Jardin; Jan; Even et al., 2011; Le Marechal; Seyffert; Jardin; Hernandez et al., 2011) foram identificados em diferentes cepas de *S. aureus* e três delas (ETA, ETB e ETD) estão associadas a infecções humanas (Imanishi; Nicolas; Caetano; Castro et al., 2019; Ladhani; Joannou; Lochrie; Evans et al., 1999; Yamaguchi; Nishifuji; Sasaki; Fudaba et al., 2002), sendo responsáveis por síndromes estafilocócicas que variam em nível de gravidade, de impetigo bolhoso localizado à síndrome

da pele escaldada estafilocócica (SSSS) (Mariutti; Souza; Ullah; Caruso et al., 2015), além de causar a esfoliação epidérmica em ratos (Melish; Glasgow, 1970). Dentre as três, a ETA e ETB são as que causam maiores danos à pele humana e estão presentes na maioria das cepas infectantes (Lee; Schmidt; Johnson-Winegar; Spero et al., 1987).

Já as cepas produtoras de ETD são principalmente isoladas de furúnculos e abscessos cutâneos (Yamaguchi; Nishifuji; Sasaki; Fudaba et al., 2002). Em contrapartida, a ETE foi identificada em mastites de ovinos, não sendo isolada de infecções em humanos (Imanishi; Nicolas; Caetano; Castro et al., 2019). Entretanto, em ensaios *in vitro* e *in silico*, a ETE de *S. aureus* foi capaz de clivar a Dsg-1 caprina, ovina, humana, murina e suína, apresentando baixa especificidade ao substrato em comparação as demais ETs já descritas (Imanishi; Nicolas; Caetano; Castro et al., 2019).

Além de *S. aureus*, outras ETs foram isoladas de espécies de Staphylococci (Bukowski; wladyka; dubin, 2010) e Mammaliicocci (Madhaiyan; Wirth; Saravanan, 2020). Por exemplo, cepas de *S. hyicus* produzem pelo menos seis ETs diferentes, denominados ExhA, ExhB, ExhC, ExhD, ShetA e ShetB (Fudaba; Nishifuji; Andresen; Yamaguchi et al., 2005). Estas estão associados à ocorrência da EE em suínos causada pela degradação da Dsg-1 suína (Fudaba; Nishifuji; Andresen; Yamaguchi et al., 2005). Ultimamente, a incidência de EE tem aumentado através de grandes surtos que causaram alta mortalidade em leitões e, consequentemente, perdas significativas para a agropecuária (Schwarz; Loncaric; Brunthaler; Knecht et al., 2021).

Sequências nucleotídicas que codificam as toxinas ExhA, ExhB, ExhC, ExhD e ShetB foram comparadas com os genes das três toxinas de *S. aureus* associadas a infecções humanas (Ahrens; Andresen, 2004). Todas essas toxinas estão relacionadas entre si com uma identidade de 49 a 54%. Curiosamente, a ExhC apresentou maior identidade com ETA do que com as demais toxinas esfoliativas de *S. hyicus* e a ExhD possui substituição do resíduo de aminoácido *D102* tríade catalítica por *E102* (Ahrens; Andresen, 2004).

ExhC também foi encontrado em uma cepa altamente patogênica de *Staphylococcus sciuri* (recentemente reclassificada como *Mammaliicoccus sciuri* (Madhaiyan; Wirth; Saravanan, 2020) responsável por um surto de EE em suínos na China (Chen; Wang; Chen; Yang et al., 2007). Além disso, uma variante de ExhB, diferindo em apenas dois resíduos em relação àquele expresso em *S. hyicus*, foi identificada em *S. chromogenes* (Andresen; Ahrens; Daugaard; Bille-Hansen, 2005), enquanto duas novas ETs, ExpA (=EXI) (Iyori; Futagawa-Saito; Hisatsune; Yamamoto et al., 2011) e ExpB (Iyori; Hisatsune; Kawakami; Shibata et al.,

2010), com capacidade de clivar Dsg-1 canina, foram identificados em cepas de *S. pseudintermedius*.

1.5 Toxinas esfoliativas e suas propriedades multifuncionais

Algumas ETs apresentam outras funções biológicas além da atividade esfoliativa e, portanto, podem ser classificadas como *moonlighting proteins*, ou seja, uma mesma proteína possui múltiplas funções dependendo do contexto biológico em que se encontra (Singh; Bhalla, 2020). Por exemplo, ETA (Rago; Vath; Bohach; Ohlendorf et al., 2000) e ETB (Monday; Vath; Ferens; Deobald et al., 1999) foram relatadas como superantígenos e possuem atividade mitogênica de linfócitos T, enquanto ExhC induz necrose em múltiplas linhagens celulares (Li; Wang; Ding; Zheng, 2011), propriedades que, juntamente com a clivagem de Dsg-1, podem ser importantes para a patogênese dos microrganismos que expressas essas toxinas. Na tentativa de identificar a região de ExhC responsável por esta atividade, diversas variantes truncadas foram produzidas e seu potencial necrótico foi avaliado (Li; Wang; Ding; Zheng, 2011).

Foi demonstrado que a deleção ou bloqueio do fragmento $ExhC_{79\text{-}128}$ (ou fragmento $ExhC_{44\text{-}92}$ na ausência do peptídeo sinal) inibe a necrose induzida pela ExhC (Li; Wang; Ding; Zheng, 2011), no entanto, a conformação, distribuição de carga e interações específicas responsáveis por isso não foram caracterizadas. O alinhamento das sequências de resíduos de aminoácidos de ETs já descritas com a ExhC de *M. sciuri* indicou a presença de 33 resíduos de aminoácidos conservados, sendo que dez deles estão localizados entre os resíduos 44-92 (20%), o que indica que essa região é a mais conservada (Li; Wang; Ding; Zheng, 2011). Por meio do alinhamento de sequências, também foi possível observar que a sequência de resíduos de aminoácidos da ExhC de *M. sciuri* (GenBank ID: JF755400) e da ExhC de *S. hyicus* (GenBank ID: AF515455) são idênticas (Li; Li; Lu; Wang et al., 2011).

Curiosamente, cepas virulentas da espécie *S. hyicus* têm sido isoladas consistentemente em lesões da síndrome de necrose da orelha (SNO) em suínos, na margem entre o tecido normal e necrótico (Morés, 2011). Embora a necrose seja atribuída, principalmente, a invasão de Streptococcus beta hemolítico, a patogenia da doença ainda não está totalmente elucidada (Andresen, 2005; Carvalho et al., 2007).

Além disso, foi demonstrado que a ExhC inibe a fagocitose de Mammaliicocci por macrófagos *in vitro* (Li; Li; Lu; Wang et al., 2011), sugerindo que essa toxina além de causar a esfoliação epidérmica e necrose celular, também pode desempenhar um papel importante na inibição da resposta imune inata do hospedeiro contra a infecção por *M. sciuri* (Li; Li; Lu;

Wang et al., 2011). Diferente do observado em morte celular por apoptose, a morte celular por necrose gera uma resposta imune no hospedeiro, envolvendo a ação de macrófagos no tecido alvo de necrose celular (Alberts et al., 2017).

2. Objetivos

O objetivo principal deste projeto de pesquisa foi utilizar ferramentas de bioinformática para projetar mutantes de ExhC a fim de identificar os resíduos de aminoácidos específicos envolvidos na necrose celular, além de avaliar a influência desse potencial de ação na atividade esfoliativa causada pela toxina. Ademais, investigou-se se outras ETs são capazes de causar necrose celular, ou se essa propriedade é uma característica exclusiva da ExhC. Para obtenção dos resultados, foram utilizadas técnicas de biologia molecular e estrutural, cristalografia de proteínas e ensaios *in vitro* com diferentes linhagens celulares, incluindo a BHK-21 já testada (Li; Wang; Ding; Zheng, 2011), que foram incubadas com a ExhC, selvagem e mutantes, com substituições no conjunto de aminoácidos R47, N49, Q51 e R89. Após a identificação dos resíduos de aminoácidos da ExhC responsáveis pela atividade necrótica, possíveis proteínas parceiras envolvidas no desencadeamento da morte celular e potenciais ligantes que modulassem as múltiplas funções dessa toxina, foram investigados e planejados, respectivamente, por meio do uso da técnica de phage display.

Devido a amplitude dos resultados obtidos, o estado oligomérico da ExhC e ETE e o papel da formação de dímeros em solução, foi investigado por técnicas experimentais e computacionais. Esses dados em conjunto com os resultados obtidos por meio da construção de uma variante mutante ETE_{P186G}, possibilitou a elaboração de novas hipóteses a respeito do funcionamento do mecanismo de ação esfoliativo das ETs.

3. Materiais e métodos

3.1 Alinhamento de sequências primárias e modelos estruturais

A sequência de aminoácidos da ExhC de *M. sciuri* foi comparada com as demais ETs descritas na literatura a partir do alinhamento de múltiplas sequências realizado no servidor Muscle (www.ebi.ac.uk/Tools/msa/muscle/) (Edgar, 2004). O servidor SMS2 (Stothard, 2000) foi usado para visualizar o alinhamento dos resíduos 44-92 da ExhC ($ExhC_{44-92}$) e as regiões análogas de outras ETs de estafilocócicos. As sequências lineares das proteínas foram obtidas no banco de dados UniProt (//www.uniprot.org/): ExhC de *M. sciuri* (UniProt ID: F6M8N2); ETA (ID UniProt: P09331), ETB (ID UniProt: P09332), ETD (ID UniProt: Q8GAX8) e ETE

(ID UniProt: L0RUV7) de *S. aureus*; ExhA (ID UniProt: Q8GNJ7), ExhB (ID UniProt: Q8GNJ6), ExhC (ID Uniprot: Q8GNJ5), ExhD (ID UniProt: Q8GNJ4), ShetA (ID UniProt: E5RQL1) e ShetB (ID UniProt: Q9KWH0) de *S. hyicus*; ExhB (UniProt ID: Q8GNJ6) de *S. chromogenes*; EXI (ID UniProt: D0VXY8) e ExpB (ID Uniprot: E2RZJ7) de *S. pseudintermedius*. As proteínas ETC de *S. aureus* (Uniprot ID: Q9KWG9) e SIET de *S. pseudintermedius* (Uniprot ID: A0A8F4Q3I3) não foram adicionados ao alinhamento primário, visto que ambas não são CLSPs (Nishifuji; Sugai; Amagai, 2008). Os modelos estruturais da $ExhC_{wt}$ e do domínio necrótico $ExhC_{44\text{-}92}$ (Li; Wang; Ding; Zheng, 2011) foram gerados usando AlphaFold2 (Jumper; Evans; Pritzel; Green et al., 2021).

3.2 Síntese de vetores de expressão em *Escherichia coli*

Inicialmente, dois alvos foram construídos: a ExhC selvagem ($ExhC_{wt}$) (i) e com mutação em R47, N49, Q51 e R89 ($ExhC_{mut4}$) (ii) de *M. sciuri*. Após a obtenção de resultados preliminares que demonstram que a $ExhC_{mut4}$ (ii) não influência na viabilidade celular da linhagem BHK-21, quatro novos vetores de expressão que codificam a ExhC de *M. sciuri* com mutações pontuais $ExhC_{R47A}$ (iii), $ExhC_{N49F}$ (iv), $ExhC_{Q51K}$ (v) e $ExhC_{R89A}$ (vi) foram confeccionados. Os plasmídeos pET-28a (+) contendo as sequências nucleotídicas codificadoras das proteínas alvo fusionadas a uma cauda de histidina (His-tag) e apresentando sítio de clivagem para TEV-protease, foram sintetizados pela Genscript.

As toxinas ETA (vii) e ETE (viii) de *S. aureus* também foram incluídas no projeto de pesquisa para comparações com os resultados obtidos por meio de ensaios *in vitro* com a ExhC selvagem e mutantes. Também foi construído o vetor para a síntese da ETE_{P186G} (ix) com o objetivo de avaliar a importância do resíduo *P192* para atividade enzimática das ETs.

A ETA e ETE_{P186G} foram sintetizadas pela Genscript em plasmídeo pET-28a (+), enquanto a ETE foi sintetizada pela GenScript em plasmídeo pD441. As três proteínas foram fusionadas a uma cauda de histidina (His-tag), mas apenas a ETA possui sítio de clivagem para a TEV-protease. Na Tabela 1 estão representadas as sequências de aminoácidos que codificam as proteínas alvo.

Tabela 1. Sequência de aminoácidos das proteínas recombinantes $ExhC_{wt}$ (i), $ExhC_{mut4}$ (ii), $ExhC_{R47A}$ (iii), $ExhC_{N49F}$ (iv), $ExhC_{Q51K}$ (v), $ExhC_{R89A}$ (vi), ETA (vii), ETE (viii) e ETE_{P186G} (ix). Em destaque, a cauda de histidina (amarelo), o sítio de clivagem para a TEV-protease (azul-claro), os resíduos de aminoácidos que diferenciam a $ExhC_{wt}$ das suas variantes mutantes (rosa) e o resíduo P186 da ETE que foi substituído por G186 na ETE_{P186G} (cinza).

(I) MDEESDLKDHRDKWNKYYGVSPDQLSKDLFDKVSPEQIKNSPYQSVGRLNVQGEAVATGVFIGKNTV
VTNHHIAKEAKNNPSKIIFSPGRHADESNTGTVLPHGTFEASEIIDAPFGTGVDISVIIFKPNAEGKSIGDVIKAADL
GNSNSLKKGDTANLIGYPYDFDSKNMYRSQVEFQSTDFGLKYYGYTVPGNSGSGIFNSEGKFVGLHIGKAKHINS
QNEINYAVSFNDFLIRDLKQLIKGGENLYFQSHHHHHH

(II) MDEESDLKDHRDKWNKYYGVSPDQLSKDLFDKVSPEQIKNSPYQSVGALFVKGEAVATGVFIGKNTVV
TNHHIAKEAKNNPSKIIFSPGAHADESNTGTVLPHGTFEASEIIDAPFGTGVDISVIIFKPNAEGKSIGDVIKAADLG
NSNSLKKGDTANLIGYPYDFDSKNMYRSQVEFQSTDFGLKYYGYTVPGNSGSGIFNSEGKFVGLHIGKAKHINSQ
NEINYAVSFNDFLIRDLKQLIKGGENLYFQSHHHHHH

(III) MDEESDLKDHRDKWNKYYGVSPDQLSKDLFDKVSPEQIKNSPYQSVGALNVGEAVATGVFIGKNTVV
TNHHIAKEAKNNPSKIIFSPGRHADESNTGTVLPHGTFEASEIIDAPFGTGVDISVIIFKPNAEGKSIGDVIKAADLG
NSNSLKKGDTANLIGYPDFDSKNMYRSQVEFQSTDFGLKYYGYTVPGNSGSGIFNSEGKFVGLHIGKAKHINSQN
EINYAVSFNDFLIRDLKQLIKGGENLYFQSHHHHHH

(IV) MDEESDLKDHRDKWNKYYGVSPDQLSKDLFDKVSPEQIKNSPYQSVGRLFVGEAVATGVFIGKNTVVT
NHHIAKEAKNNPSKIIFSPGRHADESNTGTVLPHGTFEASEIIDAPFGTGVDISVIIFKPNAEGKSIGDVIKAADLGN
SNSLKKGDTANLIGYPDFDSKNMYRSQVEFQSTDFGLKYYGYTVPGNSGSGIFNSEGKFVGLHIGKAKHINSQNEI
NYAVSFNDFLIRDLKQLIKGGENLYFQSHHHHHH

(V) MDEESDLKDHRDKWNKYYGVSPDQLSKDLFDKVSPEQIKNSPYQSVGRLNVKEAVATGVFIGKNTVV
TNHHIAKEAKNNPSKIIFSPGRHADESNTGTVLPHGTFEASEIIDAPFGTGVDISVIIFKPNAEGKSIGDVIKAADLG
NSNSLKKGDTANLIGYPDFDSKNMYRSQVEFQSTDFGLKYYGYTVPGNSGSGIFNSEGKFVGLHIGKAKHINSQN
EINYAVSFNDFLIRDLKQLIKGGENLYFQSHHHHHH

(VI) MDEESDLKDHRDKWNKYYGVSPDQLSKDLFDKVSPEQIKNSPYQSVGRLNVGEAVATGVFIGKNTVV
TNHHIAKEAKNNPSKIIFSPGAHADESNTGTVLPHGTFEASEIIDAPFGTGVDISVIIFKPNAEGKSIGDVIKAADLG
NSNSLKKGDTANLIGYPDFDSKNMYRSQVEFQSTDFGLKYYGYTVPGNSGSGIFNSEGKFVGLHIGKAKHINSQN
EINYAVSFNDFLIRDLKQLIKGGENLYFQSHHHHHH

(VII) HHHHHHMGGENLYFQSEVSAEEIKKHEEKWNKYYGVNAFNLPKELFSKVDEKDRQKYPYNTIGNVFV
KGQTSATGVLIGKNTVLTNRHIAKFANGDPSKVSFRPSINTDDNGNTETPYGEYEVKEILQEPFGAGVDLALIRLKP
DQNGVSLGDKISPAKIGTSNDLKDGDKLELIGYPFDHKVNQMHRSEIELTTLSRGLRYYGFTVPGNSGSGIFNSNG
ELVGIHSSKVSHLDREHQINYGVGIGNYVKRIINEKNELE

(VIII) MGSSHHHHHHMIEYTDEEIQKKRDFFKTRPSDSELFSKIQDTTRSPYSSVGTVFVKGKTIATGILIGKNTV
ITNKHIARLAENDPNKVIFTPGSTRDEGSLVVKKPFGEFIAEEINEAPYGGGTDLSIIKLKPNQYGKSAGDLVTPAAI
PDNVDVQKGDKISLLGYPYNTSTHSLYKSQIEVFNNQTFQYFAYTEPGNSGSGIFNLHGELVGIHSGKGGQYGLP
FGILFNRQIGSSYSTDKTVTTLAIDLKNKAKTQEQ

(IX) MGSSHHHHHHMIEYTDEEIQKKRDFFKTRPSDSELFSKIQDTTRSPYSSVGTVFVKGKTIATGILIGKNTV
ITNKHIARLAENDPNKVIFTPGSTRDEGSLVVKKPFGEFIAEEINEAPYGGGTDLSIIKLKPNQYGKSAGDLVTPAAI
PDNVDVQKGDKISLLGYPYNTSTHSLYKSQIEVFNNQTFQYFAYTEGGNSGSGIFNLHGELVGIHSGKGGQYGLP
FGILFNRQIGSSYSTDKTVTTLAIDLKNKAKTQEQ

Fonte: própria autora.

3.3 Transformação bacteriana

Linhagens celulares competentes de *E. coli* C43(DE3), C41(DE3), BL21(DE3), BL21 Star-pLysS, BL21-CodonPlus (DE3)-RIL, BL21(DE3)-T1, Origami, Rosetta, Rosetta-pLysS e Nova Blue (DE3), foram utilizadas em experimentos de transformação celular para encontrar o melhor rendimento da produção em larga escala das proteínas alvo. Inicialmente 80 ng de plasmídeos recombinantes foram adicionados a 50 µL de solução contendo células competentes. Essa mistura permaneceu no gelo durante 30 minutos, e então, foi submetida ao choque de temperatura a 42 ºC por 1 minuto seguida de 2 minutos no gelo. Após esse procedimento, as bactérias foram submetidas a agitação a 250 rpm por 1 hora a 37 ºC em meio líquido SOC. Em seguida, as bactérias foram selecionadas após cultivo em meio sólido seletivo Luria Bertani (LB) contendo canamicina (30 µg/ml) e o antibiótico de resistência de cada linhagem celular. As colônias obtidas foram inoculadas em tubos individuais contendo 10 ml de meio líquido LB seletivo sob agitação e aeração constantes e utilizadas para expressão e purificação. Posteriormente, as amostras proteicas purificadas foram aplicadas em SDS-PAGE 15% para a confirmação da presença das proteínas recombinantes pela metodologia *Western blot*.

3.4 *Western blot*

Para o procedimento, as amostras foram misturadas com tampão de amostra (1:1) e aplicadas em gel SDS-PAGE 15%. As proteínas, contidas no gel, foram transferidas para a membrana de nitrocelulose a partir de um eletrotransferidor (*Mini TransBlot Electrophoretic Transfer Cell Bio-Rad*) operando sob voltagem constante (100 V) por 2 horas. O tampão de transferência utilizado foi o Tris-Glicina (Tris 0,025 M, Glicina 0,192 M) contendo metanol 20%. A membrana foi incubada com antissoro monoclonal Antipoli-histidina diluído 1:10.000 (produzido em camundongo; Sigma) e posteriormente com antissoro secundário anti-IgG de camundongo conjugado com fosfatase alcalina (Sigma) diluído 1:30.000. A detecção da reação antígeno/anticorpo foi feita com a mistura de BCIP/NBT (*5-bromo-4chloro-3-indolylphosphate /nitro-bluetetrazolium*) até o aparecimento do resultado da reação (coloração azul púrpura) que ocorre apenas na presença da fosfatase alcalina.

3.5 Expressão das proteínas alvos

As colônias isoladas com vetores contendo os genes codificadores das proteínas alvos foram cultivadas por 16 horas a 37 ºC em meio LB contendo canamicina (50 mg/ml) e o

antibiótico de resistência de cada linhagem celular. Posteriormente, essas culturas foram diluídas 100 vezes em meio LB fresco contendo os antibióticos e incubadas a 37 ºC até a DO_{600nm} atingir 0,5, obtida em espectrofotômetro Biomate 3S (*Thermo Fisher Scientific*). Após atingir essa densidade, as culturas contendo os vetores dos genes codificadores da $ExhC_{wt}$, suas variantes mutantes, ETA, ETE e ETE_{P186G} foram induzidas com 0,4 mM de IPTG e mantidas a 30 ºC sob agitação durante 5 horas. As culturas foram centrifugadas a 2600 x g durante 10 minutos a 4 ºC, os sobrenadantes foram descartados e as células precipitadas contendo as proteínas alvo foram suspensas em tampão de lise 50 mM Fostato de Sódio; 400 mM de NaCl; 10% de glicerol e 10 mM de Imidazol (pH: 7,5). Após a adição do tampão, as células foram submetidas ao processo de lise celular por sonicação (4 pulsos de 40 segundos), e centrifugadas a 15.000 x g por 30 minutos a 4 ºC. O sobrenadante foi coletado para a realização das etapas de purificação.

3.6 Purificação das proteínas alvos

A primeira etapa de purificação das proteínas recombinantes foi por cromatografia de afinidade em Ni-Sepharose (GE *healthcare*) e posteriormente, em uma segunda etapa, os contaminantes das amostras contendo as proteínas alvos foram removidos por meio da cromatografia de exclusão molecular em coluna Superdex G75 10/300 (GE) utilizando o sistema AKTA-pure.

3.6.1 Cromatografia de afinidade por Ni-Sepharose

A purificação em coluna de afinidade utilizando resina Ni-Sepharose (*GE healthcare*) foi realizada conforme instruções do fabricante. Primeiramente, um volume de 2 ml de resina (Ni-Sepharose – *GE healthcare*) presentes em cada coluna de cromatografia de afinidade, foram equilibradas com o mesmo tampão utilizado no processo de lise das células cultivadas. Em seguida, cada coluna de purificação (coluna de gravidade Bio-Rad) foi carregada com o sobrenadante do lisado celular obtido após centrifugação, para que as proteínas recombinantes fossem adsorvidas. Após adsorção, a resina foi lavada utilizando tampões de lavagem com aumento gradativo da concentração de Imidazol, buscando remover contaminantes adsorvidos de forma inespecífica. Primeiramente, um volume de 10 ml de tampão de lise foi adicionado em cada coluna. O aumento gradativo de Imidazol iniciou-se com uma primeira lavagem de 10 ml de uma solução contendo 50 mM de Fostato de Sódio; 400 mM de NaCl; 10% de glicerol e 40 mM de Imidazol (pH 7,5), seguida de uma segunda lavagem de 10 ml de uma solução

contendo 50 mM de Fosfato de Sódio, 400 mM de NaCl, 10% de glicerol e 60 mM de Imidazol (pH 7,5). A terceira lavagem foi feita com 5 ml de uma solução contendo 50 mM de Fosfato de Sódio, 400 mM de NaCl, 10% de glicerol e 80 mM de Imidazol (pH 7,5). Por fim, para a eluição total das proteínas alvo, foi adicionado 15 ml de uma solução contendo 50 de mM Fostato de Sódio; 400 mM de NaCl; 10% de glicerol e 400 mM de Imidazol (pH 7,5). Todas as amostras obtidas por meio das diferentes lavagens com variações crescentes na concentração de Imidazol, foram coletadas e analisadas em gel de SDS-PAGE 15%, para verificar em qual delas as proteínas alvo foram eluídas.

3.6.2 Cromatografia de exclusão molecular

As amostras contendo as proteínas alvos obtidas a partir da purificação por cromatografia de afinidade, foram concentradas em dispositivos para concentração em centrífuga (*Amicon Ultra Centrifugal Filters* - 10 kDa) e submetidas à segunda etapa de purificação por gel filtração em coluna Superdex G75 10/300 (GE) equilibrada com tampão 20 mM de Tris HCl e 100 mM de NaCl em pH 7,5, utilizando o sistema AKTA-pure (*GE-Healthcare Life Sciences*). As amostras referentes aos picos do cromatograma foram coletadas e analisadas por meio de SDS-PAGE 15%.

Essa etapa de cromatografia também foi utilizada, juntamente com outras metodologias experimentais, para avaliar o estado oligomérico da $ExhC_{wt}$, suas variantes mutantes, ETE e ETE_{P186G}. Amostras contendo 10 mg/mL de proteínas alvo após purificação, foram submetidas a gel filtração com um fluxo de 0,7 mL por minuto usando a coluna Superdex G75 10/300 GL (GE) no sistema AKTA-pure (*GE-Healthcare Life Sciences*) e a absorbância foi monitorada por meio da avaliação do comprimento de onda de 280 nm.

3.7 Diálise das toxinas alvos

A fim de caracterizar biofisicamente a $ExhC_{wt}$ e suas variantes mutantes, assim como a ETE, ETE_{P186G} e ETA, após as etapas de purificação, as amostras das proteínas alvo, foram concentradas em dispositivos e dialisadas em diferentes soluções tampões (Tabela 2) em uma temperatura constante de 4 ºC, utilizando a membrana de diálise a Ficherbrand (MWCO 6.000-8.000). Primeiramente, foram realizadas duas diálises de duas horas cada, e posteriormente, essas amostras foram dialisadas por 16 horas. As variações nas soluções tamponantes foram necessárias de acordo com as condições dos experimentos realizados (Tabela 2).

Tabela 2. Tampões utilizados para a diálise das toxinas alvo de acordo com condições experimentais. Cromatografia de exclusão molecular (SEC); Dicroísmo Circular (CD); Espalhamento Dinâmico de Luz (DLS) e Viabilidade celular (ensaios de MTT).

Solução tamponante	Sal	pH	Experimento(s)
20 mM Tris Hidrocloreto	100 mM de NaCl	7,50	SEC; DLS
5 mM Dihidrigenofosfato de Sódio	50 mM de NaF	7,40	CD
10 mM Na_2HPO_4 + 1,8 mM KH_2PO_4	150 mM de NaCl	7,40	Viabilidade celular

Fonte: própria autora.

3.8 Análise de pureza e estado oligomérico das toxinas por gel de poliacrilamida

Para a verificação das etapas de purificação foi realizada a eletroforese em gel de poliacrilamida 15% (SDS-PAGE 15%), possibilitando a separação das proteínas de acordo com a massa molecular. A eletroforese foi conduzida por aproximadamente 90 minutos a 100 V no aparelho *Protean* II (Bio-Rad). Após esse procedimento, o gel foi fixado e corado em mistura contendo metanol, água, ácido acético (50:50:10) e *Comassie Brilliant Blue* 0,25% e posteriormente descorado com metanol e água (1:1) para visualização do resultado.

Além disso, após a etapa de purificação por cromatografia de exclusão molecular, também foram realizadas eletroforeses Blue-Native-PAGE, possibilitando a avalição do estado oligomérico das proteínas $ExhC_{wt}$ (i), $ExhC_{mut4}$ (ii), ETE (iii) e ETE_{P186G} (iv) sem a desnaturação proteica. Para a realização desse experimento, utilizou-se o protocolo previamente estabelecido (Westermeier, 2016), sendo que 2 mg/ml das proteínas alvo foram incubadas com *Comassie Brilliant Blue* 0,25%, para a aquisição de carga negativa pelas proteínas recombinantes, devido à ausência de detergente SDS no tampão de corrida (25 mM Tris e 192 mM glicina, pH 8,3) e no gel de poliacrilamida 15%. A eletroforese ocorreu em ambiente de 4ºC, com o objetivo de evitar a desnaturação proteica durante o experimento, e foi conduzida por aproximadamente 90 minutos a 100 V no aparelho *Protean* II (Bio-Rad). As proteínas ß-tripsina de pâncreas bovino (SigmaAldrich) com aproximadamente 24 kDa, e a albumina de soro bovino (SigmaAldrich), com aproximadamente 66 kDa, ambas já descritas como monômeros em solução (Palmer; Bonner, 2007) foram utilizadas como amostras controles, ou seja, como marcadores de peso molecular.

3.9 Espalhamento dinâmico de luz

Para realização dos experimentos, amostras das proteínas (50 µM) em tampão 20 mM de Tris HCl e 100 mM de NaCl em pH 7,5, foram previamente centrifugadas durante 30 minutos a 14.000 x g em temperatura de 10 ºC (Karow; Gotzl; Garidel, 2015). As medidas de DLS foram

realizadas utilizando o aparelho Zetasizer Nano series S90 (*Malvern Instruments*) e a aquisição de dados foi feita a partir de 10 ciclos de 30 segundos em temperatura constante de 37 °C. O cálculo da D_H foi realizado utilizando o software Zetasizer.

3.10 Dicroísmo circular

Os espectros UV-CD de ExhC e suas variantes, ETA e ETE foram obtidos utilizando um espectropolarímetro Jasco J-710 (Jasco, Brasil) a uma temperatura constante de 25 ºC usando uma cubeta de quartzo de 0,5 mm. As amostras de proteínas foram diluídas para 9-11 µM em 5 mM NaH_2PO_4 e 50 mM NaF (pH 7,4). Para a coleta de dados foi utilizada a faixa espectral de 190-260 nm com taxa de varredura de 50 nm/min, tempo de resposta de 1,0 segundo, largura de banda espectral de 1 nm, resolução espectral de 0,2 nm e 10 acumulações. Após a obtenção dos dados, todos os espectros de CD registrados foram corrigidos subtraindo-se o espectro da solução tampão obtida em condições idênticas, e convertidos em elipticidade molar na unidade de 10^3.deg.cm^2.dmol^{-1} pelo software OriginLab 7. As porcentagens de estruturas secundárias foram estimadas (Greenfield, 2006) pelo software CONTINLL do pacote CDPro, utilizando o conjunto de referência de proteínas SP29 (Sreerama; Woody, 2004).

3.11 Linhagens celulares e condições de cultura

As linhagens celulares de fibroblastos renais de hamsters recém-nascidos (BHK-21) e de carcinoma de células renais epiteliais humanas (786-O) foram cultivadas em meio *Dulbeccos's Modified Eagle Medium* (DMEM-Vitrocell) e em meio RPMI1640 (GIBCO), respectivamente. Já as linhagens celulares de câncer renal ACHN (ATCC: CRL-1611) e a de rim de porco PK-15 (ATCC: CCL-33) foram cultivadas em meio EMEM *Modified Eagle Medium* (MEM - Cultilab, Brasil). Esses meios foram suplementados com 10% de soro fetal bovino (Vitrocell), 1% de aminoácidos não essenciais (GIBCO) e 200 U/ml de penicilina e estreptomicina (Vitrocell). A linhagem celular epitelial renal derivada de rins normais (HK-2) foi cultivada em meio sem soro de queratinócitos K-SFM (*Thermo Fisher Scientific*), o qual foi suplementado com 5% de soro fetal bovino (Vitrocell), 25 mg/ml de extrato de hipófise bovina (*Thermo Fisher Scientific*) e 2,5 µg/ml de fator de crescimento epitelial recombinante humano (*Thermo Fisher Scientific*).

Todas as linhagens celulares foram mantidas sob condições de 37 ºC em atmosfera úmida com 5% de CO_2 até atingir a confluência celular. Posteriormente, os meios de cultura foram removidos e as células foram individualizadas com solução Tripsina-EDTA (0,25% ou

0,05%) (GIBCO) para a realização da contagem celular. As soluções contendo Tripsina-EDTA (0,25% ou 0,05%) foram removidas após a centrifugação a 16.000 x g durante 5 minutos e, posteriormente, novos meios de cultura foram adicionados para o plaqueamento em placas de cultivo e realização dos ensaios experimentais.

3.12 Ensaios de viabilidade celular

As toxinas $ExhC_{wt}$ e suas variantes mutantes, ETA e ETE diluídas em PBS 1X foram submetidas a testes *in vitro* com a BHK-21, linhagem celular utilizada nos experimentos de Li et. al (2011ª), que sofreu necrose celular na presença da $ExhC_{wt}$ de *M. sciuri*. Testes *in vitro* com as linhagens 786-O, PK-15, ACHN e HK-2 foram realizados apenas com a $ExhC_{wt}$, para avaliar a capacidade dessa proteína em causar diminuição da viabilidade celular em outras linhagens celulares. O perfis de citotoxicidade foram mensurados por meio do método de MTT (*3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyl tetrazolium bromide*) (Sigma- Aldrich, Missouri, USA). Para cada ensaio, vinte e quatro horas antes do tratamento 1×10^4 de células foram plaqueadas em placas de cultivo de 96 poços. No dia seguinte, adicionou-se novo meio de cultivo contendo cada uma das proteínas de estudo em uma concentração máxima inicial de 60 µM com decaimento em $\log_2$ até atingir concentração mínima de 0,94 µM, garantindo que a concentração de 15 µM utilizada por LI et al. (2011ª) fosse novamente testada. Amostras controle de células sem as toxinas alvo na presença e ausência do tampão PBS 1X também foram avaliadas. As células foram incubadas por 48 horas e cada experimento foi realizado em triplicata e em três ensaios independentes. Posteriormente, o meio de cultivo contendo as proteínas foi removido de cada poço e substituído por solução MTT (1 mg/ml em 100 µL de meio de cultura sem suplementação) (Sigma- Aldrich, Missouri, USA), na qual as células foram incubadas por 30 minutos a 37 ºC. Após a formação dos cristais de formazan, a solução MTT foi removida e adicionou-se 100 µL de DMSO (Synth São Paulo, Brasil) em cada poço. Em seguida, a absorbância foi medida a 572 nm, utilizando um espectrofotômetro de placa (FLUOstar Omega/BMG LABTECH, Offenburg, BW, DE).

3.13 Ensaios de apoptose

Para verificar se a morte das células BHK-21 causada pela ExhC é devido a necrose, realizou-se experimentos com o inibidor irreversível da pan-caspase Z-VAD-FMK (Promega). Primeiramente, 1×10^4 células/poço em placas de 96 poços foram previamente tratadas durante 2 horas com 50 µM de Z-VAD-FMK (Promega) e depois incubadas com 60 µM de $ExhC_{wt}$

durante 24 horas a 37 °C e 5% de CO_2 (Li; Wang; Ding; Zheng, 2011). O controle positivo foi $1x10^4$ células/poço incubadas com 100 µM de peróxido de hidrogênio, que induz apoptose em células BHK-21 (Burdon; Gill; Alliangana, 1996). Os experimentos foram executados em triplicata e três ensaios independentes. Após os tratamentos, foram realizados ensaios de viabilidade celular conforme descrito na seção 3.13.

3.14 Análises estatísticas

Os ensaios de viabilidade celular expressos como média ± desvio padrão foram analisados estatisticamente utilizando o software GraphPad Prism 8 (GraphPad Software Inc., San Diego, CA, EUA). Com base no teste de normalidade D'Agostino & Pearson, os dados foram classificados em paramétricos ou não paramétricos, e os testes de Dunnett e Kruskall Wallis Dunn foram utilizados para análise dos dados paramétricos e não paramétricos, respectivamente. Valores de $P < 0,01$ foram considerados estatisticamente significativos.

3.15 Ensaios enzimáticos

Apesar da alta especificidade das ETs às Dsgs-1, experimentos com o objetivo de avaliar a atividade catalítica das ETs indicaram que ETA e ETB apresentam atividade estereolítica moderada contra o substrato Boc-L-Glu-↓OPhe (Bailey; Redpath, 1992; Cavarelli; Prevost; Bourguet; Moulinier et al., 1997), o qual quando hidrolisado pelas toxinas, libera um gurpo fenol que absorve em 270 nm. Além disso, essas toxinas são conhecidas por clivar alguns oligopeptídeos, por exemplo, hormônios estimuladores de melanócitos α e β (α-MSH e β-MSH) (Rago; Vath; Bohach; Ohlendorf et al., 2000). Em conjunto, esses dados sugerem que as ETs podem apresentar uma atividade hidrolítica reduzida contra algumas moléculas pequenas, sem a necessidade de ativação prévia. A partir desses dados, o substrato Boc-L-Glu-↓OPhe foi utilizado em ensaios enzimáticos com a $ExhC_{wt}$, $ExhC_{mut4}$, ETE_{wt}, ETE_{P186G}. Primeiramente, 1 M de Boc-L-Glu-↓OPhe foi diluído em 1,4-Dioxano (Sigma) e utilizado como solução estoque de substrato. As reações estereolíticas foram realizadas em condição de 50 mM de HEPES (pH 7,4) e 2% de 1,4-Dioxano (v/v) utilizando 1 µg/ml de toxina, variando a concentração de substrato entre 1-20 mM, com decaimento na concentração a cada 5 mM. Todas as concentrações de substrato testadas foram realizadas em ensaios independentes e triplicata. A medição da atividade estereolítica do NaOH de Boc-L-Glu-↓OPhe foi determinada a partir do coeficiente de absorção molar de 1,5 mM/cm em 270 nm (Cavarelli; Prevost; Bourguet; Moulinier et al., 1997), utilizando o espectrofotômetro Cary UV-Vis Compact Peltier (Agilent),

regulado para manter a temperatura constante de 37 ºC durante todos os ensaios enzimáticos. As leituras foram feitas a cada 15 segundos durante 20 minutos. Os valores de V_o, KM, K_{cat} foram estimados a partir da equação de Hanes-Woolf, método usado para linearização da hipérbola de Michaelis-Menten. A eficiência catalítica (K_{cat}/KM) foi calculada apenas a partir da equação de Hanes-Woolf e não pela equação de Michaelis-Menten pois, devido a baixa solubilidade do substrato, não foram atingidas concentrações de saturação enzimática o que impediu a obtenção da hipérbole completa de Michaelis-Menten.

3.16 Phage display

3.16.1 Biblioteca TriCO-16

A exibição de fagos com a biblioteca TriCo-16 (Creative Biolabs, Shirley, EUA) foi realizada utilizando a ExhC. Experimentalmente, foram feitos em três rodadas de seleção para selecionar apenas os peptídeos de alta afinidade de ligação ao alvo. Resumidamente, para cada etapa de seleção, poços de uma placa de alta ligação – 3D-NHS (PolyAn – Molecular Surface Engineering) foram sensibilizados por 30 minutos com 30 pM de proteína dissolvida em 100 µL de tampão acetato de sódio 10 mM pH 5,2. Em seguida, a solução contendo o alvo foi removida e 100 µL de tampão *Quenching* (etanolamina*HCl 1 M pH 8,5) foi adicionado por 30 minutos. Posteriormente, esta solução foi descartada e 100 µL de tampão de bloqueio (HEPES 25 mM, NaCl 150 mM, pH 7,5 e BSA 1%) foi adicionado por 30 minutos. O descarte do tampão de bloqueio foi seguido por três lavagens usando 200 µL de tampão de lavagem (HEPES 25 mM, NaCl 150 mM, pH 7,5 e BSA 0,1%). Em seguida, a ExhC foi incubada com 100 µL de $1{,}66x10^{11}$ da biblioteca TriCo-16 (Creative Biolabs, Shirley, EUA) por 30 minutos. Após a incubação, iniciou-se a etapa de *biopanning*: foram realizadas cinco lavagens com tampão de lavagem para cada poço, no primeiro ciclo de seleção, enquanto no segundo e terceiro ciclos foram realizadas 15 lavagens em cada poço. O objetivo desta etapa foi garantir a remoção de fagos que não interagiam com o alvo e os fagos de alta ligação foram eluídos pela incubação com 100 µL de solução 0,2 M de glicina-HCl pH 2,2 por 10 minutos. A neutralização desta solução de baixo pH foi realizada com 25 µL de 1M de Tris pH 9,0, o que garantiu a obtenção de 125 µL do eluato não amplificado da ExhC (TS). Todas as etapas foram realizadas à temperatura ambiente (TA).

Durante o primeiro ciclo de seleção, $1.66x10^{11}$ da biblioteca TriCo-16 foi utilizado como controle experimental. Esse controle foi denominado *empty control* (ES) e no segundo e terceiro ciclos esse controle foi realizado com os fagos amplificados do *empty control* do

primeiro ciclo (ES1) e com os fagos amplificados do *empty control* do segundo ciclo (ES2), respectivamente. Além disso, um segundo controle foi adicionado às etapas experimentais do segundo ciclo de seleção, que foi denominado *direct control* (DC). Esse controle foi produzido utilizando os fagos amplificados que eluíram dos poços contendo a ExhC, com o objetivo de avaliar, logo após a etapa de sequenciamento, quais peptídeos interagem apenas com a ExhC e quais interagem com os componentes da placa de alta ligação.

3.16.2 Amplificação dos fagos

A amplificação dos fagos selecionados foi realizada utilizando a *E. coli* - ER2738 em 20 ml de meio de cultura LB contendo tetraciclina (20 mg/ml), individualmente para cada eluato não amplificado (TS1, TS2, TS3, ES1, ES2, ES3, DC2 e DC3). Após atingir uma DO_{600} de 0,1, 120 µL de cada eluato não amplificado foram adicionados às culturas bacterianas e incubados por 4 horas sob agitação constante a 37 ºC. Em seguida, as culturas foram centrifugadas a 4°C a 5.000 rpm por 20 minutos e os sobrenadantes foram transferidos para tubos estéreis contendo uma solução de 7 ml de PEG 8.000 a 25% e NaCl 2,5 M (PEG/NaCl). As soluções foram incubadas durante a noite a 4°C para precipitação dos fagos amplificados e depois centrifugadas a 5.000 rpm durante 1 hora a 4°C para descartar os sobrenadantes.

Os fagos foram ressupendidos em 1 ml de tampão de seleção (HEPES 25 mM e NaCl 150 mM, pH 7,5) e precipitados novamente a partir da adição de 200 µL de PEG/NaCl por 1 hora no gelo. Posteriormente, as amostras foram centrifugadas a 5.000 rpm a 4°C e os sobrenadantes foram descartados. Os precipitados foram ressuspendidos em 100 µL do tampão de seleção, obtendo-se assim os fagos amplificados. A concentração de cada fago amplificado foi medida com espectrofotômetro UV-VIS UV-1800 (Shimadzu Deutschland GmbH) usando absorbância de 220 e 340 nm. Foram realizadas diluições dos fagos amplificados para garantir uma concentração de $1.66.10^{11}$, para ser utilizada no segundo e terceiro ciclo de seleção da ExhC.

3.16.3 Titulação dos fagos

Para avaliar se os peptídeos acoplados aos capsídeos virais estavam interagindo com o alvo ExhC (TS) durante as três etapas de seleção de exibição de fagos com a biblioteca TriCo-16, 1 µL de cada o eluato obtido foi utilizado em experimentos de titulação entre 10^{-3} e 10^{-11}. Experimentos controle também foram avaliados utilizando esta técnica (ES e DC). Diluições em série foram incubadas com 100 µl de *E. coli* - ER2738 em 0,6 OD_{600} e plaqueadas em meio

LB contendo IPTG (0,5 mM) e X-gal (40 µg/ml) juntamente com 800 µL de Agarose Top. Após incubação em estufa por 16 horas a 37 ºC, as colônias azuis foram contadas para avaliar o sucesso experimental.

3.16.4 Ensaio de imunoabsorção enzimática (ELISA)

Ensaios de ELISA foram realizados com os fagos amplificados obtidos após a eluição do primeiro, segundo e terceiro estágios de seleção com o alvo ExhC. O objetivo foi verificar a interação da ExhC com os peptídeos selecionados em cada etapa da seleção da exibição do fago. Para isso, oito poços foram separados e denominados: TS1+, TS2+, TS3+, TS1-, TS2-, TS3-, C+ e C-. Primeiramente, 30 pM de ExhC foram diluídos em 10 mM de acetato de sódio (pH 5,2), respectivamente, e adicionados aos poços com sinal positivo (+), enquanto apenas os diluentes foram adicionados aos poços com sinal negativo (-). Estas soluções foram incubadas durante uma hora à temperatura ambiente. Em seguida, a solução contendo os alvos foi removida e 100 µL de tampão *Quenching* foram adicionados a cada poço e incubados por 30 minutos. Posteriormente, essas soluções foram descartadas e 100 µL de tampão de bloqueio (HEPES 25 mM, NaCl 150 mM, pH 7,5 e BSA 1%) foram adicionados por 30 minutos.

O descarte do tampão de bloqueio foi seguido por três lavagens usando 200 µL de tampão de lavagem (HEPES 25 mM, NaCl 150 mM, pH 7,5 e BSA a 0,1%) para cada lavagem. Posteriormente, os poços denominados TS1, TS2 e TS3 foram incubados durante 1 hora com 100 µL de $1.66x10^{11}$ fagos amplificados obtidos a partir das eluições de TS1, TS2 e TS3 durante o phage display. Durante essa etapa, os poços de controle na ausência de fagos (C+ e C-) foram completados com 100 µL de tampão de seleção. Ao final deste processo, foram realizadas cinco lavagens com 200 µL de tampão de lavagem para cada poço e, em seguida, 100 µL de anticorpo bacteriófago M13 1:2500 (11973-MMO5T-H) por poço foram incubados por 1 hora e posteriormente adicionados 100 µL de uma solução 0,1 mg/ml de 3,3',5,5'-Tetrametilbenzidina (TMB), Fosfato-Citrato e 10% de DMSO. Imediatamente, foi observada uma mudança de cor devido à interação da reação do TMB com o anticorpo bacteriófago M13. A continuidade da reação foi inibida com a adição de 100 µL de 2M de H_2SO_4. Finalmente, a absorbância foi medida por um leitor de placas a 450 nm e os resultados foram plotados utilizando o software GraphPad Prism 8.

3.16.5 Extração de DNA e PCR

Dez microlitros dos fagos amplificados (TS1, TS2, TS3, ES1, ES2, ES3, DC2 e DC3), obtidos a partir de experimentos de exibição de fagos com a ExhC, foram misturados com 90

µl de TBS 1X e 200 µL de PEG/NaCl e incubados por 20 minutos em temperatura ambiente. Posteriormente, as amostras foram centrifugadas a 14.000 rpm durante dez minutos a 4°C e o sobrenadante foi completamente removido. O palete obtido de cada amostra foi ressuspendido em 200 µL de 3M de acetato de sódio e 0,1 X TAE, em seguida foram adicionados 500 µL de etanol 99% e as amostras foram incubadas por 15 minutos nesta solução. Novamente, as amostras foram centrifugadas a 14.000 rpm durante dez minutos a 4°C e o sobrenadante foi completamente removido. Após a retirada do sobrenadante, foram adicionados 250 µL de etanol 70% por tubo e, após centrifugação nas mesmas condições já descritas, o mesmo foi removido. A partir da extração do DNA, as amostras foram secas a 30 °C e 40 µL de água milliQ foram adicionados a cada amostra para diluir o ácido nucleico. A concentração de DNA de cada extração foi medida utilizando o NanoFotômetro (IMPLEN).

Para a amplificação do DNA viral obtido na extração foi realizada a técnica de PCR. Cada reação de PCR foi preparada a partir de: 12,5 µL Kapa HiFi HotStart ReadyMix (Roche), 5 µM de primer reverse (TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGC CCTCATAGTTAGCGTAAC), 5 µM de primer forward (TCGTCGGCAGCGTCAGAT GTGTATAAGAGACAGCGCAATTCCTTTAGTGGTACC) e 25 ng de DNA. Todas as reações foram completadas com água milliQ em um volume final de 25 µL. O termociclador utilizado Thermal Cycler C1000 Touch (BIO-RAD) para as reações de PCR foi programado nas seguintes condições:

Steps	Cycles	Temperature (°C)	Minutes
Initial denaturation	Start	95	3
Denaturation	25x	95	0.5
Annealing	25x	55	0.5
Elongation	25x	72	0.5
Final elongation	End	72	5
Storage	Hold	4	-

Posteriormente a realização do PCR, a concentração de cada amostra foi medida pelo NanoFotômetro (IMPLEN), e em todos os casos foram obtidas concentrações superiores a 500 ng/µL, sendo suficientes para a realização do sequenciamento. Além disso, as amostras foram avaliadas utilizando gel de agarose a 1%. O sequenciamento de nova geração (NSG) foi realizado pela Universität Düsseldorf (HHU).

3.16.5 Análises das sequências peptídicas

Sequências de DNA resultantes do sequenciamento de nova geração (NGS) foram transcritas em sequências peptídicas, conforme descrito por SATUR et al. (2021). As

sequências transcritas foram filtradas com base em três parâmetros: seu aumento de frequência relativa dentro do TS (biblioteca TriCO-16 < TS1 < TS2 < TS3) (i), a frequência relativa de uma sequência observada em um TS em comparação com sua abundância dentro do respectivo DC (ii) e a frequência relativa de uma sequência observada em um TS em comparação com sua abundância dentro do respectivo ES (iii). As sequências obtidas no TS3 que passaram pela filtragem foram classificadas de acordo com seu *enrichment factor* (TS3/biblioteca = *enrichment factor*) e *empty score* (TS3/ES3 = *empty score*). As sequências filtradas também foram adicionadas ao software de agrupamento Hammock (Krejci; Hupp; Lexa; Vojtesek et al., 2016). Cluster *motif* logos para o *phage display* feito com a ExhC foram gerados usando o aplicativo WebLogo 3.4 (Crooks; Hon; Chandonia; Brenner, 2004).

Após análises, sete peptídeos foram selecionados para ensaios primários de interação com ExhC de acordo com: valores de *empty score* (1) e *enrichment factor* (2), assim como sua similaridade com o substrato original (desmoglein-1 suína) (3) e presença dos resíduos frequentemente observados nos Cluster *motif* logos (4). Parâmetros físico-químicos também foram investigados, tais como ponto isoelétrico (5) e solubilidade dos peptídeos (6), com o objetivo de obtermos maior sucesso nas etapas experimentais.

3.17 Ensaios de interação com peptídeos

Anteriormente aos ensaios de interação, ensaios de atividade enzimática com a ExhC foram reformulados, utilizando Dimetilsulfóxido de dimetilo (DMSO) (v/v) ao invés de 1,4-Dioxano (v/v) para a diluição do substrato Boc-L-Glu-↓OPhe. Essa mudança foi realizada, visto que a maioria dos peptídeos selecionados (descrição item 3.17.5) eram solúveis em DMSO de acordo com o fabricante (Genscript).

Primeiramente, 1M de Boc-L-Glu-↓OPhe foi diluído em DMSO (Sigma) e usado como solução estoque de substrato. As reações enzimáticas foram realizadas sob condições de 20 mM de HEPES pH 7,4 (i) e 20 mM de MES pH 6,5 (ii) contendo 2% de DMSO (v/v) e utilizando 10 µg/mL de toxina, variando a concentração do substrato entre 0 a 20 mM, com passos de 5 mM. Todas as concentrações de substrato foram testadas em triplicada. A liberação de fenol (coeficiente de absorção molar = 1,5 mM/cm) do substrato sintético Boc-L-Glu-↓OPhe (λ = 270 nm) foi medida durante 30 minutos, a cada 30 segundos, a uma temperatura constante de 37 ºC, utilizando o leitor de placas Infinite 200 PRO (Tecan, Männedorf, Suíça). Os valores de Vo, KM, Kcat e a eficiência catalítica (K_{cat}/KM) foram estimados a partir da equação de Michaelis-Menten.

Após essa etapa, ensaios primários para avaliar a interação entre ExhC-peptídeos foram realizados com 10 µg/ml de $ExhC_{wt}$ e 2,5 mM de Boc-L-Glu-↓OPhe em tampão 20 mM de HEPES (pH 7,4) (i) e 20 mM de MES (pH 6,5) (ii), contendo 2% de DMSO (v/v). Todos os peptídeos foram adicionados em uma concentração final de 40 µM. Foram testados três períodos de incubação (10 minutos, 30 minutos e 2 horas) entre os peptídeos e a toxina. Após a incubação, o substrato sintético foi adicionado a cada solução e a absorbância de 270 nm foi medida em intervalos de 30 segundos durante 30 minutos utilizando o leitor de placas Infinite 200 PRO (Tecan, Männedorf, Suíça) a 37 °C. Todos os ensaios de inibição foram realizados em triplicata.

3.18 Análises por *Basic Local Alignment Search Tool* (BLAST)

Sabendo-se que a ExhC é uma proteína *moonlight* (Singh; Bhalla, 2020), com atividade necrótica, esfoliativa (Li; Wang; Ding; Zheng, 2011) e de inibição do sistema imune inato do hospedeiro (Li; Li; Lu; Wang et al., 2011), o objetivo desta etapa *in silico* foi buscar potenciais proteínas parceiras envolvidas nesses mecanismos de ação. Para isso, as sequências peptídicas obtidos por phage display com maiores valores de *Empty score* e *Enrichment factor* (Tabela 3) foram avaliadas pela ferramenta BLAST *protein* (blast.ncbi.nlm.nih.gov/Blast.cgi). Nessa análise, apenas proteínas de organismos hospedeiros de *M. sciuri* foram selecionadas para a busca: *Sus scrofa domesticus*, *Mus musculus* e *Mesocricetus auratus*. As proteínas com maior identidade com os alvos peptídicos foram investigadas de acordo com a suas funções e estrutura. Por fim, selecionou-se a Pcyox1l de *Mus musculus* para análises computacionais de interação proteína-proteína (descrição no item 3.22.4).

Tabela 3. Peptídeos com maiores valores de *empty score* e *enrichment factor* selecionados pelo phage-display realizado com a ExhC.

Sequence	Empty score	Enrichment factor
WAQMPQPFWKHSPFER	1253.91	1358.46
YTSPWWYYQMDEHMQG	353.55	47.88
DMQYWLGDSQQTRIS	277.53	300.67
VMPWSSNVYGYMPNYR	271.12	146.86
VSWSYWDPIHPWQKRDR	223.49	121.06

3.19 Cristalização, coleta de dados, processamento e determinação de estruturas

Os cristais de proteína foram obtidos por difusão de vapor e concentrações proteicas de 15 mg/mL para $ExhC_{wt}$ (i) e 10 mg/mL para $ExhC_{mut4}$ (ii) e 30 mg/mL para ETE (iii) em tampão

Tris-HCl 20 mM, pH 7,5 contendo 100 mM de NaCl, equilibrado contra soluções reservatório de HEPES 0,1M pH 7,5; Tartarato de Potássio/Sódio 0,8 M (i), $MgCl_2.6H_2O$ 0,2 M; Tris-HCl 0,1 M pH 8,5 e PEG 4.000 a 30% p/v com adição de Sarcosina 0,1 M (ii) e Formato de Sódio 4 M, pH 6,0 (iii), respectivamente. Dados de difração de raios-X foram coletados de cristais congelados na linha de luz MANACA (SIRIUS, Campinas, Brasil). As 3600 imagens coletadas de cada cristal foram indexadas usando o pacote XDS (Kabsch, 2010). As estruturas cristalinas de $ExhC_{mut4}$ e ETE foram resolvidas por substituição molecular empregando as coordenadas atômicas de ETD-like (PDB ID: 5C2Z) (Mariutti; Souza; Ullah; Caruso et al., 2015). Após o refinamento, a estrutura de $ExhC_{mut4}$ foi utilizada para determinar a estrutura da $ExhC_{wt}$ por substituição molecular. Os parâmetros estereoquímicos foram avaliados usando o servidor MolProbity (Chen; Arendall; Headd; Keedy *et al.*, 2010). As coordenadas atômicas $ExhC_{wt}$, $ExhC_{mut4}$ e ETE foram depositadas no Protein Data Bank (PDB) e receberam os códigos de acesso 8T3J, 8T3I e 8DAX, respectivamente.

3.20 Espalhamento de raios-X em baixos ângulos

Experimentos de espalhamento de raios-X em baixos ângulos **(**SAXS) foram realizados para avaliar as propriedades estruturais da ETE em solução próxima às condições fisiológicas em termos de concentração de salina e pH. Os testes foram realizados utilizando o sistema SAXS de bancada Xeuss 2.0 da empresa XENOCS, equipado com uma fonte de raios-X - Xenocs Genix 3D com ânodo de cobre, o qual produz um feixe com comprimento de onda de λ = 1,5419 Å. O sistema de colimação, composto por um espelho monocromador e um conjunto de fendas sem dispersão, produz um feixe de aproximadamente $0,7 \times 0,7$ mm^2. O sistema utiliza um detector Pilatus 300K Dectris para todas as medições realizadas com amostras sob condições de vácuo. A distância amostra-detector foi calibrada por meio de medições de um padrão behenato de prata (AgBeh) que possui picos na região SAXS fornecendo uma distância amostra-detector de 1193 ± 8 mm. Para medição, utilizou-se um volume de 100 µL de ETE a uma concentração de 5 mg/mL em 20 Mm de MES pH 7,0 e 150 Mm de NaCl. Os dados SAXS foram obtidos a partir de 7 ciclos com tempos de exposição de 1800 s a uma temperatura constante de 25 °C. As imagens 2D SAXS foram integradas ao programa Fit2D (Hammersley, 2016). As integrações de dados foram processadas e normalizadas em escala absoluta com o programa SUPERSAXS (Ezquerra; Garcia-gutierrez; Nogales; Gomez, 2009). As curvas tratadas foram analisadas com o software ATSAS (Petoukhov; Franke; Shkumatov; Tria et al.,

2012), permitindo-nos obter curvas de distribuição de distâncias pareadas e analisar a presença de estados oligoméricos.

Os modelos de ETE preditos por ensaios *in silico* (descrição item 3.22.2) foram ajustados aos perfis de intensidade obtidos por SAXS. Para avaliar a presença de monômeros, homodímeros ou mesmo misturas no sistema, utilizou-se o programa Oligomer (Petoukhov; Franke; Shkumatov; Tria et al., 2012), que avaliou a qualidade do ajuste através de valores de qui-quadrado ($\chi 2$) e estimou a porcentagem de homodímeros associados a cada modelo analisado. Foi permitido adicionar uma constante durante o ajuste dos dados. Em princípio, os modelos homodímeros que produziram perfis de intensidade teórica com valores mais baixos de $\chi 2$ em relação à curva experimental têm maior probabilidade de estar próximos da conformação real da ETE em solução (Ezquerra; Garcia-Gutierrez; Nogales; Gomez, 2009).

3.21 Análises computacionais das toxinas esfoliativas

3.21.1 *Proteins, Interfaces, Structures and Assemblies* (PDBePISA) e CavityPlus

Como foi observado que a $ExhC_{wt}$ e suas variantes mutantes, assim como a ETE_{wt} são homodímeros em solução, o servidor PDBePISA (www.ebi.ac.uk/msdsrv/prot_int/pistart.html) que dispõem de um método padrão para a classificação de interfaces entre proteínas (Krissinel; Henrick, 2007), foi utilizado para avaliar se a disposição dos oligômeros observados nas unidades assimétricas das estruturas cristalinas (descrição item 3.20) eram energeticamente favoráveis. A mesma avaliação foi realizada para a interface do homodímero de ETE proposto por ferramentas computacionais (descrição item 3.22.2). A ferramenta CavityPlus (www.pkumdl.cn:8000/cavityplus/index.php) (Xu; Wang; Hu; Gao et al., 2018) foi usada para investigação de cavidades superficiais na ExhC.

3.21.2 Predição homodímeros

Cada monômero da estrutura cristalina ETE (PDB: 8DAX) foi submetido para *ab initio docking* nos servidores web de cinco algoritmos diferentes de acoplamento proteína-proteína: ClusPro (cluspro.bu.edu/) (Kozakov; Hall; Xia; Porter et al., 2017), HDOCK (hdock.phys.hust.edu.cn/) (Yan; Zhang; Zhou; Li et al., 2017), HADDOCK2.4 (bianca.science.uu.nl/haddock2.4/) (Dominguez; Boelens; Bonvin, 2003), GalaxyTongDock (GalaxyWeb: galaxy.seoklab.org/) (Ko; Park; Heo; Seok, 2012; Park; Baek; Lee; Seok, 2019) e LZerD (lzerd.kiharalab.org/) (Christoffer; Bharadwaj; Luu; Kihara, 2021), na ordem acima e empregando em cada caso em parâmetros padrões do programa. Vale ressaltar que as

simulações de acoplamento HADDOCK2.4 foram restritas a uma região de proteína que foi prevista conter resíduos formadores de interface de acordo com o CPORT (alcazar.science.uu.nl/services/CPORT/) (De Vries; Bonvin, 2011). As poses de ancoragem LZerD foram geradas usando duas abordagens diferentes, uma delas envolvendo ancoragem simétrica C2 e a outra não assumindo simetria específica. Além disso, modelos baseados em homologia da estrutura do homodímero foram obtidos no site GalaxyWeb (Ko; Park; Heo; Seok, 2012; Park; Baek; Lee; Seok, 2019). As melhores poses não redundantes, ou seja, aquelas com RMSDs pareados em relação aos seus átomos pesados > 4 Å entre as dez soluções mais bem classificadas geradas por cada algoritmo de acoplamento, foram selecionadas para etapas de refinamento. Essas etapas consistem em simulações MD, bem como MM-GBSA e cálculos de US de energia livre, a fim de prever a conformação mais provável dos dímeros de ETE. A técnica SAXS (descrição item 3.21) foi empregada como critério adicional para seleção da conformação do homodímero. Os modelos que produziram valores mais baixos de $\chi 2$ e, ao mesmo tempo, apresentaram valores favoráveis nos cálculos US de energia livre, foram selecionados para as próximas etapas, que consistiram em simulações MD de 1 μs de comprimento e cálculos adicionais de energia livre.

Além do uso da ferramenta PISA (descrição item 3.22.1), a estrutura central coletada da em simulações MD de 1 μs foi submetida a análises bioinformáticas adicionais para avaliar se corresponde ou não a uma montagem biológica, como o servidor web PRODIGY-CRYSTAL (bianca.science.uu.nl/prodigy/cryst) (Jimenez-Garcia; Elez; Koukos; Bonvin et al., 2019). Além disso, o SPPIDER foi empregado para prever se as conformações dos dímeros continham IFRs (Porollo; Meller, 2007). O SPPIDER faz uso de acessibilidade a solventes, conservação de resíduos de aminoácidos, conservação de carga, conservação de tamanho de resíduos de aminoácidos, contatos e hidrofobicidade para alimentar uma rede neural para classificar resíduos na interface ou superfície livre. O modelo finalmente proposto como a conformação mais confiável de ETE em solução foi aquele que passou por todas as etapas do protocolo elaborado.

3.21.3 Cálculo do potencial eletrostático superficial e hidrofobicidade da ExhC e toxinas esfoliativas homólogas

As estruturas cristalinas de $ExhC_{wt}$ e $ExhC_{mut4}$ e os modelos AlphaFold2 (Jumper; Evans; Pritzel; Green et al., 2021) de cada variante mutante de ExhC, foram usados para calcular os potenciais eletrostáticos de superfície com Pymol 1.8.4.0 (Schrödinger; Delano,

2020) e APBS Tools2.1 (Baker; Sept; Joseph; Holst et al., 2001). Parâmetros padrão, incluindo uma concentração de 0,15 M de NaCl, foram definidos para os cálculos. A análise anterior foi estendida para incluir outras ETs estafilocócicos homólogas, cujas estruturas foram recuperadas do PDB – como por exemplo a ETE (PDB ID: 8DAX) – ou geradas pelo AlphaFold2 (Jumper; Evans; Pritzel; Green et al., 2021) com base em sequências depositadas no UniProtKB. Além disso, a hidrofobicidade para a região ligada à necrose de $ExhC_{wt}$ e suas variantes mutantes, assim como para as regiões equivalentes de todas as outras proteínas homólogas analisadas, foram geradas usando Pymol 1.8.4.0 com o *script* disponível em pymolwiki.org/index.php/Color_h.

3.21.4 Previsão de modos de ligação e simulações de Dinâmica Molecular

Entre os peptídeos identificados através da phage display contra a ExhC, um deles, YTSPWWYYQMDEHMQG, exibiu estrutura secundária bem definida com base em previsões de estrutura 3D realizadas com AlphaFold2, apresentando dobramento α-hélice. Posteriormente, recorreu-se ao acoplamento às cegas proteína-peptídeo usando três algoritmos (kozakov; Hall; Xia; Porter et al., 2017; Park; Baek; Lee; Seok, 2019; Yan; Zhang; Zhou; Li et al., 2017), isto é, ClusPro, GalaxyTongDock e HDOCK, para prever os modos de ligação do peptídeo à ExhC. Após determinar os modelos de ligação por cada algoritmo de acoplamento, a estabilidade do modo de ligação proteína-peptídeo foi investigada usando simulações MD com pmemd.cuda de Amber22, seguido por cálculos de energia livre MM-GBSA (modelo GBneck, implementado em MMPBSA.py de Amber22). A mesma abordagem foi utilizada para prever a estrutura do suposto complexo ExhC-Pcyox1l. No entanto, para reduzir a demanda computacional associada a esta predição, apenas o algoritmo de acoplamento (ClusPro) que levou ao modo de ligação mais estável do peptídeo à ExhC foi empregado nesta análise. A estabilidade dos modos de ligação previstos durante as simulações MD do peptídeo e da proteína parceira foi monitorada através de cálculos RMSD com cpptraj de Amber22.

4. Referências

ADEGOKE, G. O. Comparative characteristics of *Staphylococcus sciuri*, *Staphylococcus lentus* and *Staphylococcus gallinarum* isolated from healthy and sick hosts. **Vet Microbiol**, 11, n. 1-2, p. 185-189, Feb 1986.

AHRENS, P.; ANDRESEN, L. O. Cloning and sequence analysis of genes encoding *Staphylococcus hyicus* exfoliative toxin types A, B, C, and D. **J Bacteriol**, 186, n. 6, p. 1833-1837, Mar 2004.

ALBERTS, B.; J. A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. Biologia Molecular da Célula. 6ª Edição. **São Paulo: ARTMED EDITORA LTDA**. 2017.

ANDRESEN, L. O. Production of exfoliative toxin by isolates of *Staphylococcus hyicus* from different countries. **J Veterinary record**, 157, n. 13, p. 376-378, 2005.

ANDRESEN, L. O.; AHRENS, P.; DAUGAARD, L.; BILLE-HANSEN, V. Exudative epidermitis in pigs caused by toxigenic *Staphylococcus chromogenes*. **Vet Microbiol**, 105, n. 3-4, p. 291-300, Feb 25 2005.

BAILEY, C. J.; REDPATH, M. B. The esterolytic activity of epidermolytic toxins. **Biochem J**, 284 (Pt 1), n. Pt 1, p. 177-180, May 15 1992.

BAKER, N. A.; SEPT, D.; JOSEPH, S.; HOLST, M. J. *et al.* Electrostatics of nanosystems: application to microtubules and the ribosome. 98, n. 18, p. 10037-10041, 2001.

BOONE, D. R.; CASTENHOLZ, R. W.; GARRITY, G. M.; STANLEY, J. **Bergey's Manual® of Systematic Bacteriology: Volume One The Archaea and the Deeply Branching and Phototrophic Bacteria**. Springer, 2001. 0387987711.

BUKOWSKI, M.; WLADYKA, B.; DUBIN, G. Exfoliative toxins of *Staphylococcus aureus*. **Toxins (Basel)**, 2, n. 5, p. 1148-1165, May 2010.

BURDON, R. H.; GILL, V.; ALLIANGANA, D. Hydrogen peroxide in relation to proliferation and apoptosis in BHK-21 hamster fibroblasts. **Free Radic Res**, 24, n. 2, p. 81-93, Feb 1996.

CARVALHO, L. F. O. S.; MORENO, A. M.; SOBESTIANSKY, J.; BARCELLOS, D. Ear necrosis. **In swine diseases**, p. 403-04, 2007.

CAVARELLI, J.; PREVOST, G.; BOURGUET, W.; MOULINIER, L. *et al.* The structure of *Staphylococcus aureus* epidermolytic toxin A, an atypic serine protease, at 1.7 A resolution. **Structure**, 5, n. 6, p. 813-824, Jun 15 1997.

CHEN, S.; WANG, Y.; CHEN, F.; YANG, H. *et al.* A highly pathogenic strain of *Staphylococcus sciuri* caused fatal exudative epidermitis in piglets. **PLoS One**, 2, n. 1, p. e147, Jan 10 2007.

CHEN, V. B.; ARENDALL, W. B., 3rd; HEADD, J. J.; KEEDY, D. A. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. **Acta Crystallogr D Biol Crystallogr**, 66, n. Pt 1, p. 12-21, Jan 2010.

CHRISTOFFER, C.; BHARADWAJ, V.; LUU, R.; KIHARA, D. LZerD Protein-Protein Docking Webserver Enhanced With de novo Structure Prediction. **Front Mol Biosci**, 8, p. 724947, 2021.

CROOKS, G. E.; HON, G.; CHANDONIA, J. M.; BRENNER, S. E. WebLogo: a sequence logo generator. **Genome research**, v. 14, n. 6, p. 1188-1190, 2004.

DAKIC, I.; MORRISON, D.; VUKOVIC, D.; SAVIC, B. *et al.* Isolation and molecular characterization of *Staphylococcus sciuri* in the hospital environment. **J Clin Microbiol**, 43, n. 6, p. 2782-2785, Jun 2005.

DE VRIES, S. J.; BONVIN, A. M. CPORT: a consensus interface predictor and its performance in prediction-driven docking with HADDOCK. **PLoS One**, 6, n. 3, p. e17695, Mar 25 2011.

DEMIDYUK, I. V.; CHUKHONTSEVA, K. N.; KOSTROV, S. V. Glutamyl Endopeptidases: The Puzzle of Substrate Specificity. **Acta Naturae**, 9, n. 2, p. 17-33, Apr-Jun 2017.

DOMINGUEZ, C.; BOELENS, R.; BONVIN, A. M. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. **J Am Chem Soc**, 125, n. 7, p. 1731-1737, Feb 19 2003.

EDGAR, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. **BMC Bioinformatics**, 5, p. 113, Aug 19 2004.

EZQUERRA, T. A.; GARCIA-GUTIERREZ, M. C.; NOGALES, A.; GOMEZ, M. **Applications of synchrotron light to scattering and diffraction in materials and life sciences**. Springer Science & Business Media, 2009. 354095967X.

FUDABA, Y.; NISHIFUJI, K.; ANDRESEN, L. O.; YAMAGUCHI, T. *et al*. *Staphylococcus hyicus* exfoliative toxins selectively digest porcine desmoglein 1. **J Microbial pathogenesis**, 39, n. 5-6, p. 171-176, 2005.

GETSIOS, S.; HUEN, A. C.; GREEN, K. J. Working out the strength and flexibility of desmosomes. **Nat Rev Mol Cell Biol**, 5, n. 4, p. 271-281, Apr 2004.

GREENFIELD, N. J. Using circular dichroism spectra to estimate protein secondary structure. **Nat Protoc**, 1, n. 6, p. 2876-2890, 2006.

HAMMERSLEY, A. FIT2D: a multi-purpose data reduction, analysis and visualization program. **J Journal of Applied Crystallography**, 49, n. 2, p. 646-652, 2016.

HANAKAWA, Y.; SCHECHTER, N. M.; LIN, C.; NISHIFUJI, K. *et al.* Enzymatic and molecular characteristics of the efficiency and specificity of exfoliative toxin cleavage of desmoglein 1. **J Biol Chem**, 279, n. 7, p. 5268-5277, Feb 13 2004.

HAUSCHILD, T.; WOJCIK, A. Species distribution and properties of staphylococci from canine dermatitis. **Res Vet Sci**, 82, n. 1, p. 1-6, Feb 2007.

HEDIN, G.; WIDERSTROM, M. Endocarditis due to *Staphylococcus sciuri*. **Eur J Clin Microbiol Infect Dis**, 17, n. 9, p. 673-675, Sep 1998.

IMANISHI, I.; NICOLAS, A.; CAETANO, A. B.; CASTRO, T. L. P. *et al.* Exfoliative toxin E, a new *Staphylococcus aureus* virulence factor with host-specific activity. **Sci Rep**, 9, n. 1, p. 16336, Nov 8 2019.

IYORI, K.; FUTAGAWA-SAITO, K.; HISATSUNE, J.; YAMAMOTO, M. *et al. Staphylococcus pseudintermedius* exfoliative toxin EXI selectively digests canine desmoglein 1 and causes subcorneal clefts in canine epidermis. **Vet Dermatol**, 22, n. 4, p. 319-326, Aug 2011.

IYORI, K.; HISATSUNE, J.; KAWAKAMI, T.; SHIBATA, S. *et al.* Identification of a novel *Staphylococcus pseudintermedius* exfoliative toxin gene and its prevalence in isolates from canines with pyoderma and healthy dogs. **FEMS Microbiol Lett**, 312, n. 2, p. 169-175, Nov 2010.

JIMENEZ-GARCIA, B.; ELEZ, K.; KOUKOS, P. I.; BONVIN, A. M. *et al.* PRODIGY-crystal: a web-tool for classification of biological interfaces in protein complexes. **Bioinformatics**, 35, n. 22, p. 4821-4823, Nov 1 2019.

JUMPER, J.; EVANS, R.; PRITZEL, A.; GREEN, T. *et al.* Highly accurate protein structure prediction with AlphaFold. **Nature**, 596, n. 7873, p. 583-589, Aug 2021.

JUNQUEIRA, L. C.; CARNEIRO, J. Histologia básica. 13ª edicao. Rio de Janeiro - RJ: Guanabara Koogan, 2018.

KABSCH, W. Xds. **Acta Crystallogr D Biol Crystallogr**, 66, n. Pt 2, p. 125-132, Feb 2010.

KAROW, A. R.; GOTZL, J.; GARIDEL, P. Resolving power of dynamic light scattering for protein and polystyrene nanoparticles. **Pharm Dev Technol**, 20, n. 1, p. 84-89, Jan 2015.

KLOOS, W. E.; BALLARD, D. N.; WEBSTER, J. A.; HUBNER, R. J. *et al.* Ribotype delineation and description of *Staphylococcus sciuri* subspecies and their potential as reservoirs of methicillin resistance and staphylolytic enzyme genes. 47, n. 2, p. 313-323, 1997.

KLOOS, W. E.; SCHLEIFER, K. H.; SMITH, R. F. Characterization of *Staphylococcus sciuri* sp. nov. and its subspecies. **J International Journal of Systematic Evolutionary Microbiology**, 26, n. 1, p. 22-37, 1976.

KO, J.; PARK, H.; HEO, L.; SEOK, C. GalaxyWEB server for protein structure prediction and refinement. **Nucleic Acids Res**, 40, n. Web Server issue, p. W294-297, Jul 2012.

KOZAKOV, D.; HALL, D. R.; XIA, B.; PORTER, K. A. *et al.* The ClusPro web server for protein-protein docking. **Nat Protoc**, 12, n. 2, p. 255-278, Feb 2017.

KREJCI, A.; HUPP, T. R.; LEXA, M.; VOJTESEK, B. *et al.* Hammock: a hidden Markov model-based peptide clustering algorithm to identify protein-interaction consensus motifs in large datasets. **Bioinformatics**, 32, n. 1, p. 9-16, Jan 1 2016.

KRISSINEL, E.; HENRICK, K. J. J. o. m. b. Inference of macromolecular assemblies from crystalline state. 372, n. 3, p. 774-797, 2007.

LADHANI, S.; JOANNOU, C. L.; LOCHRIE, D. P.; EVANS, R. W. *et al.* Clinical, microbial, and biochemical aspects of the exfoliative toxins causing staphylococcal scalded-skin syndrome. **J Clinical microbiology reviews**, 12, n. 2, p. 224-242, 1999.

LE MARECHAL, C.; JARDIN, J.; JAN, G.; EVEN, S. *et al. Staphylococcus aureu*s seroproteomes discriminate ruminant isolates causing mild or severe mastitis. **Vet Res**, 42, n. 1, p. 35, Feb 15 2011.

LE MARECHAL, C.; SEYFFERT, N.; JARDIN, J.; HERNANDEZ, D. *et al.* Molecular basis of virulence in *Staphylococcus aureus* mastitis. **PLoS One**, 6, n. 11, p. e27354, 2011.

LEE, C. Y.; SCHMIDT, J. J.; JOHNSON-WINEGAR, A. D.; SPERO, L. *et al.* Sequence determination and comparison of the exfoliative toxin A and toxin B genes from *Staphylococcus aureus*. **J Bacteriol**, 169, n. 9, p. 3904-3909, Sep 1987.

LI, H.; LI, X.; LU, Y.; WANG, X. *et al. Staphylococcus sciuri* exfoliative toxin C is a dimer that modulates macrophage functions. **Can J Microbiol**, 57, n. 9, p. 722-729, Sep 2011.

LI, H.; WANG, Y.; DING, L.; ZHENG, S. J. *Staphylococcus sciuri* exfoliative toxin C (ExhC) is a necrosis-inducer for mammalian cells. **PLoS One**, 6, n. 7, p. e23145, 2011.

LIENEN, T.; SCHNITT, A.; HAMMERL, J. A.; MAURISCHAT, S. *et al.* Mammaliicoccus spp. from German Dairy Farms Exhibit a Wide Range of Antimicrobial Resistance Genes and Non-Wildtype Phenotypes to Several Antibiotic Classes. **Biology (Basel)**, 11, n. 2, Jan 18 2022.

MADHAIYAN, M.; WIRTH, J. S.; SARAVANAN, V. S. Phylogenomic analyses of the Staphylococcaceae family suggest the reclassification of five species within the genus Staphylococcus as heterotypic synonyms, the promotion of five subspecies to novel species, the taxonomic reassignment of five Staphylococcus species to Mammaliicoccus gen. nov., and the formal assignment of Nosocomiicoccus to the family Staphylococcaceae. **Int J Syst Evol Microbiol**, 70, n. 11, p. 5926-5936, Nov 2020.

MARIUTTI, R. B.; SOUZA, T. A.; ULLAH, A.; CARUSO, I. P. *et al.* Crystal structure of *Staphylococcus aureus* exfoliative toxin D-like protein: Structural basis for the high specificity of exfoliative toxins. **Biochem Biophys Res Commun**, 467, n. 1, p. 171-177, Nov 6 2015.

MARIUTTI, R. B.; TARTAGLIA, N. R.; SEYFFERT, N.; DE PAULA CASTRO, T. L. *et al.* Exfoliative toxins of *Staphylococcus aureus*. **J The Rise of Virulence Antibiotic Resistance in Staphylococcus aureus**, 2017.

MELISH, M. E.; GLASGOW, L. A. The staphylococcal scalded-skin syndrome. **N Engl J Med**, 282, n. 20, p. 1114-1119, May 14 1970.

MONDAY, S. R.; VATH, G. M.; FERENS, W. A.; DEOBALD, C. *et al.* Unique superantigen activity of staphylococcal exfoliative toxins. **J Immunol**, 162, n. 8, p. 4550-4559, Apr 15 1999.

MORÉS, N. Necrose de orelha em suínos: uma visão crítica do problema. **In: Embrapa Suínos e Aves-Artigo em anais de congresso. In: SIMPÓSIO INTERNACIONAL DE SUINOCULTURA**, 6., Porto Alegre, RS. Produção, reprodução e sanidade suína: anais Porto Alegre: UFRGS, p. 199-205, 2011.

NEMEGHAIRE, S.; VANDERHAEGHEN, W.; ARGUDIN, M. A.; HAESEBROUCK, F. *et al.* Characterization of methicillin-resistant *Staphylococcus sciuri* isolates from industrially raised pigs, cattle and broiler chickens. **J Antimicrob Chemother**, 69, n. 11, p. 2928-2934, Nov 2014.

NISHIFUJI, K.; SUGAI, M.; AMAGAI, M. Staphylococcal exfoliative toxins: "molecular scissors" of bacteria that attack the cutaneous defense barrier in mammals. **J Dermatol Sci**, 49, n. 1, p. 21-31, Jan 2008.

O'TOOLE, P. W.; FOSTER, T. J. Nucleotide sequence of the epidermolytic toxin A gene of Staphylococcus aureus. **J Bacteriol**, 169, n. 9, p. 3910-3915, Sep 1987.

PALMER, T.; BONNER, P. L. **Enzymes: Biochemistry, Biotechnology, Clinical Chemistry**. Elsevier Science, 2007. 9781904275275.

PAPAGEORGIOU, A. C.; PLANO, L. R.; COLLINS, C. M.; ACHARYA, K. R. Structural similarities and differences in *Staphylococcus aureus* exfoliative toxins A and B as revealed by their crystal structures. **Protein Sci**, 9, n. 3, p. 610-618, Mar 2000.

PARK, T.; BAEK, M.; LEE, H.; SEOK, C. GalaxyTongDock: Symmetric and asymmetric ab initio protein-protein docking web server with improved energy parameters. **J Comput Chem**, 40, n. 27, p. 2413-2417, Oct 15 2019.

PETOUKHOV, M. V.; FRANKE, D.; SHKUMATOV, A. V.; TRIA, G. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. **J Appl Crystallogr**, 45, n. Pt 2, p. 342-350, Apr 1 2012.

POROLLO, A.; MELLER, J. Prediction-based fingerprints of protein-protein interactions. **Proteins**, 66, n. 3, p. 630-645, Feb 15 2007.

RAGO, J. V.; VATH, G. M.; BOHACH, G. A.; OHLENDORF, D. H. *et al.* Mutational analysis of the superantigen staphylococcal exfoliative toxin A (ETA). **J Immunol**, 164, n. 4, p. 2207-2213, Feb 15 2000.

SCHRÖDINGER, L.; DELANO, W. PyMOL. 2020.

SCHWARZ, L.; LONCARIC, I.; BRUNTHALER, R.; KNECHT, C. *et al.* Exudative Epidermitis in Combination with Staphylococcal Pyoderma in Suckling Piglets. **Antibiotics (Basel)**, 10, n. 7, Jul 9 2021.

SHITTU, A.; LIN, J.; MORRISON, D.; KOLAWOLE, D. Isolation and molecular characterization of multiresistant *Staphylococcus sciuri* and *Staphylococcus haemolyticus* associated with skin and soft-tissue infections. **J Journal of medical microbiology**, 53, n. 1, p. 51-55, 2004.

SINGH, N.; BHALLA, N. Moonlighting Proteins. **Annu Rev Genet**, 54, p. 265-285, Nov 23 2020.

SREERAMA, N.; WOODY, R. W. On the analysis of membrane protein circular dichroism spectra. **Protein Sci**, 13, n. 1, p. 100-112, Jan 2004.

STANLEY, J. R.; AMAGAI, M. Pemphigus, bullous impetigo, and the staphylococcal scalded-skin syndrome. **N Engl J Med**, 355, n. 17, p. 1800-1810, Oct 26 2006.

STEPANOVIC, S.; DAKIC, I.; DJUKIC, S.; LOZUK, B. *et al.* Surgical wound infection associated with *Staphylococcus sciuri*. **Scand J Infect Dis**, 34, n. 9, p. 685-686, 2002.

STEPANOVIĆ, S.; VUKOVIĆ, D.; JEŽEK, P.; PAVLOVIĆ, M. *et al.* Influence of dynamic conditions on biofilm formation by staphylococci. **J European Journal of Clinical Microbiology**, 20, p. 502-504, 2001.

STOTHARD, P. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. **Biotechniques**, 28, n. 6, p. 1102, 1104, Jun 2000.

VATH, G. M.; EARHART, C. A.; MONIE, D. D.; IANDOLO, J. J. *et al.* The crystal structure of exfoliative toxin B: a superantigen with enzymatic activity. **Biochemistry**, 38, n. 32, p. 10239-10246, Aug 10 1999.

VATH, G. M.; EARHART, C. A.; RAGO, J. V.; KIM, M. H. *et al.* The structure of the superantigen exfoliative toxin A suggests a novel regulation as a serine protease. **J Biochemistry**, 36, n. 7, p. 1559-1566, 1997.

WALLET, F.; STUIT, L.; BOULANGER, E.; ROUSSEL-DELVALLEZ, M. *et al.* Peritonitis due to *Staphylococcus sciuri* in a patient on continuous ambulatory peritoneal dialysis. **Scand J Infect Dis**, 32, n. 6, p. 697-698, 2000.

WEGENER, H. C.; ANDRESEN, L. O.; BILLE-HANSEN, V. *Staphylococcus hyicus* virulence in relation to exudative epidermitis in pigs. **Can J Vet Res**, 57, n. 2, p. 119-125, Apr 1993.

WESTERMEIER, R. **Electrophoresis in practice: a guide to methods and applications of DNA and protein separations**. John Wiley & Sons, 2016. 3527338802.

XU, Y.; WANG, S.; HU, Q.; GAO, S. *et al.* CavityPlus: a web server for protein cavity detection with pharmacophore modelling, allosteric site identification and covalent ligand binding ability prediction. **Nucleic Acids Res**, 46, n. W1, p. W374-W379, Jul 2 2018.

YAMAGUCHI, T.; NISHIFUJI, K.; SASAKI, M.; FUDABA, Y. *et al.* Identification of the *Staphylococcus aureus* etd pathogenicity island which encodes a novel exfoliative toxin, ETD, and EDIN-B. **Infect Immun**, 70, n. 10, p. 5835-5845, Oct 2002.

YAN, Y.; ZHANG, D.; ZHOU, P.; LI, B. *et al.* HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. **Nucleic Acids Res**, 45, n. W1, p. W365-W373, Jul 3 2017.

YIN, T.; GREEN, K. J. Regulation of desmosome assembly and adhesion. **Semin Cell Dev Biol**, 15, n. 6, p. 665-677, Dec 2004.

CAPÍTULO II. Necrotic activity of Exhc from *Mammaliicoccus sciuri* is mediated by specific amino acid residues

Necrotic activity of ExhC from *Mammaliicoccus sciuri* is mediated by specific amino acid residues

Carolina Gismene[a]; Jorge Enrique Hernández González[a]; Marília de Freitas Calmon[b]; Andrey Fabricio Ziem Nascimento[c]; Angela Rocio Niño Santisteban[a]; Felipe Antunes Calil[a]; Alana Della Torre da Silva[d]; Paula Rahal[b], Rejane Maira Góes[d]; Raghuvir Krishnaswamy Arni[a]; Ricardo Barros Mariutti[a]

[a]Multiuser Center for Biomolecular Innovation, São Paulo State University - UNESP, São José do Rio Preto, SP, Brazil.

[b]Laboratory of Genomic Studies, São Paulo State University - UNESP, São José do Rio Preto, SP, Brazil.

[c]Brazilian Synchrotron Light Laboratory (LNLS), Brazilian Center for Research in Energy and Materials (CNPEM), Campinas, SP, Brazil.

[d]Department of Biological Sciences, São Paulo State University - UNESP, São José do Rio Preto, SP, Brazil.

Highlights

- Nonconserved residues R47, N49, Q51 and R89 are crucial for the necrotic activity of ExhC.
- Necrotic activity is independent of ExhC enzymatic activity.
- Residues R47, N49, Q51 and R89 lie in a cavity-shaped region of ExhC from *Mammaliicoccus sciuri*, as confirmed by the crystal structures of wild-type and mutant protein.
- Necrosis-linked region in ExhC is highly-positively charged, differently from other exfoliative toxins deprived of necrotic activity.

Abstract

Mammaliicoccus sciuri, a commensal and pathogenic bacterium of significant clinical and veterinary relevance, expresses exfoliative toxin C (ExhC), a specific glutamyl endopeptidase belonging to the chymotrypsin family as the principal virulence factor. However, unlike most members of this family, ETs are inactive against a wide range of substrates and possess exquisite specificity for desmoglein-1 (Dsg1), a cadherin-like adhesion molecule that is crucial to maintain tissue integrity, thereby preventing the separation of skin cells and the entry of pathogens. ExhC is of clinical importance since in addition to causing exfoliation in pigs and mice, it induces necrosis in multiple mammalian cell lines, a property not observed for other ETs. Previous experiments have implicated the $ExhC_{79\text{-}128}$ fragment in causing necrosis. Site-directed mutagenesis of specific residues within this fragment were studied and led to the design

of an ExhC variant containing four-point mutations ($ExhC_{mut4}$) lacking necrotic potential but retaining nearly wild-type (wt) levels of enzymatic activity. Moreover, the determination of the $ExhC_{wt}$ and $ExhC_{mut4}$ crystal structures identified the conformation in the necrosis-linked region. These results constitute an important step toward the understanding of the mechanisms underlying the necrotic and epidermolytic activity of ExhC.

Keywords: Exfoliative toxin C; necrotic activity; moonlighting protein; enzymatic activity; necrosis-linked region; *Mammaliicoccus sciuri*.

Abbreviations

$ExhC_{wt}$, wild-type ExhC; $ExhC_{mut4}$, ExhC variant containing four point mutations; $ExhC_{R47A}$, ExhC with single-point mutation R47A; $ExhC_{N49F}$, ExhC with single-point mutation N49F; $ExhC_{Q51K}$, ExhC with single-point mutation Q51K; $ExhC_{R89A,}$ ExhC with single-point mutation R89A; ETs, exfoliative toxins; Dsg1, desmoglein 1; SSSS, staphylococcal scalded skin syndrome; GEPs, glutamyl endopeptidases; CLSP, chymotrypsin-like serine proteinase; EE, exudative epidermitis; LB, lysogeny broth; OD_{600}, optical density measured at 600 nm; SEC, size-exclusion chromatography; BN-PAGE, Blue-Native-PAGE; EV_{max}, maximum elution volume; TBS, Tris-buffered saline; DLS, dynamic light scattering; D_H, hydrodynamic diameter; CD, circular dichroism; BHK-21, baby hamster kidney fibroblast cell line; DMEM, modified eagle medium; ACHN, renal cancer cell line; PK-15, cell line isolated from a pig kidney; 786-0, renal cancer cell line; HK-2, cell line derived from a normal, human adult male kidney; MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide; PBS, phosphate-buffered saline; PDB, Protein Data Bank; SI, sequence identity; MW, molecular weight; BSA, bovine serum albumin; AU, asymmetric units.

1. Introduction

Staphylococcal exfoliative toxins (ETs) are virulence factors expressed by certain strains of infectious Staphylococci, which specifically cleave the desmosomal cadherin desmoglein 1 (Dsg1) resulting in skin damage [1, 2]. Four ETs referred to as ETA [3] ETB [4], ETD [5] and ETE [6, 7] have been identified in different *Staphylococcus aureus* strains. These proteins are glutamyl endopeptidases (GEPs) belonging to the chymotrypsin-like serine proteinase (CLSP) family [8] and possess the characteristic structural features that are common to CLSPs, such as the overall chymotrypsin fold, which consists of two domains, each containing six antiparallel β-strands that form a β-barrel [8, 9] and the catalytic triad comprising residues *H57-D102-S195*

(residues in italic are based on chymotrypsin numbering) [8]. However, ETs display unique structural features, such as an extended highly-charged N-terminal α-helix and a flipped conformation of *P192* carbonyl oxygen, resulting in the formation of a nonfunctional oxyanion hole [8, 10-12]. The deletion of the extended N-terminal α-helix in ETA abrogates enzymatic activity, highlighting the functional importance of this structural element [10]. Moreover, the flipped conformation of residue *P192* has been considered to be relevant for the observed specificity of ETs for Dsg1, and it has been proposed that a functional oxyanion hole is formed upon binding to Dsg1 [2], [10], [11]. Recent enzyme activity assays conducted by our group downplays the role of residue *P192* and stresses the importance of other molecular events, such as homodimerization [13]. This oligomeric state may play a relevant role in regulating the exfoliative activity of ETs, although conclusive evidence supporting this hypothesis is currently lacking [13].

ETA-producing strains can cause the staphylococcal scalded skin syndrome (SSSS), a serious disease characterized by fever [14], malaise [15, 16], loss of appetite [16], followed by erythematous rash and widespread skin blistering [15-18]. ETB and ETD have been associated with the onset of more localized skin diseases, such as bullous impetigo [16] and furuncles [5], respectively. The fourth toxin, ETE, was identified in *S. aureus* strains isolated from ewe mastitis and is believed to facilitate the colonization of teat canal epithelium [7].

Other ETs have been isolated from species of Staphylococci [16] and Mammaliicoccus [19]. For example, strains of *Staphylococcus hyicus* produce at least six different ETs, termed ExhA, ExhB, ExhC, ExhD, ShetA and ShetB [20]. These are associated with the occurrence of exudative epidermitis (EE) in pigs caused by the degradation of swine Dsg1 [20], a disease with symptoms similar to those of SSSS in humans [21], such as exfoliation accompanied by epidermal cell separation, erythema and serous exudation [22]. Lately, the incidence of EE has been increasing through large outbreaks that have caused high mortality in piglets and, consequently, significant loss to livestock [23].

ExhC was also found in a highly pathogenic strain of *Staphylococcus sciuri* (recently reclassified as *Mammaliicoccus sciuri* [19]) responsible for a serious outbreak of EE in pigs in China [24]. Moreover, a variant of ExhB, differing in only two residues with respect to that expressed in *S. hyicus*, was identified in *S. chromogenes* [25], while two new ETs, ExpA (=EXI) [26] and ExpB [27], with the ability to cleave canine Dsg1, were identified in *S. pseudintermedius* strains.

Some ETs display other biological functions in addition to exfoliative activity and, therefore, can be classified as moonlighting proteins [28]. For example, ETA [29] and ETB [30] are reported as superantigens and possess T lymphocyte mitogenic activity, whereas ExhC induces necrosis in multiple cell lines [31], properties that, together with Dsg1 cleavage, may be important for pathogenesis. In an attempt to identify the region of ExhC responsible for this activity, several truncated variants were produced and their necrotic potential was evaluated [31].

Deletion or blocking of the $ExhC_{79\text{-}128}$ fragment (or $ExhC_{44\text{-}92}$ fragment in the absence of the signal peptide) was shown to abolish ExhC induced necrosis or skin lesions in newborn mice [31], however, the conformation, charge distribution and specific interactions responsible for this were not characterized. In addition, it was shown that this toxin inhibits the phagocytosis of *Mammaliicocci* by macrophages *in vitro* [32]. These findings suggest that ETs play key roles in modulating the host immune response to benefit the infecting bacteria.

The comprehension of the role of ETs in pathogenesis requires the study of their moonlighting functions. In this work, we use bioinformatic tools to design ExhC mutants in order to evaluate their necrotic potential. As a result, we identify four-point mutations (R47A, N49F, Q51K and R89A) that abrogate the necrotic activity of ExhC but do not significantly modify the esterolytic activity of the protein, indicating that the necrotic and catalytic activities are independent and involve distinct sites. Remarkably, crystal structures of wild-type and mutant ExhC were obtained. The analysis of ExhC wild-type revealed that the four key residues subjected to mutation lie at the surface of a pocket-shaped region with high density of positive charges. Furthermore, the mutated variant displayed striking differences in terms of electrostatic potential and hydrophobicity of the amino acids of the region implicated in necrosis, suggesting that the mutations might interfere with the recognition and binding of an effector molecule responsible for cell death. Nonetheless, this hypothesis is not supported by the fact that ExhC fragments are able to induce necrosis.

The results presented here contribute to understanding the underlying mechanism of necrotic and EE activities observed in ExhC. Unraveling the detailed events in the necrotic and exfoliative processes at the atomic level is important for designing highly specific inhibitors that can potentially be used clinically and also for enhancing diagnostics involving pathogens encoding exfoliative proteins.

2. Materials and Methods

2.1 Sequence alignment and structural models

The amino acid sequence of ExhC from *M. sciuri* was compared to other ETs using multiple sequence alignment performed at the Muscle server (www.ebi.ac.uk/Tools/msa/muscle/) [33]. SMS2 [34] was used to visualize the alignment of residues 44-92 of ExhC of *M. sciuri* ($ExhC_{44-92}$) and the analogous regions in the other staphylococcal ETs. The linear sequences of the proteins were obtained from the UniProt database (https://www.uniprot.org/): ExhC of *M. sciuri* (UniProt ID: F6M8N2); ETA (UniProt ID: P09331), ETB (UniProt ID: P09332), ETD (UniProt ID: Q8GAX8) and ETE (UniProt ID: L0RUV7) of *S. aureus*; ExhA (UniProt ID: Q8GNJ7), ExhB (UniProt ID: Q8GNJ6), ExhC (Uniprot ID: Q8GNJ5), ExhD (UniProt ID: Q8GNJ4), ShetA (UniProt ID: E5RQL1) and ShetB (UniProt ID: Q9KWH0) of *S. hyicus*; ExhB (UniProt ID: Q8GNJ6) of *S. chromogenes*; EXI (UniProt ID: D0VXY8) and ExpB (Uniprot ID: E2RZJ7) of *S. pseudintermedius*. ETC of *S. aureus* (Uniprot ID: Q9KWG9) and SIET of *S. pseudintermedius* (Uniprot ID: A0A8F4Q3I3) were not added to the sequence alignments, as these two proteins are not CLSPs [35]. Structural models of wild-type ExhC ($ExhC_{wt}$) and $ExhC_{44-92}$ [31] were generated using AlphaFold2 [36].

2.2. Protein expression and purification

Escherichia coli BL21 (DE3) competent cells were transformed with the pET-28a(+) expression vector with kanamycin resistance (GenScript) for the $ExhC_{wt}$ and an ExhC variant containing the following mutations: R47A, N49F, Q51K and R89A ($ExhC_{mut4}$). A single bacterial colony of each bacterial transformation was selected and grown overnight at 37 ºC in Lysogeny Broth (LB) medium supplemented with kanamycin (50 mg/mL) [37]. Subsequently, the cultures were diluted 100-fold with a fresh LB medium containing kanamycin (50 mg/mL) and incubated at 37 ºC under vigorous agitation. When the optical density measured at 600 nm (OD_{600}) reached 0.5, the expressions were induced with 0.4 mM IPTG for 5 hours at 30 ºC. The cells were collected by centrifugation at 2,600 *g* for 10 min at 4 ºC and homogenized by sonication in a lysis buffer containing 5 mM $NaHPO_4$, pH 7.7, 400 mM NaCl, 10 mM imidazole and 10% (v/v) glycerol [38]. The lysed cells were centrifuged at 15,000 g for 30 min at 4 ºC, and the supernatants with a large quantity of $ExhC_{wt}$ and $ExhC_{mut4}$ were subjected to affinity chromatography using a Ni^{2+}-Sepharose column (Bio-Rad) [39]. Protein concentrations were determined to be 10 mg/mL by a spectrophotometer (Biomate 3S - Thermo scientific) and further purification was performed in a 20 mM Tris-HCl buffer, pH 7.5 containing 100 mM

NaCl using an AKTA system purifier by size-exclusion chromatography (SEC) [40] on a Superdex G75 10/300 GL column (GE - Healthcare Life Sciences) at a flow rate of 0.7 mL/min. Protein purity was evaluated by Western blot, SDS-PAGE and Blue-Native-PAGE (BN-PAGE) 15% [41]. After optimizing this purification protocol, the ExhC point mutants ($ExhC_{R47A}$, $ExhC_{N49F}$, $ExhC_{Q51K}$ and $ExhC_{R89A}$), ETA and ETE were expressed and purified under the same conditions.

2.3. Dynamic light scattering

The hydrodynamic diameters (D_H) of $ExhC_{wt}$ and ExhC mutants were determined by DLS using a ZETASIZER Nano series (Malvern Instruments) in a 20 mM Tris-HCl pH 7.5 buffer containing 100 mM NaCl (TBS) [42]. Data was acquired in 10 runs of 30 seconds each at a constant temperature of 25 ºC using a protein concentration of 50 μM. The calculation of the D_H was performed using the Zetasizer software.

2.4. Circular dichroism spectroscopy

The far UV-CD spectra of ExhC variants were obtained using a Jasco J-710 spectropolarimeter (Jasco, Brazil) at a constant temperature of 25 ºC using a 0.5 mm quartz cuvette. Proteins samples were diluted to 9-11 μM in 5 mM NaH_2PO_4 pH 7.4 and 50 mM NaF buffer. For the data collection, the spectral range of 190-260 nm was used with a scan rate of 50 nm/min, response time of 1.0 second, spectral bandwidth of 1 nm, spectral resolution of 0.2 nm and 10 accumulations. After obtaining the data, all CD spectra recorded were corrected by subtracting the spectrum from the buffer solution obtained under identical conditions, and converted into molar ellipticity in units of 10^3.deg.cm^2.dmol^{-1} by the OriginLab software 7. The percentages of secondary structures were estimated [43] by the CONTINLL software of the CDPro package, using the reference set of proteins SP29 [44].

2.5. Cell lines and culture conditions

The baby hamster kidney fibroblast cell line (BHK-21) (ATCC: CCL-10) was cultured in Dulbeccos's Modified Eagle Medium (DMEM-Vitrocell). The renal cancer cell line ACHN (ATCC: CRL-1611) and the cell line isolated from a pig kidney PK-15 (ATCC: CCL-33) were cultured EMEM medium (Cultilab, Brazil). The renal cancer cell line 786-0 (ATCC: CRL-1932) was cultured in RPMI-1640 (Cultilab, Brazil). These mediums were supplemented with 10% fetal bovine serum (Vitrocell), 1% non-essential amino acids (GIBCO) and 200 U/mL of

penicillin and streptomycin (Vitrocell). The HK-2 cell line derived from a normal, human adult male kidney (ATCC: CRL-2190) was cultured in KSFM medium (GIBCO). All the cell lines were maintained under conditions of 37 ºC in a humid atmosphere with 5% CO_2 until reaching cellular confluence.

2.6. Cell viability assay

The cell lines were seeded at the density of 1×10^4 cells/well in 96-well plates for 24 h and treated with $ExhC_{wt}$ and ExhC mutants diluted in PBS 1X at an initial maximum concentration of 60 µM with log_2 decay until reaching a minimum concentration of 0.94 µM. The control experiment was 1×10^4 cells/well with PBS 1X in the culture medium. After 48 h of treatment, 1 mg/mL of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide (MTT) (Sigma Aldrich, St Louis, MO, USA) was diluted to 100 µL in the medium within each well and incubated for 30 min at 37 °C. Then, the MTT was removed and 100 µL of DMSO (Sigma-Aldrich, St. Louis, Missouri, USA) was added to solubilize the formazan crystals. Subsequently, the absorbance was determined using a microplate reader (Epoch Microplate Spectrophotometer – Bio Tek Instruments) at a wavelength of 572 nm. All experiments were performed in triplicate and in three independent assays.

2.7. Apoptosis assay

BHK-21 cells were previously treated with the irreversible pan caspase inhibitor Z-VAD-FMK (Promega). First, 1×10^4 cells/well in 96-well plates were pretreated for 2 hours with 50 µM of Z-VAD-FMK (Promega) and then incubated with 60 µM of $ExhC_{wt}$ for 24 hours at 37 °C and 5% CO_2 [31]. The positive control was 1×10^4 cells/well incubated with 100 µM of hydrogen peroxide, which induces apoptosis in BHK-21 cells [45]. The experiments were executed in triplicate and in three independent assays. After the treatments, cell viability assays were performed as described in section 2.6.

2.8. Statistical analysis

Cell viability assays expressed as mean± standard deviation were statistically analyzed using the GraphPad Prism 8 software (GraphPad Software Inc., San Diego, CA, USA). Based on the D'Agostino & Pearson normality test, the data were classified as parametric or non-parametric, and the Dunnett's and Kruskall Wallis Dunn tests were used for parametric and nonparametric data analyses, respectively. P values < 0.01 were considered statistically significant.

2.9. Enzymatic Assays

The esterolytic activity of $ExhC_{wt}$ and $ExhC_{mut4}$ against the synthetic substrate Boc-L-Glu-↓OPh was evaluated as previously reported [10, 29, 46]. First, 1 M Boc-L-Glu-↓OPhe was diluted in 1,4-Dioxane (Sigma) and used as a stock substrate solution. Esterolytic reactions occurred under conditions of 50 mM HEPES (pH 7.4) and 2% 1,4-Dioxane (v/v) using 1 µg/mL of toxin, varying the substrate concentration between 1 to 20 mM, with strides of 5 mM. All substrate concentrations were tested in triplicate. Phenol release (molar absorption coefficient = 1.5 mM/cm) from the synthetic substrate Boc-L-Glu-↓OPhe (λ = 270 nm) was measured for 20 minutes, every 15 seconds at a constant temperature of 37 ºC, using the Cary UV-Vis Compact Peltier spectrophotometer (Agilent). The values of V_o, K_M and K_{cat} were estimated from the Hanes-Woolf linearization method.

2.10. Crystallization, data collection, processing and structure determination

The crystals were obtained by vapor diffusion and protein concentrations of 15 mg/mL for $ExhC_{wt}$ (1) and 10 mg/mL for ExhCmut4 (2) in 20 mM Tris-HCl, pH 7.5 buffer containing 100 mM NaCl, equilibrated against reservoir solutions of 0.1 M HEPES pH 7.5; 0.8 M Potassium/Sodium Tartrate (1) and 0.2 M $MgCl_2.6H_2O$; 0.1 M Tris-HCl pH 8.5 and 30% w/v PEG 4.000 with the addition of 0.1 M Sarcosine (2), respectively. X-ray diffraction data was collected from flash frozen crystals at the MANACA beamline (SIRIUS, Campinas, Brazil). The wavelengths of the incident beam were 1.32363 and 0.97718 Å for $ExhC_{wt}$ and $ExhC_{mut4}$, respectively. The 3600 images collected from each $ExhC_{wt}$ and $ExhC_{mut4}$ crystal were indexed using the XDS package [47] and the structure of the $ExhC_{mut4}$ was solved by molecular replacement employing the atomic coordinates of ETE (PDB ID: 5C2Z, 46.56% sequence identity) [48]. The refined coordinates of the structure of $ExhC_{mut4}$, were used to determine the structure of $ExhC_{wt}$ by molecular replacement. The stereochemical parameters were assessed using the MolProbity server [49]. The $ExhC_{wt}$ and $ExhC_{mut4}$ atomic coordinates have been deposited in the Protein Data Bank (PDB) and have been assigned the accession codes 8T3J and 8T3I, respectively.

2.11. Calculation of the surface electrostatic potential and hydrophobicity for ExhC and homologous exfoliative toxins

The crystal structures of $ExhC_{wt}$ and $ExhC_{mut4}$ reported and the AlphaFold2 [36] models of every ExhC single mutant generated here were used to calculate the surface electrostatic

potentials with Pymol 1.8.4.0 [50] and APBS Tools2.1 [51]. Default parameters, including an NaCl concentration of 0.15 M, were set for the calculations. The previous analysis was extended to include other homologous staphylococcal ETs, whose structures were retrieved either from the PDB or were generated by AlphaFold2 [36] based on sequences deposited at UniProtKB. Moreover, the hydrophobicity for the necrosis-linked region of $ExhC_{wt}$ and its mutated variants, as well as for the equivalent regions of all the other analyzed homologous proteins, were generated using Pymol 1.8.4.0 with the script available at https://pymolwiki.org/index.php/Color_h.

3. Results

3.1. Sequence and structural analyses of ExhC

The sequence alignment of full-length ExhC from *M. sciuri* and *S. hyicus* indicate 100% sequence identity (SI) followed by 44.2% identity with ETA from *S. aureus* (44.2% SI). On the other hand, when the alignment is restricted to the $ExhC_{44-92}$ fragment and equivalent regions in other ETs (Fig. 1.A), the highest identity is with ETB from *S. aureus* (59.2% SI). The sequence alignments also indicated significant variations in the properties of the ExhC residues at positions 47, 49, 51 and 89 with respect to other ETs. For example, the ExhC Arg residues at positions 47 and 89 are substituted by neutral residues possessing either polar or aliphatic side-chains in the other ETs (Fig. 1A). Moreover, the latter possess aromatic (Phe/Tyr) and basic residues (Lys) in lieu of the ExhC polar residues N49 and Q51, respectively (Fig. 1A).

The AlphaFold2-generated 3D model of full-length ExhC (Fig. 1.B) highlights the fact that nonconserved residues R47, N49, Q51 and R89 are in structural proximity and located on the surface of the toxin, suggesting that they likely belong to a binding site that triggers necrosis upon recognition of a hitherto unidentified molecular target. These *in silico* analyses were the basis for the design and expression of an ExhC variant with mutations in the four residues mentioned above, which is termed $ExhC_{mut4}$.

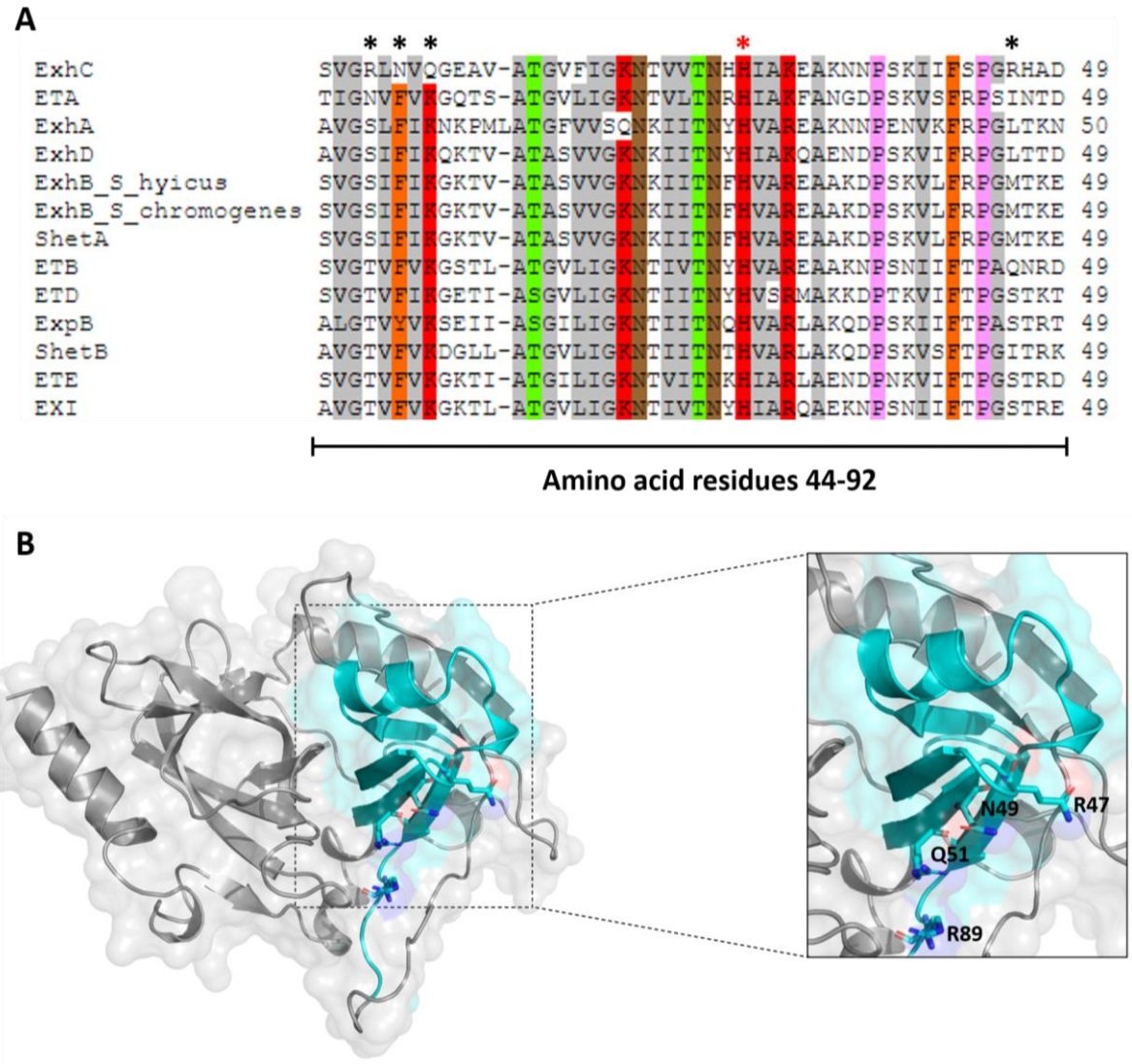


Fig 1. (A) Sequence alignment of ExhC and other ETs comprising residues from 44 to 92. Residues with ≥90% SI are shaded in different colors. The black asterisks highlight residues R47, N49, Q51 and R89 of ExhC, which differ appreciably from the equivalent residues in the other aligned ETs. The red asterisk represents *H57* which forms part of the catalytic triad. **(B)** 3D model of full-length ExhC presented as a cartoon and including the surface area. Closeup of the necrotic fragment (cyan - $ExhC_{44-92}$) [31] with R47, N49, Q51 and R89 represented by sticks.

3.2. Mutated variants of ExhC are dimers in solution

SEC chromatograms indicate that $ExhC_{wt}$ and its mutated variants (MW ~29 kDa) eluted at the same volumes (between 9-12 mL), with approximately-equal maximum elution volumes (EV_{max}) (Fig. 2.A) and similar hydrodynamic diameters (Table 1). These results indicate that the four-point mutations do not significantly alter the overall shapes or the oligomeric states of these enzymes. Likewise, the elution peaks of $ExhC_{wt}$ and of the mutated variants overlap with that of serum albumin (Sigma-Aldrich) (BSA = MW ~66 kDa), which was employed as a

control (Fig. 2.A), thus indicating that the former proteins are dimers in solution, as previously reported in experiments conducted with ExhC from *M. sciuri* [32].

These results were reinforced by the analysis of $ExhC_{wt}$ and $ExhC_{mut4}$ by SDS-PAGE (Fig. 2.B) and BN-PAGE (Fig. 2.C). While both enzymes migrate as ~30 kDa molecules under denaturing conditions, their molecular weights roughly double under native conditions, i.e., BSA, $ExhC_{wt}$ and $ExhC_{mut4}$ exhibited similar migration profiles, clearly different from bovine pancreatic β-trypsin (Sigma-Aldrich) (MW ~24 kDa). Thus, the four-point mutations in ExhC are not likely to be located at the dimer interface, as the mutations performed here do not interfere the dimerization process.

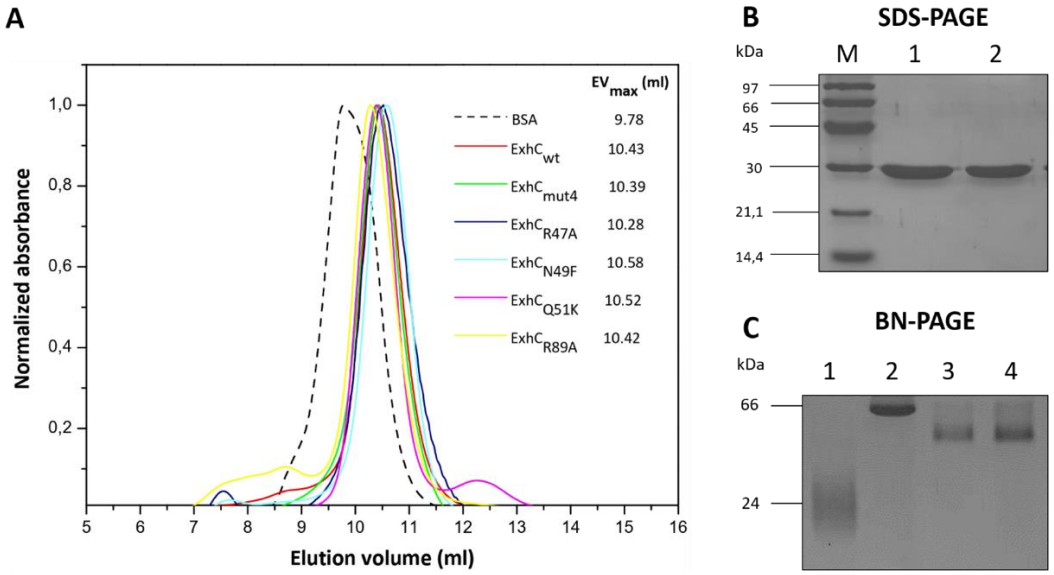


Fig 2. (A) Normalized SEC profiles of $ExhC_{wt}$ and mutants (solid lines) together with BSA (dashed line). Absorbance was measured at 280 nm. EV_{max} values for each protein are indicated within the graph. **(B)** SDS-PAGE of $ExhC_{wt}$ (1) and $ExhC_{mut4}$ (2) using the protein marker LMW-SDS Marker KIT—GE Healthcare (M). **(C)** BN-PAGE of $ExhC_{wt}$ (3) and $ExhC_{mut4}$ (4) using bovine pancreatic β-trypsin (1) and BSA (2) as molecular markers.

Table 1. Hydrodynamic properties (nm) of ExhC and its mutant variants in TBS buffer using DLS.

Protein	Hydrodynamic diameter (nm)
$ExhC_{wt}$	5.77
$ExhC_{mut4}$	5.69
$ExhC_{R47A}$	6.27
$ExhC_{N49F}$	7.63
$ExhC_{Q51K}$	5.36
$ExhC_{R89A}$	5.14

3.3. *In vitro* inhibition of ExhC necrotic activity in cell line BHK-21 by point mutations

In vitro assays in the presence of the apoptosis inhibitor Z-VAD-FMK (Fig 3.A) demonstrated that cell death caused by $ExhC_{wt}$ is due to necrosis, as previously described by Li et al. (2011a) [31]. Treatment with $ExhC_{wt}$ in the absence (-) and presence (+) of the apoptosis inhibitor, showed similar cell death results, i.e., pretreatment with Z-VAD-FMK did not favor the increase in cell viability, unlike the results obtained following the addition of H_2O_2 (positive control) (Fig 3.A), in which the use of Z-VAD-FMK decreased the percentage of cell death by apoptosis caused by H_2O_2 in BHK-21 cells [45].

Furthermore, cytotoxicity assays indicated that significant death of BHK-21 cells is caused following the addition of 15 µM $ExhC_{wt}$ and that this effect depends on the toxin concentration (Fig. 3.B). In fact, BHK-21 cells treated with 15 µM, 30 µM and 60 µM of $ExhC_{wt}$ exhibited reduction in cell viability of approximately 40%, 60% and 70%, respectively (Fig 3.B). Conversely, treatment of cells with $ExhC_{mut4}$ did not induce significant cell death. In this case, approximately 93% to 96% of the cells remained viable with the addition of concentrations between 15-60 µM of $ExhC_{mut4}$ (Fig 3.B). Therefore, the absence of amino acids R47, N49, Q51 and R89 from ExhC abrogated the necrotic activity within the tested toxin concentration range.

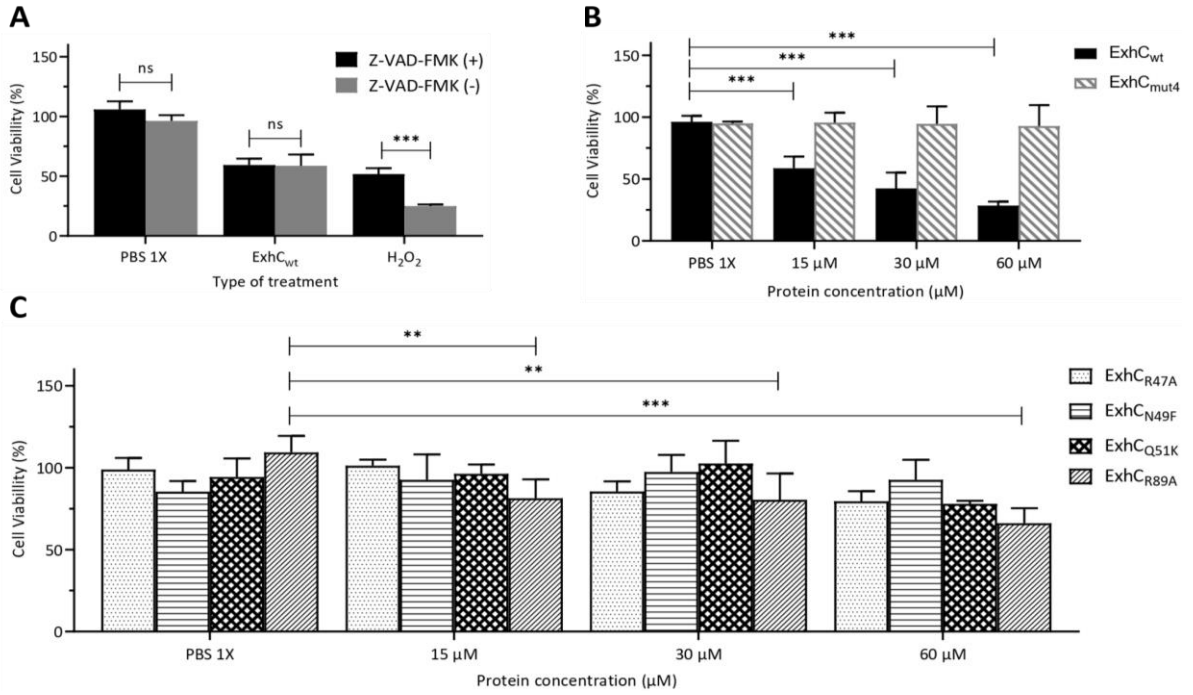


Fig 3. Viability assays of BHK-21 cells. **(A)** Comparison between cells with/without Z-VAD-FMK pretreatment with subsequent addition of: PBS 1X, 60 µM of $ExhC_{wt}$ and 100 µM of H_2O_2. **(B)** Treatment with $ExhC_{wt}$ and $ExhC_{mut4}$ at concentrations of 15, 30 and 60 µM. **(C)** Treatment with $ExhC_{R47A}$, $ExhC_{N49F}$, $ExhC_{Q51K}$ and $ExhC_{R89A}$

at 15, 30 and 60 µM concentrations. In all cases (A, B and C) the plotted results represent three independent MTT experiments shown as mean ± SD. Percentages of cell viability were calculated from the control experiment in the absence of PBS 1X (only in culture medium). Statistically significant differences between the control experiment and treated samples were calculated using Dunnett's test or Kruskall Wallis Dunn test. **, $p<0.01$; *** $p < 0.001$; ns, not significant.

From this initial design, four variants with point mutations were also tested (Fig 3.C) and it was observed that $ExhC_{R47A}$, $ExhC_{N49F}$ and $ExhC_{Q51K}$ did not cause significant decrease in cell viability when compared to the control experiments. These results suggest that residues R47, N49 and Q51 are individually crucial for the necrotic activity of ExhC from *M. sciuri*. On the other hand, cells treated with 15 µM, 30 µM and 60 µM of $ExhC_{R89A}$ showed statistically significant differences when compared to the control group (Fig 3.C), although the necrotic potential of this mutant is lower than that of $ExhC_{wt}$. In fact, the addition of 15 µM, 30 µM and 60 µM of $ExhC_{R89A}$ decreased cell viability by approximately 19%, 20% and 34%, respectively. Remarkably, at 15 µM of $ExhC_{wt}$, a greater number of cells died than at $ExhC_{R89A}$ concentrations of 30 and 60 µM (Table. S1, supplementary material), thereby indicating that residue R87 is less relevant for the observed necrotic potential of ExhC when compared to residues R47, N49 and Q51, but still interferes with the effectiveness of this mechanism of necrotic action.

Other cell lines were tested in the presence of $ExhC_{wt}$ and no cell death was caused by the enzyme in cell lines HK-2, 786-O, ACHN and PK-15 (Fig. S1, supplementary material). Furthermore, tests with ETA (1) and ETE (2) from *S. aureus* were also performed to evaluate the necrotic potential of other ETs (Fig. S2, supplementary material). However, the ET with the highest primary sequence identity with ExhC (1) and the one cleaving the widest range of Dsg1 isoforms (2) [7] did not cause cell death in BHK-21, which raises the probability of ExhC being the only ET possessing necrotic activity in mammalian cell lines.

In all cases, CD spectra indicated that the protein samples used in the *in vitro* tests with cell lines were properly folded (Table S2, supplementary material), i.e., these proteins were present in their functional states during the experiments. In addition, data obtained for $ExhC_{wt}$ and its mutants (Fig. S3, supplementary material) indicated similar percentages of secondary structures (Table S2, supplementary material). These results are expected since they differ only by the mutation of four amino acid residues. However, when comparing the percentage of secondary structures of $ExhC_{wt}$ and its mutant variants with the secondary structures of ETA and ETE

(Table S2, supplementary material), changes were observed in relatively significant percentages.

3.4. Esterolytic activity of ExhC$_{wt}$ and ExhC$_{mut4}$

The ability of ETs to cleave the substrate Boc-L-Glu-↓OPh [10, 13, 29] allowed us to assess whether the mutations in the necrosis-linked region affected enzymatic activity. During the enzymatic assays, we noticed that, as a consequence of the low solubility of Boc-L-Glu-↓OPh, saturation of ExhC$_{wt}$ and ExhC$_{mut4}$ activity was not attained. However, even in this scenario, it can be inferred from the V_o vs [S] graphs that the enzymes follow the typical Michaelian behavior as the substrate concentration increases (Fig 4.A). From these results, Hanes-Woolf plots [52] were generated to calculate the catalytic efficiencies for both proteins (Fig 4.B and Table 2). The results show that ExhC$_{mut4}$ retained the esterolytic activity and that its catalytic efficiency was just 1.3 times lower than that of ExhC$_{wt}$. Therefore, our results demonstrate that the four mutations have minimal effect enzyme activity and, consequently, the necrotic and catalytic activities of ExhC may be considered largely independent functions.

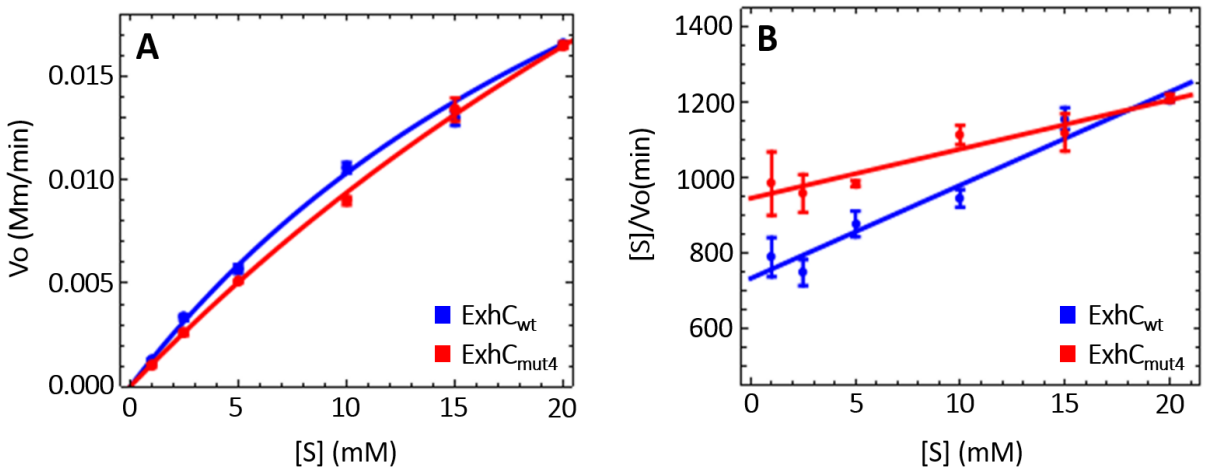


Fig 4. Hydrolysis of Boc-L-Glu-↓OPhe by ExhC$_{wt}$ and ExhC$_{mut4}$ using the enzyme kinetics model by Michaelis-Menten **(A)** and Hanes-Woolf **(B)**. Plotted results represent three independent experiments shown as mean ± SD.

Table 2. Kinetic parameters for the hydrolysis of Boc-L-Glu-↓OPhe by ExhC$_{wt}$ and ExhC$_{mut4}$.

Enzyme	K_{cat}/K_M ($mM^{-1}\cdot s^{-1}$)[a]
ExhC$_{wt}$	0.60±0.02
ExhC$_{mut4}$	0.46±0.02

[a]The mean catalytic efficiency (K_{cat}/K_M) ± SD is determined from the inverse of the *y*-intercept in the Hanes-Woolf plots.

3.5. The crystallographic structures of $ExhC_{wt}$ and $ExhC_{mut4}$

The crystallographic structures of $ExhC_{wt}$ and $ExhC_{mut4}$ were refined at 2.71 Å and 1.57 Å (Table 3), respectively, and represent the first structures of ETs with specificity for porcine and murine Dsgs1. The $ExhC_{mut4}$ crystallographic data correspond to the highest resolution obtained for ETs to this date.

The $ExhC_{wt}$ crystals were tetragonal (I 4 2 2) whereas $ExhC_{mut4}$ crystals were monoclinic (P 1 21 1) (Table 3). The Matthew coefficients calculated for $ExhC_{mut4}$ and $ExhC_{wt}$ were 2.02 and 3.23 A^3/Da, and indicated that the asymmetric units (AU) consisted of a dimeric and a trimeric assembly with solvent contents of 39.11% and 61.95%, respectively (Table 3) (Fig. 5).

As $ExhC_{wt}$ and $ExhC_{mut4}$ are dimers in solution (described in item 3.2), the protein-protein interactions observed in the AU of the crystals were evaluated by the PISA server [53], however, the complexes observed (Fig. 5) were not considered to be stable in solution (Table 4) and, therefore, the arrangement of these molecules in the AU is probably a crystallization artifact.

Table 3. Crystallographic data collection and refinement statistics of $ExhC_{wt}$ and $ExhC_{mut4}$.

Data collection and refinement	**$ExhC_{wt}$**	**$ExhC_{mut4}$**
Resolution range (Å)	47.28 - 2.71 (2.81 - 2.71)	39.55 - 1.57 (1.63 - 1.57)
Space group symmetry	I 4 2 2	P 1 21 1
Cell dimensions:		
a, b, c (Å)	167.09, 167.09, 157.74	45.15, 71.40, 71.40
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 94.71, 90.00
Total number of reflections	799434 (79981)	411983 (37881)
Number of unique reflections	30572 (3017)	63061 (6315)
Multiplicity	26.2 (26.5)	6.5 (6.0)
Completeness range (%)	99.92 (99.83)	99.90 (99.65)
Mean I/σI	8.40 (0.79)	15.17 (1.14)
R-merge[a]	0.3359 (3.813)	0.06174 (1.278)
R-meas	0.3427 (3.887)	0.06707 (1.402)
Correlation coefficient ½	0.995 (0.45)	0.999 (0.506)
R(work)/R(free)[b]	0.2113/0.2701	0.1813/0.2229
Number of protein chains in the asymmetric unit	3	2
Number of solvent water molecules	56	283
R.M.S:		
Bond lengths (Å)	0.009	0.016
Bond angles (°)	1.71	2.19
Ramachandran analysis:		
Favored (%)	95.35	97.90
Allowed (%)	4.65	2.10
Disallowed (%)	0.00	0.00

Average B-factor (Å2)	83.15	29.40
PDB accession code	8T3J	8T3I

[a]Rmerge I(h)I {I(h)}/{I(h)}, where Ih is the observed intensity of the ih measurement of reflection h, and {I(h)} is the mean intensity of reflection h calculated after scaling.

[b]R(free) is the R(work) based on 5% of the data excluded from the refinement.

Values in parentheses represents the highest-resolution shell.

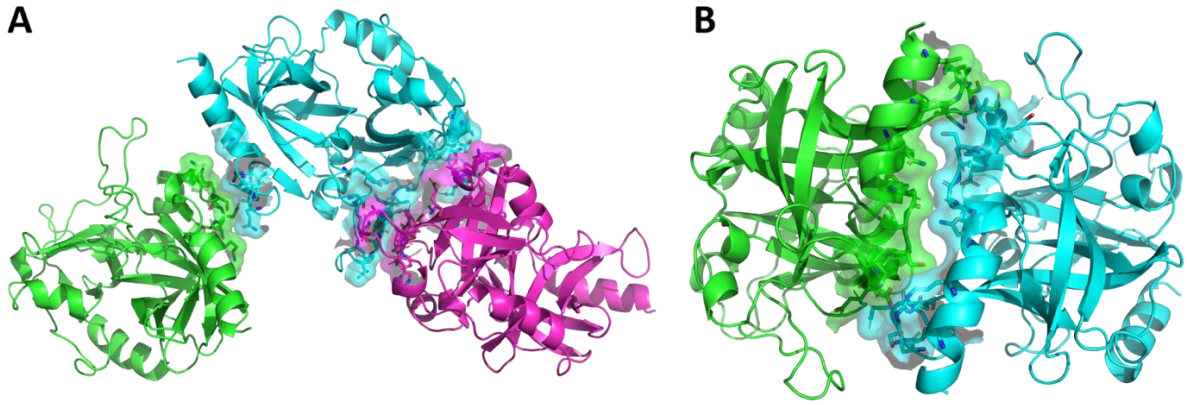


Figure 5. Representation of the protein assemblies in the AUs of $ExhC_{wt}$ (A) e $ExhC_{mut4}$ (B). Each chain is shown in a secondary structure with different colors: green, cyan and magenta. Interface regions of protein-protein interactions are depicted as surfaces.

Table 4. Stability of the interface of oligomers present in the AU of the crystalline structures $ExhC_{wt}$ and $ExhC_{mut4}$ according to the ΔG^{diss} values.

Structure	**ΔG^{diss} (kcal/mol)**	**Oligomer stability**
$ExhC_{wt}$	-16.3	Not stable
$ExhC_{mut4}$	-9.3	Not stable

The structures of $ExhC_{wt}$ and $ExhC_{mut4}$ were also analyzed by the CavityPlus server [54], and the results indicated the formation of nine superficial cavities. Among them, residues R47, N49, Q51 and R89 stand out in $ExhC_{wt}$, which, together with other residues of the target protein, form a cavity opposite the cavity that contains the active site formed by *H57*, *D102* and *S195*, which are responsible for the epidermolytic activity (Fig. 6). The detection of a cavity containing all the residues that, when absent, abrogate or reduce the necrotic potential of ExhC, reinforces the concept that this region of the enzyme is likely to interact with molecular targets of BHK-21 cells.

From the comparison of the necrosis-linked regions in $ExhC_{wt}$ and $ExhC_{mut4}$ by the ProteinPlus server [55] (Table S3, supplementary material), it can be concluded that the absence of residues

R47, N49, Q51 and R89 in $ExhC_{mut4}$ caused a significant change in the properties of the affected protein region, which may hinder the interaction with yet unknown cell target(s).

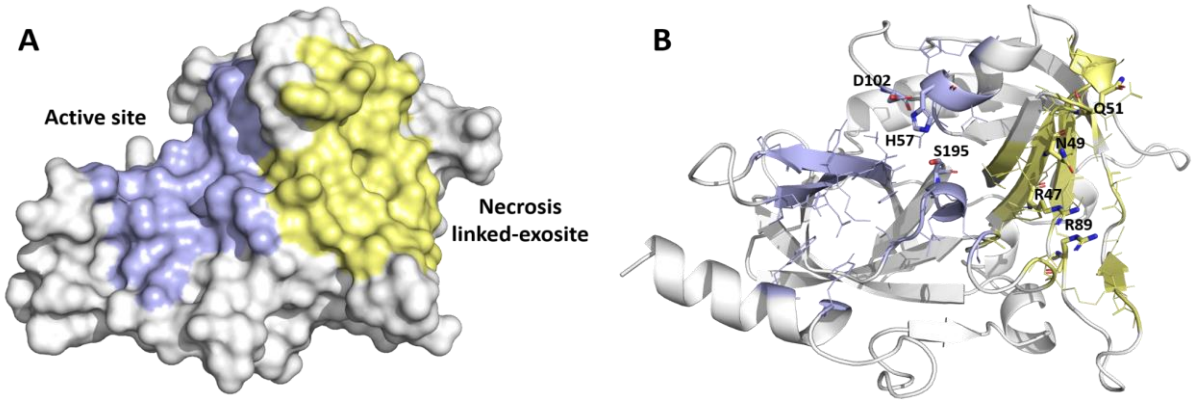


Fig 6. Representation of the cavities corresponding to the active site (blue) and the necrosis linked-region (yellow) detected using the crystal structure of $ExhC_{wt}$. Cavities were detected using the CavityPlus server. Representation of the surface area **(A)** and cartoon representation of the protein backbone, with residues whose side-chains form both cavities depicted as lines **(B)**. In sticks, the residues of the catalytic triad (*H57, D102* and *S195*) and those involved in necrotic activity (R47, N49, Q51 and R89).

3.6. Analysis of physico-chemical surface necrosis-liked region of ExhC and its homologous regions in other ETs

Two of the mutations described here involved positively-charged residues substituted by neutral residues, e.g., R47A and R89A, and a third mutation, a neutral residue substituted by a positively-charged one (Q51K). Therefore, the charge density on the surface of the necrotic-linked region might play an important role in the interaction with the molecule(s) responsible for triggering necrosis. Fig. 7 represents the surface electrostatic potential on the necrosis-linked region of $ExhC_{wt}$ and its mutated variants. As can be observed, charge-modifying mutations R47A and R89A alter the charge density of the necrosis-linked region by turning it less positively-charged, whereas the impact on the electrostatic surface potential of other mutations, e.g., N49F and Q51K, is less apparent. On the other hand, it can be shown that the mutations also alter the hydrophobicity (Fig. 8), which is another property relevant for the binding to other molecules. This is particularly the case for mutations N49F and Q51K, as they seem to affect the pocket hydrophobicity more than the electrostatic potential.

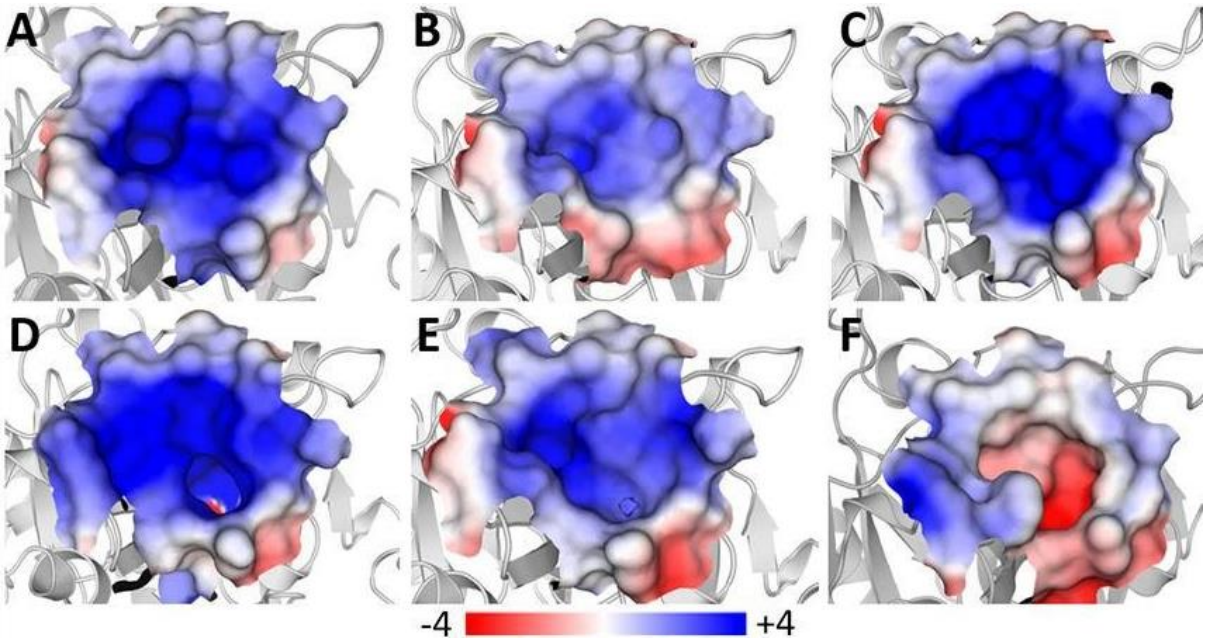


Fig 7. Structural representation of the electrostatic surface potential of the necrosis-linked region of **(A)** $ExhC_{wt}$, (**B)** $ExhC_{R47A}$, (**C)** $ExhC_{N49F}$, (**D)** $ExhC_{Q51K}$, (**E)** $ExhC_{R89A}$ and (**F)** $ExhC_{mut4}$. The protein surfaces are colored according to the electrostatic charge density, expressed in k_bT/e units, where k_b, T and e stand for the Boltzmann's constant, the temperature (298.15 K) and the electron charge, respectively.

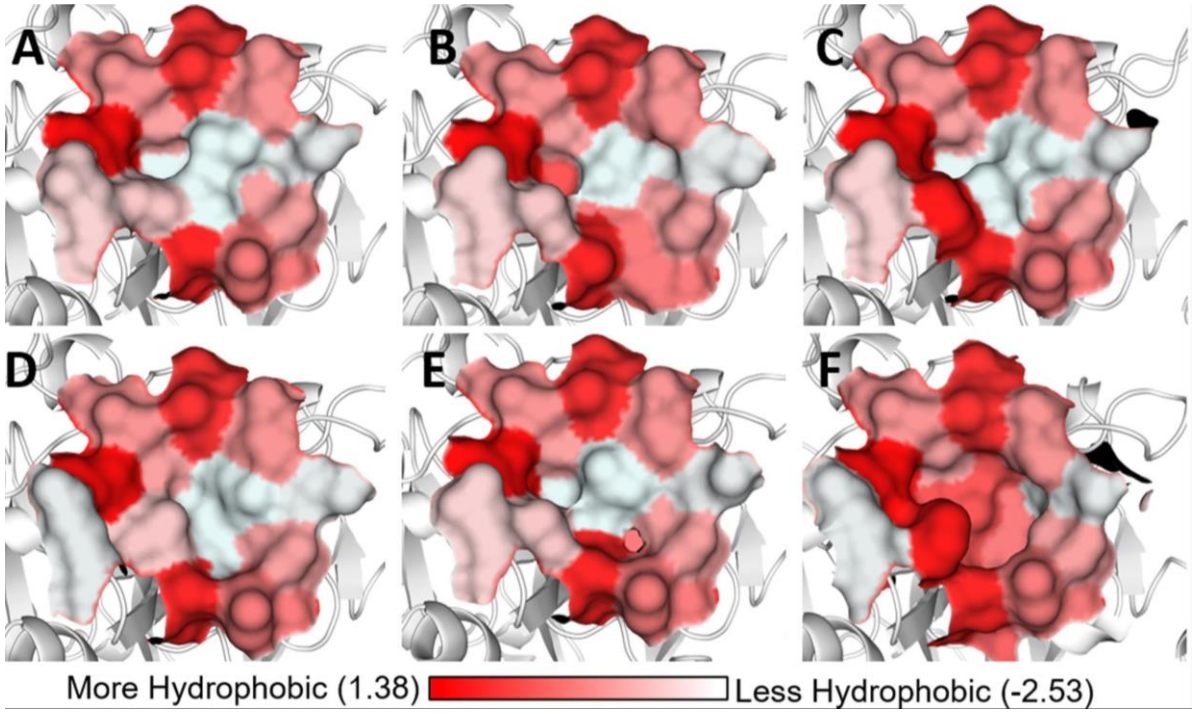


Fig 8. Structural representation of the necrosis-linked region of **(A)** $ExhC_{wt}$, (**B)** $ExhC_{R47A}$, (**C)** $ExhC_{N49F}$, (**D)** $ExhC_{Q51K}$, (**E)** $ExhC_{R89A}$ and (**F)** $ExhC_{mut4}$, with the surface colored by residue hydrophobicity, according to the gradient below the figure. The normalized consensus hydrophobicity scale for each protein residue was employed [56].

As mentioned earlier, ExhC is the only ET for which *in vitro* cell necrosis activity has been confirmed and it was shown here that this property is determined by specific residues located in the necrosis-linked region. Our results also demonstrated that ETA and ETE do not possess necrotic potential when incubated with BHK-21 cells. The analysis of the electrostatic potentials of the equivalent regions in other ETs revealed that most of them, except for ExhA and ExhD, do not possess a high density of positive charge in these regions (Fig. S4, supplementary material). In particular, non-necrotic toxins such as ETA and ETE possess negatively-charged patches on the analyzed surface, thus making it unlikely the charge complementarity with the same putative molecular(s) partner(s) of ExhC associated with the necrotic effect. Significant differences in the hydrophobicity of the region associated with the necrosis-linked region of ExhC and the corresponding regions in other ETs were observed (Fig. S5, supplementary material). It becomes clear that the analyzed region is less hydrophobic in ExhC than in most of the ETs represented in the figure. In conclusion, the results suggest that the particularly high density of positively-charged residues encountered at the necrotic-linked region of ExhC is crucial for the observed necrotic potential and can also indicate the reason for the absence of necrotic activity in other staphylococcal ETs.

4. Discussion

Following the initial description of SSSS [57], approximately one hundred years elapsed before the enzymes which triggered this activity were discovered [58], and the substrate of these enzymes was identified only three decades later [59]. Even after the determination of numerous high-resolution three-dimensional structures of these toxins [9, 11-13, 48] no inhibitors have yet been identified and the specificities and mechanisms of action have not been determined. Through the discovery of the moonlighting [28] properties of these toxins [29-31], such as the unusual ability of ExhC from *M. sciuri* to induce necrosis in multiple cell lines *in vitro* [31], the need to understand the mechanism of action of these factors of virulence in the host has become even more relevant.

The necrotic property of ExhC, not observed previously in any other ET, suggests that ExhC may participate in diverse functions during *M. sciuri* infections. From the knowledge of the fragment responsible for *in vitro* cell necrosis, comprising residues from 44 to 92 [31], we were able to identify four non-conserved residues implicated in the necrotic potential of ExhC. This brings us closer to deciphering the specific interactions of this mechanism of action.

The primary sequence alignment with $ExhC_{44-92}$ indicated the importance of four residues with distinct physico-chemical properties from those observed at the analogous positions in other ETs, which established the basis for the design of an ExhC variant with four mutations, R47A, N49F, Q51K and R89A which exhibited no necrotic activity. *In vitro* assays indicated that point mutations at residues R47, N49 and Q51 eliminated the necrotic potential, whereas mutation of residue R89 resulted in a protein with reduced necrotic activity.

ExhC from *M. sciuri* forms homodimers in solution [32]. Interestingly, the mutations conducted here for this enzyme, which drastically affected its necrotic potential, did not alter its oligomeric state. These results indicate that the mutated residues are not located at the protein-protein complex interface required for forming the dimer, and also suggest that the presence of the monomeric form is not essential for interaction with cellular target(s) that cause necrosis. Data that differ from those proposed for the exfoliative activity of ETE from *S. aureus*, in which monomerization of the ETE homodimer might play a key role in the cleavage of Dsg1 [13].

Previous experiments by Li et al. (2011a) [31] concluded that an ExhC variant with a mutation of the catalytic His residue did not cause exfoliation in the epidermis of newborn rats, but continued to cause cell necrosis. Further evidence presented here, highlights the orthogonality of the cell necrosis and skin exfoliation caused by this toxin, since both $ExhC_{wt}$ and $ExhC_{mut4}$ have the ability to cleave the substrate Boc-L-Glu-↓OPh with similar catalytic efficiencies. Although the esterolytic activity can be taken as an indicator of the proteolytic activity in general, assays with porcine or murine Dsg1 and $ExhC_{mut4}$ need to be conducted to confirm this finding.

The crystallographic structures of $ExhC_{wt}$ and $ExhC_{mut4}$ provided crucial, additional data for understanding the factors important for the observed necrotic potential, indicating that the residues R47A, N49F, Q51K and R89A responsible for cell necrosis, together with other amino acid residues, are located in a cavity on the surface of the protein which is distinct from the catalytic site. This cavity was termed the necrosis-linked region which may be involved in triggering necrosis upon binding of an effector molecule to the region. The mutations in $ExhC_{mut4}$ resulted in a change in the contact surface and in the volume of this cavity, which can potentially hinder the interaction with the target(s).

Furthermore, electrostatic surface analyses showed that the necrosis-linked region of ExhC is highly-positively charged, different from the charge of analogous regions in other ETs, such as ETA and ETE of *S. aureus*. Interestingly, ETA and ETE were also assayed with BHK-21 cells and did not cause necrosis. Changes in the hydrophobicity of the necrosis-linked region were

also evidenced, underscoring the differences in physico-chemical properties of this region in the analyzed proteins. This in turn might explain why necrotic activity is not a general characteristic of the ETs.

Li et al. (2011a) demonstrated that monoclonal antibodies produced from full-length ExhC can block the necrotic activity, suggesting that the identified pocket functions as a necrosis-linked site that mediates the recognition of an unknown target within the cell, which in turn triggers the necrotic events. On the other hand, truncated fragments of ExhC retained necrotic potential similar to that of the intact protein [31]. The predicted structures of the smallest necrosis-inducing fragment (data not shown) revealed that, in contrast to the full-length protein, the surface in the vicinity of the key residues do not form surface pockets. This second finding is compatible with an alternative mechanism in which necrosis is a molecular recognition phenomena that is not significantly influenced by the integrity and shape of the site containing the mutated residues but is limited to a few structural elements that are preserved even following truncation. Overall, due to the diverse functional criteria further studies warranted to establish the mode of action by which ExhC triggers cell necrosis.

5. Conclusion

Through bioinformatics, recombinant protein production, spectroscopic analysis, and in vitro assays, we have identified key residues and structural features that are crucial for the expression of necrotic activity caused by ExhC. This knowledge will prove valuable in the design of inhibitory molecules with potential clinical applications.

6. Acknowledgements

We are very grateful to Dr. Caio C. Oliveira and Me. Lucas dos Santos Cunha for the synthesis of the substrate Boc-L-Glu-↓Oph. We also thank Dr. Yves Le Loir for providing the recombinant ETE plasmid. This research was supported by grants from the São Paulo Research Foundation (FAPESP) 2019/10230-0, 2020/13921-0, 2020/08615-8, 2020/10214-1, 2022/03901-8 and 2022/14362-0, by grant 309940/2019-2 from CNPq and from UNESP/PROPE.

7. References

[1] S. Ladhani, C.L. Joannou, D.P. Lochrie, R.W. Evans, S.M. Poston, Clinical, microbial, and biochemical aspects of the exfoliative toxins causing staphylococcal scalded-skin syndrome, J Clinical microbiology reviews 12(2) (1999) 224-242.

[2] R.B. Mariutti, N.R. Tartaglia, N. Seyffert, T.L. de Paula Castro, R.K. Arni, V.A. Azevedo, Y. Le Loir, K. Nishifuji, Exfoliative toxins of *Staphylococcus aureus*, J The Rise of Virulence Antibiotic Resistance in *Staphylococcus aureus* (2017).
[3] P.W. O'Toole, T.J. Foster, Nucleotide sequence of the epidermolytic toxin A gene of *Staphylococcus aureus*, J Bacteriol 169(9) (1987) 3910-5.
[4] C.Y. Lee, J.J. Schmidt, A.D. Johnson-Winegar, L. Spero, J.J. Iandolo, Sequence determination and comparison of the exfoliative toxin A and toxin B genes from *Staphylococcus aureus*, J Bacteriol 169(9) (1987) 3904-9.
[5] T. Yamaguchi, K. Nishifuji, M. Sasaki, Y. Fudaba, M. Aepfelbacher, T. Takata, M. Ohara, H. Komatsuzawa, M. Amagai, M. Sugai, Identification of the *Staphylococcus aureus* etd pathogenicity island which encodes a novel exfoliative toxin, ETD, and EDIN-B, Infect Immun 70(10) (2002) 5835-45.
[6] C. Le Marechal, J. Jardin, G. Jan, S. Even, C. Pulido, J.M. Guibert, D. Hernandez, P. Francois, J. Schrenzel, D. Demon, E. Meyer, N. Berkova, R. Thiery, E. Vautor, Y. Le Loir, *Staphylococcus aureus* seroproteomes discriminate ruminant isolates causing mild or severe mastitis, Vet Res 42(1) (2011) 35.
[7] I. Imanishi, A. Nicolas, A.B. Caetano, T.L.P. Castro, N.R. Tartaglia, R. Mariutti, E. Guedon, S. Even, N. Berkova, R.K. Arni, N. Seyffert, V. Azevedo, K. Nishifuji, Y. Le Loir, Exfoliative toxin E, a new *Staphylococcus aureus* virulence factor with host-specific activity, Sci Rep 9(1) (2019) 16336.
[8] I.V. Demidyuk, K.N. Chukhontseva, S.V. Kostrov, Glutamyl Endopeptidases: The Puzzle of Substrate Specificity, Acta Naturae 9(2) (2017) 17-33.
[9] G.M. Vath, C.A. Earhart, D.D. Monie, J.J. Iandolo, P.M. Schlievert, D.H. Ohlendorf, The crystal structure of exfoliative toxin B: a superantigen with enzymatic activity, Biochemistry 38(32) (1999) 10239-46.
[10] J. Cavarelli, G. Prevost, W. Bourguet, L. Moulinier, B. Chevrier, B. Delagoutte, A. Bilwes, L. Mourey, S. Rifai, Y. Piemont, D. Moras, The structure of *Staphylococcus aureus* epidermolytic toxin A, an atypic serine protease, at 1.7 A resolution, Structure 5(6) (1997) 813-24.
[11] G.M. Vath, C.A. Earhart, J.V. Rago, M.H. Kim, G.A. Bohach, P.M. Schlievert, D.H. Ohlendorf, The structure of the superantigen exfoliative toxin A suggests a novel regulation as a serine protease, J Biochemistry 36(7) (1997) 1559-1566.
[12] A.C. Papageorgiou, L.R. Plano, C.M. Collins, K.R. Acharya, Structural similarities and differences in *Staphylococcus aureus* exfoliative toxins A and B as revealed by their crystal structures, Protein Sci 9(3) (2000) 610-8.
[13] C. Gismene, J.E. Hernandez Gonzalez, A.R.N. Santisteban, A.F. Ziem Nascimento, L. Dos Santos Cunha, F.R. de Moraes, C.L.P. de Oliveira, C.C. Oliveira, P. Jocelan Scarin Provazzi, P.G. Pascutti, R.K. Arni, R. Barros Mariutti, *Staphylococcus aureus* Exfoliative Toxin E, Oligomeric State and Flip of P186: Implications for Its Action Mechanism, Int J Mol Sci 23(17) (2022).
[14] C.M. Nusman, C. Blokhuis, D. Pajkrt, D.H.J.A. Visser, Staphylococcal Scalded Skin Syndrome in Neonates: Case Series and Overview of Outbreaks, J Antibiotics 12(1) (2022) 38.
[15] P.W. O'Toole, T.J. Foster, Molecular cloning and expression of the epidermolytic toxin A gene of *Staphylococcus aureus*, J Microbial pathogenesis 1(6) (1986) 583-594.
[16] M. Bukowski, B. Wladyka, G. Dubin, Exfoliative toxins of *Staphylococcus aureus*, Toxins (Basel) 2(5) (2010) 1148-65.
[17] Q. Qian, J. Shen, A. Vinturache, G. Ding, Staphylococcal scalded skin syndrome in a child, J The Lancet Infectious Diseases 23(5) (2023) 634.
[18] E. Mueller, M. Haim, T. Petnehazy, B. Acham-Roschitz, M. Trop, An innovative local treatment for staphylococcal scalded skin syndrome, Eur J Clin Microbiol Infect Dis 29(7) (2010) 893-7.
[19] M. Madhaiyan, J.S. Wirth, V.S. Saravanan, Phylogenomic analyses of the Staphylococcaceae family suggest the reclassification of five species within the genus Staphylococcus as heterotypic synonyms, the promotion of five subspecies to novel species, the taxonomic reassignment of five Staphylococcus species to Mammaliicoccus gen. nov., and the formal assignment of Nosocomiicoccus to the family Staphylococcaceae, Int J Syst Evol Microbiol 70(11) (2020) 5926-5936.
[20] Y. Fudaba, K. Nishifuji, L.O. Andresen, T. Yamaguchi, H. Komatsuzawa, M. Amagai, M. Sugai, *Staphylococcus hyicus* exfoliative toxins selectively digest porcine desmoglein 1, J Microbial pathogenesis 39(5-6) (2005) 171-176.

[21] F.M. Aarestrup, H.C. Wegener, Association between production of fibrinolysin and virulence of *Staphylococcus hyicus* in relation to exudative epidermitis in pigs, J Acta Veterinaria Scandinavica 38 (1997) 295-297.
[22] P. Ahrens, L.O. Andresen, Cloning and sequence analysis of genes encoding *Staphylococcus hyicus* exfoliative toxin types A, B, C, and D, J Bacteriol 186(6) (2004) 1833-7.
[23] L. Schwarz, I. Loncaric, R. Brunthaler, C. Knecht, I. Hennig-Pauka, A. Ladinig, Exudative Epidermitis in Combination with Staphylococcal Pyoderma in Suckling Piglets, Antibiotics (Basel) 10(7) (2021).
[24] S. Chen, Y. Wang, F. Chen, H. Yang, M. Gan, S.J. Zheng, A highly pathogenic strain of *Staphylococcus sciuri* caused fatal exudative epidermitis in piglets, PLoS One 2(1) (2007) e147.
[25] L.O. Andresen, P. Ahrens, L. Daugaard, V. Bille-Hansen, Exudative epidermitis in pigs caused by toxigenic *Staphylococcus chromogenes*, Vet Microbiol 105(3-4) (2005) 291-300.
[26] K. Iyori, K. Futagawa-Saito, J. Hisatsune, M. Yamamoto, M. Sekiguchi, K. Ide, W.G. Son, T. Olivry, M. Sugai, T. Fukuyasu, T. Iwasaki, K. Nishifuji, *Staphylococcus pseudintermedius* exfoliative toxin EXI selectively digests canine desmoglein 1 and causes subcorneal clefts in canine epidermis, Vet Dermatol 22(4) (2011) 319-26.
[27] K. Iyori, J. Hisatsune, T. Kawakami, S. Shibata, N. Murayama, K. Ide, M. Nagata, T. Fukata, T. Iwasaki, K. Oshima, M. Hattori, M. Sugai, K. Nishifuji, Identification of a novel *Staphylococcus pseudintermedius* exfoliative toxin gene and its prevalence in isolates from canines with pyoderma and healthy dogs, FEMS Microbiol Lett 312(2) (2010) 169-75.
[28] N. Singh, N. Bhalla, Moonlighting Proteins, Annu Rev Genet 54 (2020) 265-285.
[29] J.V. Rago, G.M. Vath, G.A. Bohach, D.H. Ohlendorf, P.M. Schlievert, Mutational analysis of the superantigen staphylococcal exfoliative toxin A (ETA), J Immunol 164(4) (2000) 2207-13.
[30] S.R. Monday, G.M. Vath, W.A. Ferens, C. Deobald, J.V. Rago, P.J. Gahr, D.D. Monie, J.J. Iandolo, S.K. Chapes, W.C. Davis, D.H. Ohlendorf, P.M. Schlievert, G.A. Bohach, Unique superantigen activity of staphylococcal exfoliative toxins, J Immunol 162(8) (1999) 4550-9.
[31] H. Li, Y. Wang, L. Ding, S.J. Zheng, *Staphylococcus sciuri* exfoliative toxin C (ExhC) is a necrosis-inducer for mammalian cells, PLoS One 6(7) (2011) e23145.
[32] H. Li, X. Li, Y. Lu, X. Wang, S.J. Zheng, Staphylococcus sciuri exfoliative toxin C is a dimer that modulates macrophage functions, Can J Microbiol 57(9) (2011) 722-9.
[33] R.C. Edgar, MUSCLE: a multiple sequence alignment method with reduced time and space complexity, BMC Bioinformatics 5 (2004) 113.
[34] P. Stothard, The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences, Biotechniques 28(6) (2000) 1102, 1104.
[35] K. Nishifuji, M. Sugai, M. Amagai, Staphylococcal exfoliative toxins: "molecular scissors" of bacteria that attack the cutaneous defense barrier in mammals, J Dermatol Sci 49(1) (2008) 21-31.
[36] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature 596(7873) (2021) 583-589.
[37] G.L. Rosano, E.A. Ceccarelli, Recombinant protein expression in *Escherichia coli*: advances and challenges, J Front Microbiol, 5 (2014) 172.
[38] M. Shehadul Islam, A. Aryasomayajula, P.R.J.M. Selvaganapathy, A review on macroscale and microscale cell lysis methods, 8(3) (2017) 83.
[39] E.L. Rodriguez, S. Poddar, S. Iftekhar, K. Suh, A.G. Woolfork, S. Ovbude, A. Pekarek, M. Walters, S. Lott, D.S. Hage, Affinity chromatography: A review of trends and developments over the past 50 years, J Chromatogr B Analyt Technol Biomed Life Sci 1157 (2020) 122332.
[40] R.R. Burgess, A brief practical review of size exclusion chromatography: Rules of thumb, limitations, and troubleshooting, Protein Expr Purif 150 (2018) 81-85.
[41] K. Peters, H.P. Braun, Comparative analyses of protein complexes by blue native DIGE, Methods Mol Biol 854 (2012) 145-54.

[42] A.R. Karow, J. Gotzl, P. Garidel, Resolving power of dynamic light scattering for protein and polystyrene nanoparticles, Pharm Dev Technol 20(1) (2015) 84-9.
[43] N.J. Greenfield, Using circular dichroism spectra to estimate protein secondary structure, Nat Protoc 1(6) (2006) 2876-90.
[44] N. Sreerama, R.W. Woody, On the analysis of membrane protein circular dichroism spectra, Protein Sci 13(1) (2004) 100-12.
[45] R.H. Burdon, V. Gill, D. Alliangana, Hydrogen peroxide in relation to proliferation and apoptosis in BHK-21 hamster fibroblasts, Free Radic Res 24(2) (1996) 81-93.
[46] C.J. Bailey, M.B. Redpath, The esterolytic activity of epidermolytic toxins, Biochem J 284 (Pt 1)(Pt 1) (1992) 177-80.
[47] W. Kabsch, Xds, Acta Crystallogr D Biol Crystallogr 66(Pt 2) (2010) 125-32.
[48] R.B. Mariutti, T.A. Souza, A. Ullah, I.P. Caruso, F.R. de Moraes, L.M. Zanphorlin, N.R. Tartaglia, N. Seyffert, V.A. Azevedo, Y. Le Loir, M.T. Murakami, R.K. Arni, Crystal structure of *Staphylococcus aureus* exfoliative toxin D-like protein: Structural basis for the high specificity of exfoliative toxins, Biochem Biophys Res Commun 467(1) (2015) 171-7.
[49] V.B. Chen, W.B. Arendall, 3rd, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, Acta Crystallogr D Biol Crystallogr 66(Pt 1) (2010) 12-21.
[50] L. Schrödinger, W. DeLano, PyMOL, 2020.
[51] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A.J.P.o.t.N.A.o.S. McCammon, Electrostatics of nanosystems: application to microtubules and the ribosome, 98(18) (2001) 10037-10041.
[52] C.S. Hanes, Studies on plant amylases: The effect of starch concentration upon the velocity of hydrolysis by the amylase of germinated barley, Biochem J 26(5) (1932) 1406-21.
[53] E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state, J Mol Biol 372(3) (2007) 774-97.
[54] Y. Xu, S. Wang, Q. Hu, S. Gao, X. Ma, W. Zhang, Y. Shen, F. Chen, L. Lai, J. Pei, CavityPlus: a web server for protein cavity detection with pharmacophore modelling, allosteric site identification and covalent ligand binding ability prediction, Nucleic Acids Res 46(W1) (2018) W374-W379.
[55] K. Schoning-Stierand, K. Diedrich, R. Fahrrolfes, F. Flachsenberg, A. Meyder, E. Nittinger, R. Steinegger, M. Rarey, ProteinsPlus: interactive analysis of protein-ligand binding interfaces, Nucleic Acids Res 48(W1) (2020) W48-W53.
[56] D. Eisenberg, E. Schwarz, M. Komaromy, R. Wall, Analysis of membrane and surface protein sequences with the hydrophobic moment plot, J Mol Biol 179(1) (1984) 125-42.
[57] G. Ritter von Rittershain, Die exfoliativa dermatitis jungerer Saulinge, J Centralz Kinderhheilk 2 (1878) 3-23.
[58] M.E. Melish, L.A. Glasgow, The staphylococcal scalded-skin syndrome, N Engl J Med 282(20) (1970) 1114-9.
[59] M. Amagai, N. Matsuyoshi, Z.H. Wang, C. Andl, J.R. Stanley, Toxin in bullous impetigo and staphylococcal scalded-skin syndrome targets desmoglein 1, Nat Med 6(11) (2000) 1275-7.

CAPÍTULO III. *Staphylococcus aureus* Exfoliative Toxin E, oligomeric state and flip of P186: implications for its action mechanism

Staphylococcus aureus Exfoliative Toxin E, Oligomeric State and Flip of P186: Implications for Its Action Mechanism

Carolina Gismene,[1,†] Jorge Enrique Hernández González,[1,†] Angela Rocio Niño Santisteban,[1] Andrey Fabricio Ziem Nascimento,[2] Lucas dos Santos Cunha,[3] Fábio Rogério de Moraes,[1] Cristiano Luis Pinto de Oliveira,[4] Caio C. Oliveira,[3] Paola Jocelan Scarin Provazzi,[5] Pedro Geraldo Pascutti,[6] Raghuvir Krishnaswamy Arni,[1,*] and Ricardo Barros Mariutti[1,*]

[1]Multiuser Center for Biomolecular Innovation, IBILCE/UNESP, São José do Rio Preto 15054-000, Brazil

[2]Brazilian Synchrotron Light Laboratory (LNLS), Brazilian Center for Research in Energy and Materials (CNPEM), Campinas 13083-970, Brazil

[3]Institute of Chemistry, University of Campinas, Campinas 13083-970, Brazil

[4]Instituto de Física, University of São Paulo, São Paulo 05314-970, Brazil

[5]Laboratory of Genomic Studies, Sao Paulo State University-UNESP, São José do Rio Preto 15054-000, Brazil

[6]Laboratório de Modelagem e Dinâmica Molecular, Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Rio de Janeiro 21941-901, Brazil

[*]Correspondence: raghuvir.arni@unesp.br (R.K.A.); ricardomariutti@yahoo.com.br (R.B.M.)

[†]These authors contributed equally to this work.

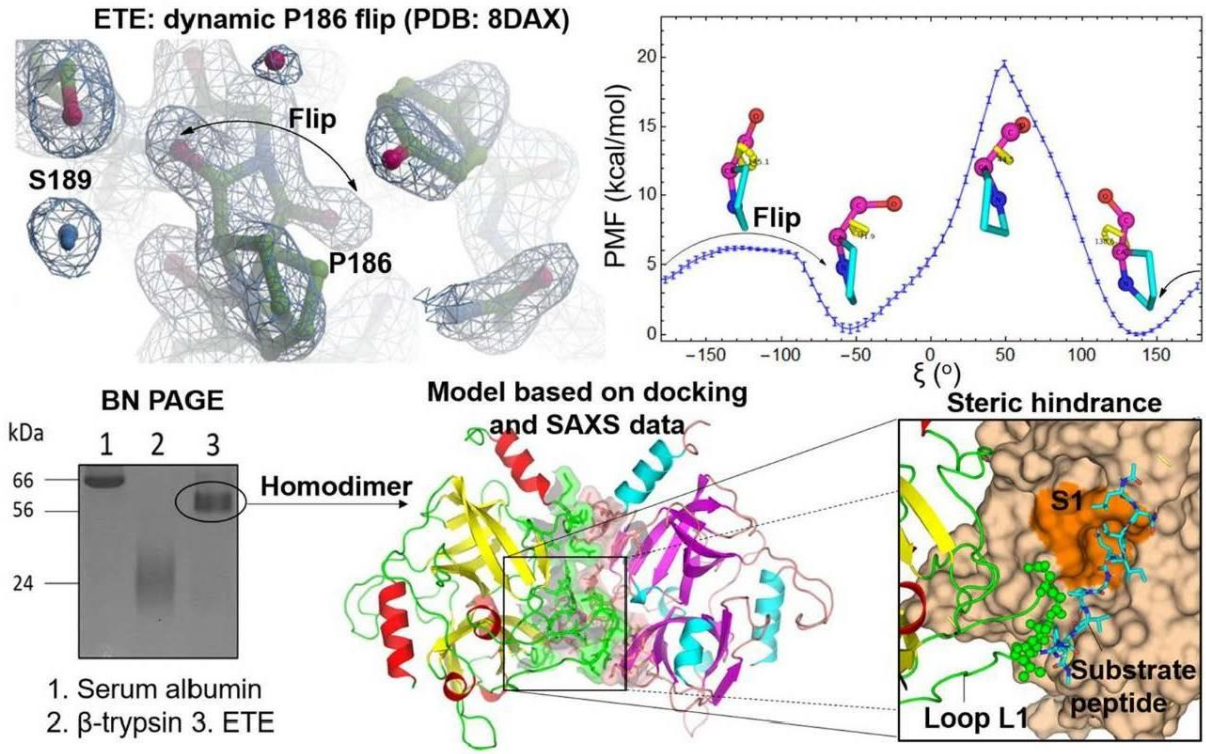
Abstract

Staphylococcal exfoliative toxins (ETs) are glutamyl endopeptidases that specifically cleave the Glu381-Gly382 bond in the ectodomains of desmoglein 1 (Dsg1) via complex action mechanisms. To date, four ETs have been identified in different *Staphylococcus aureus* strains and ETE is the most recently characterized. The unusual properties of ETs have been attributed to a unique structural feature, i.e., the 180° flip of the carbonyl oxygen (O) of the nonconserved residue 192/186 (ETA/ETE numbering), not conducive to the oxyanion hole formation. We report the crystal structure of ETE determined at 1.61 Å resolution, in which P186(O) adopts two conformations displaying a 180° rotation. This finding, together with free energy calculations, supports the existence of a dynamic transition between the conformations under the tested conditions. Moreover, enzymatic assays showed no significant differences in the esterolytic efficiency of ETE and ETE/P186G, a mutant predicted to possess a functional oxyanion hole, thus downplaying the influence of the flip on the activity. Finally, we observed the formation of ETE homodimers in solution and the predicted homodimeric structure revealed

the participation of a characteristic nonconserved loop in the interface and the partial occlusion of the protein active site, suggesting that monomerization is required for enzymatic activity.

Keywords: staphylococcal exfoliative toxin E; proline flip; homodimerization; X-ray diffraction; molecular dynamics.

Graphical Abstract



1. Introduction

Several strains of *Staphylococci* express exfoliative toxins (ETs), a group of proteins that function as virulence factors and facilitate host invasion [1,2]. These toxins cleave a desmosomal protein, desmoglein 1 (Dsg1), crucial for cell–cell adhesion in the epidermis, thus leading to blistering skin disorders [3,4,5]. Strains of *Staphylococcus aureus*, a major human pathogen, produce at least four ET variants, termed ETA [6,7], ETB [7], ETD [8] and ETE [9,10], the first three being relevant to human health [1]. ETA-producing *S. aureus* strains can cause Staphylococcal scalded skin syndrome (SSSS), a serious skin disease affecting large parts of the body and characterized by early symptoms such as lethargy, malaise, loss of appetite and fever, and followed by erythematous rash and blistering [1]. A more localized form of this affection, termed bullous impetigo, has been linked to strains expressing ETB [1]. Remarkably, it has been established using animal models that skin blistering can be induced by administering sterile filtrates containing the toxins and that the latter are solely responsible for the most

pronounced disease manifestations [1,11,12,13]. On the other hand, ETD was identified in *S. aureus* strains collected from furuncles and cutaneous abscesses, and hence is not believed to be related to SSSS or bullous impetigo [1,8]. ETE, the most recently characterized *S. aureus* ET, previously termed ETD-like, was discovered in *S. aureus* isolates from ewe mastitis [9,10]. Interestingly, it was found that ETE degrades Dsg1 expressed in ruminant teat canal epithelia, which suggests that it might provide advantages for mammary tract colonization [14]. ETs are directly responsible for serious diseases affecting livestock animals, which lead to important economic losses. In this sense, it has been found that ET-producing *S. hyicus* strains are associated with exudative epidermitis (EE) outbreaks in pigs [15]. The clinical manifestations include exfoliation, erythema and serous exudation and have been linked to several ETs, e.g., SHETA, SHETB, ExhA, ExhB, ExhC and ExhD [2]. Remarkably, in 2007, an EE outbreak in China was attributed to a highly pathogenic *S. sciuri* strain expressing ExhC [16]. In addition to skin exfoliation in newborn mice, this toxin is able to induce necrosis in multiple cell cultures [17], which suggests that its role in the pathogenesis of infections caused by ExhC-producing *Staphylococci* might not be limited to skin damage.

ETs are classified as glutamyl endopeptidases (GEPs), enzymes of the chymotrypsin-like serine protease (CLSPs) family that cleave peptide bonds following the α-carboxyl group of Glu residues [18]. They bear the CLSP typical catalytic triad (Asp-His-Ser) and their S1 pockets contain two structural elements present in all known GEPs that play a key role in the preference for Glu, i.e., His and Thr/Ser (Thr in ETs) at positions 213 and 190 (chymotrypsin numbering, hereinafter written in italic) [18]. On the other hand, nearly all known ETs possess a distinctive Lys residue at position *216* in lieu of the Ser residue encountered in many GEPs, except for ExhA of *S. hyicus*, which bears Arg at the equivalent position. It has been proposed that the amine group ($-NH_3^+$) of *K216* helps stabilize the carboxyl group of Glu at the P1 position, a role played by the $-NH_3^+$ of the N-terminal residue in other GEPs [18].

The available crystal structures of *S. aureus* ETs (PDBs: 1EXF, 1AGJ, 1DUE and 1DUA for ETA; 1DT2 and 1QTF for ETB; and 5C2Z for ETE) have revealed two unique structural features of these proteins: (i) the presence of an N-terminal α-helix containing several charged amino acids, and (ii) an unusual conformation of residues forming the oxyanion hole (Figure 1A,B) [18,19,20,21]. Concerning this last feature, it was first observed in ETA that the *P192-G193* peptide bond is flipped 180° relative to its typical position in other CLSPs [20]. Due to this unusual conformation, which will be termed I (I = inactive), the carbonyl oxygen of *P192* (*P192*(O)) points toward the -OH group of *S195* and forms a hydrogen bond (H-bond)

with it (Figure 1B), thereby hindering the access of substrates to the oxyanion hole [20]. Hence, it was proposed that the catalytic activation of ETA would require that *P192*(O) adopts the generally observed conformation, hereinafter termed A (A = active) [19,20]. On the basis of structural analyses, Vath et al. hypothesized that the *P192*(O) flip from conformation I to A could be promoted by an unknown activator that would induce a conformational change in the D-loop ($loop_{143\text{-}153}$, Figure 1A) by interacting with the N-terminal α-helix [20]. Likewise, conformation I has been observed for V183(O) and P186(O) of ETB (PDB: 1QTF) and ETE (PDB: 5C2Z), respectively (Figure 1B).

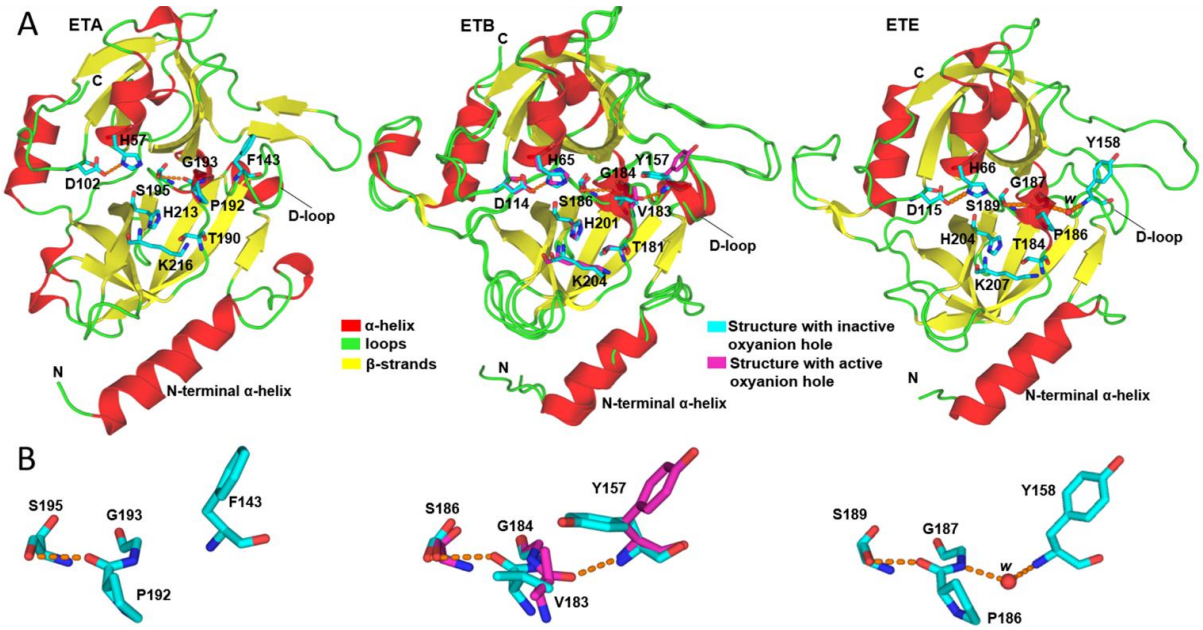


Figure 1. Structural representations of *S. aureus* ETs. (**A**) Cartoon and stick representations of the backbones and key active site residues of ETA (PDB: 1EXF), ETB (PDBs: 1DT2, magenta sticks, and 1QTF, cyan sticks) and ETE (PDB: 5C2Z), respectively. The main structural features of these proteins are depicted, i.e., the D102/114/115-H75/65/66-S195/186/189 catalytic triad, the N-terminal α-helix, the presence in the S1 pocket of residues T190/181/184, H213/201/204 and K216/204/207, which are crucial to the specificity for Glu-P1, atypical flipped conformation of P192/V183/P186(O) atom and the D-loop, which is believed to control the flip. (**B**) Stick representations of the alternate P192/V183/P186(O) conformations observed in the available crystal structures of ETA, ETB and ETE. In all cases, the orange dashed lines represent H-bonds. Residues of ETA are labeled according to the chymotrypsin numbering for ETA, as in PDB: 1EXF. Since the available PDBs of ETB and ETE do not follow the chymotrypsin numbering, residues of these toxins are labeled in sequential order. Numerical labels separated by / correspond to the positions of equivalent residues in ETA, ETB and ETE sequences, in that order. Of note, a water molecule, labeled *w*, forms a water bridge (red sphere) between G187(N) and Y158(N) in the crystal structure of ETE (PDB: 5C2Z). The actual residue numbers in the structure 5C2Z (*n*) took into account 30 residues of the signal peptide and are different from those used in this work (*n*-30).

Interestingly, in the structure of ETB (PDB: 1DT2), V183(O) occurs in conformation A (Figure 1B) [21], which suggests that an activator is not essential to promote the flip. It has also been proposed that the equilibrium between conformations A and I of V183(O) may depend on the different crystal forms [21]. Furthermore, the crystal structure of EXI, another ET from *S. pseudintermedius* that hydrolyses canine Dsg1 [22], has been deposited in the Protein Data Bank (PDB: 6E0U) and analysis of the two protein chains reveals that the Pro residue at S1 (P213) also appears in conformation A. These findings support the idea that the oxyanion hole can be properly formed in ETs. However, both the structures (PDBs: 6E0U and 1DT2) have been determined at relatively low resolutions (2.75 and 2.80 Å, respectively), which undermines the previous conclusion.

ETs do not exhibit in vitro proteolytic activity against broad-spectrum substrates of CLSPs, a fact that has been frequently attributed to the unusual conformation of their oxyanion holes [19,20]. Nonetheless, they have been found to display moderate esterolytic and peptidase activities against some small molecules, such as Boc-L-Glu-↓OPh [23] and the peptides α- and β-melanocyte-stimulating hormones (α-MSH and β-MSH), respectively [24]. Therefore, the interaction with Dsg1 is not an absolute requirement for catalytic activity. These lines of evidence, along with the possible existence of alternate conformations of residue *192* observed in the available crystal structures, suggest that the relatively high specificity of ETs, restricted to a few known molecules, may arise from molecular interactions that are independent of the oxyanion hole conformation.

Our experimental and computational study of ETE provides insights into structural aspects of this toxin that are relevant to the understanding of its action mechanism. Moreover, a high-resolution crystal structure of ETE revealed that P186 exists in two alternate conformations, A and I, thus indicating the occurrence of the P186 flip. Protein–protein docking and molecular dynamics (MD) simulations were combined with small-angle X-ray scattering (SAXS) to predict the conformation of ETE homodimers in solution. The most favorable conformation involves the participation of a nonconserved loop, characteristic of the ETs, and shows that the active site can be partially occluded by the neighboring protein chain.

2. Results

2.1. Refinement of the ETE Crystal Structure

Recombinant ETE was expressed and purified as described in Materials and Methods. Crystals were obtained with the precipitant sodium formate 4.0 M, pH 6.0 and used in X-ray diffraction experiments. The previously determined ETE structure (PDB: 5C2Z) was crystallized in the space group P 2(1) with cell dimensions a = 49.41, b = 93.14 and c = 50.48 Å and diffracted to 1.96 Å [25]. The current ETE crystal belongs to the space group P 4(3) 2(1) 2 with unit cell dimensions a = 97.53, b = 97.53 and c = 116.35 Å and was refined at 1.61 Å to crystallographic residuals of 0.21/0.24 (Rwork/Rfree) (Table 1).

Table 1. Crystallographic data processing and refinement statistics of ETE PDB structure 8DAX.

Data Collection	**ETE**
Wavelength (Å)	1.459
Resolution range (Å)	43.71–1.61 (1.67–1.61)
Space symmetry group	P 4(3) 2(1) 2
Cell dimensions:	
a, b, c (Å)	97.74, 97.74, 116.65
α, β, γ (º)	90.00, 90.00, 90.00
Total reflections	1,847,482 (175,190)
Unique reflections	72,995 (7181)
Multiplicity	25.3 (24.4)
Completeness range (%)	99.94 (99.85)
Mean I/σI	24.66 (1.07)
R-merge	0.076 (2.829)
Rmeas	0.077 (2.889)
Correlation coefficient ½	1 (0.52)
$R_{(work)}/R_{(free)}$	0.2206/0.2523
Number of protein chains in AU	2
Solvent water molecules	145
R.M.S:	
Bond lengths (Å)	0.015
Bond angles (º)	2.00
Ramachandran analysis:	
Favored (%)	95.33%
Allowed (%)	4.27%
Disallowed (%)	0.41%
Average B-factor ($Å^2$)	36.50
PDB accession code	8DAX

Coincidentally, both crystal structures of ETE (PDBs: 5C2Z and 8DAX) contain two protein chains in the asymmetric unit (AU); however, the observed protein–protein interactions (PPIs) are different (Figure S1). To assess whether any of the crystallographic homodimers correspond

to stable assemblies in solution, the complexes were submitted to the PISA and PRODIGY-CRYSTAL web servers [26,27]. In addition, both putative homodimers were subjected to molecular mechanics generalized Born surface area (MM-GBSA) and umbrella sampling (US) free energy calculations, and the results are summarized in Table S1. As can be observed, none of the homodimers present in the AUs are classified as being stable in solution. According to PRODIGY-CRYSTAL, the PPIs of 5C2Z and 8DAX have a 100 and 65.6% chance, respectively, of being merely the result of crystallographic contacts. Therefore, out of the two structures, the one reported here is more likely to correspond to a stable homodimer, although still classified as crystallographic. Likewise, it ranked better than 5C2Z based on the MM-GBSA effective free energy (ΔG_{eff}) values determined from 100 ns MD simulations of such systems ($\Delta G_{eff,100\ ns}$). Nonetheless, the free energy estimated from the potential of mean force (PMF) of the new crystal structure, calculated through US, is nearly zero, which underscores that this complex is unstable in solution (Table S1).

2.2. Oxyanion Hole and P186 Flip

The electron density for P186(O) in chain A of the new ETE structure (PDB: 8DAX) shows that this atom can coexist in two alternate conformations A and I (Figure 2A), with occupancies of 0.3 and 0.7, respectively, thus indicating that the latter is relatively more stable. It is worth noting that this crystal structure demonstrates for the first time the occurrence of a dynamic flip between conformations A and I of residue *192* of an ET. Moreover, the respective occupancies are consistent with the fact that conformation I has been more frequently observed in the available crystal structures of ETs than the alternate one. On the other hand, the flip was not detected in the second chain (chain B) of ETE present in the AU (Figure 2A). In this case, P186(O) is pointing toward the oxyanion hole, forming an H-bond with S189(OG) (Figure 2B). Moreover, as in PDB: 5C2Z (Figure 1B), a crystallographic water molecule forms a water bridge between G187(N) and Y158(N) in chain B of PDB: 8DAX. This water molecule is absent when P186 adopts conformation A (compare Figure 2A,B).

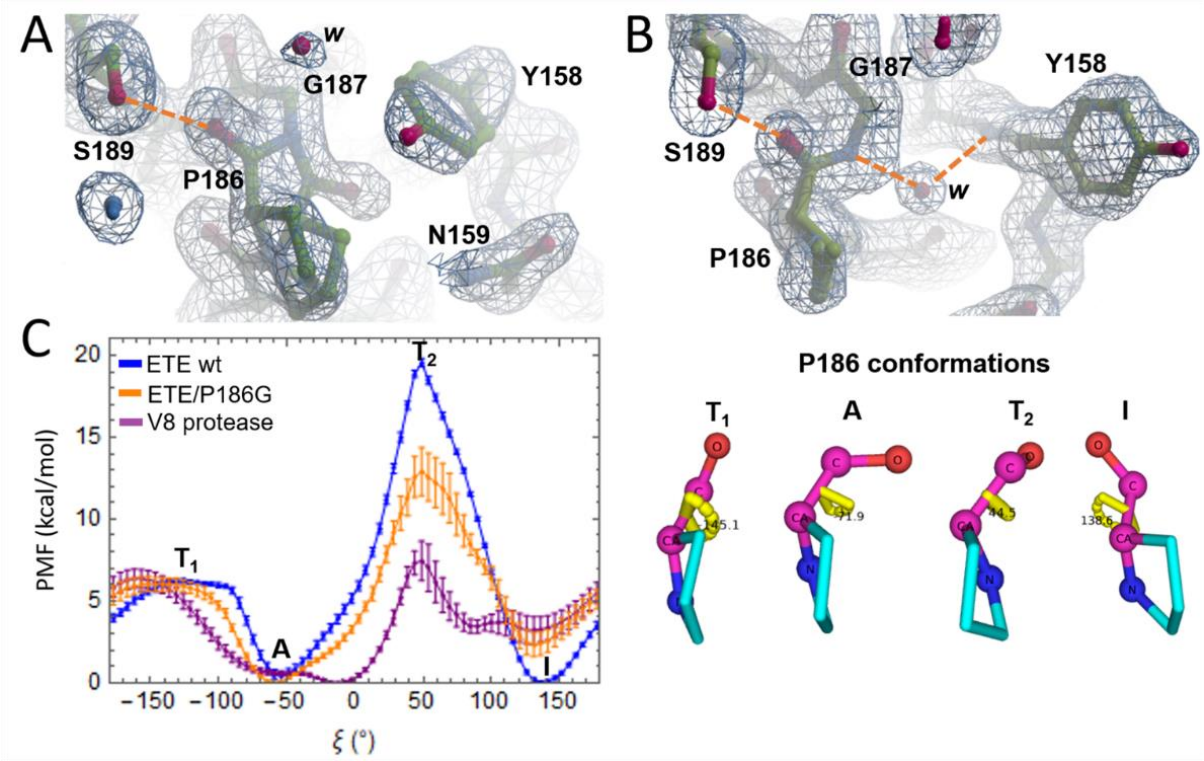


Figure 2. Crystallographic and energetic evidence of P186 dynamic flip in ETE. (**A**) Electron density map of the three-dimensional (3D) structure of ETE (PDB: 8DAX, chain A) showing the occurrence of a 180° flip of P186(O). (**B**) Electron density map of the 3D structure of ETE (PDB: 8DAX, chain B) showing conformation I of P186(O). Residues are labeled and represented as sticks. The electron density is represented as a mesh. H-bonds are depicted as orange dashed lines. Crystallographic water molecules are identified with the letter *w*. (**C**) PMFs for the 360° rotation around the N-CA-C-O dihedrals (ξ) of P186 of ETE wt, G186 of ETE/P186G mutant and G166 of V8 protease (see color legend). The two minima correspond to conformations A and I and the two maxima, T_1 and T_2, to the associated energy barriers (see the corresponding P186 conformations on the right panel). Equivalent conformations of G186 and G166 are not shown for brevity's sake.

It is worth noting that the Y158 side chain occurs in different conformations in each chain of the crystal structure, which seem to be linked to the P186(O) orientation, and will be referred to as conformations 1 (Figure 2A) and 2 (Figure 2B). To assess whether there is an actual correlation in the motion of both residues, we performed five replicate 1 μs MD simulations of ETE in water. Three replicates were started from an ETE structure bearing P186(O) in conformation I and two from a structure with P186 adopting conformation A. In all cases, Y158 was initially oriented as in Figure 2A (conformation 1). First, we observed that almost no transitions between conformations A and I of P186 occurred along the trajectories, thus indicating that crossing the energy barrier separating both states is a rare event when using conventional MD simulations in the microsecond time scale (Figure S2A). Conversely, Y158 sampled different conformations and many transitions during the MD simulations (Figure S2B).

Remarkably, the three sampled conformations of Y158 side chain, which include 1 and 2, were observed regardless of the orientation of P186 (Figure S2C,D), which in turn suggests that the motion of these two residues is uncorrelated. This result is further supported by the fact that the Y157 side chain in the crystal structure of ETB (PDB: 1DT2) adopts a conformation different from that of ETE chain B, despite V183(O) and P186(O) being similarly oriented in both structures (compare Figure 1B and Figure 2B).

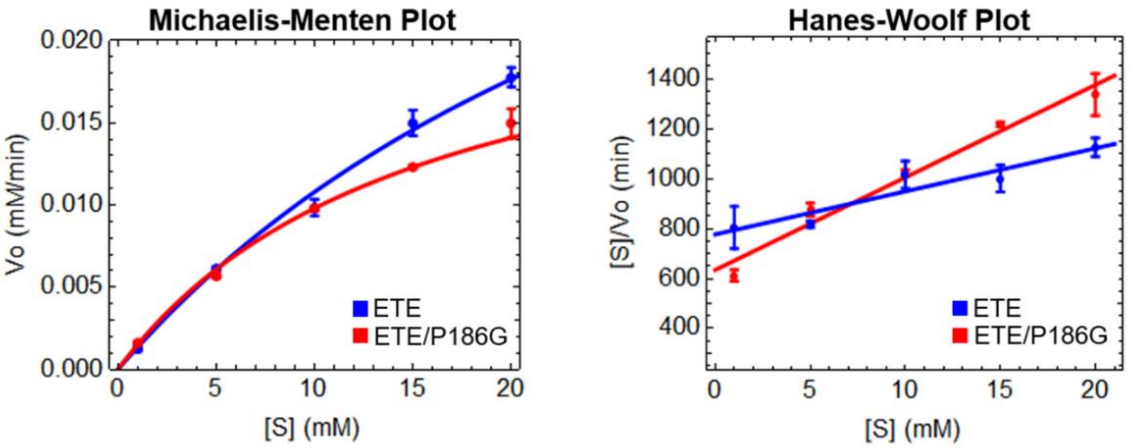
The inability of conventional MD simulations in the microsecond time scale to properly sample transitions between conformations A and I of P186 indicated the need for enhanced sampling techniques to accurately estimate the free energy associated with that conformational change. Therefore, we conducted US simulations, which allowed us to calculate the PMF corresponding to the 360° rotation of P186 N-CA-C-O dihedral (ξ) and, from this, the relative stability of conformations A and I, as well as the associated energy barriers. Prior to these calculations, we assessed the accuracy of the US free energy protocol by obtaining the PMF for the rotation of the dipeptide L-alanyl-L-proline peptide bond (Figure S3). The predicted free energy associated with the interconversion of *cis* and *trans* conformers is in excellent agreement with the experimental value obtained from the reported equilibrium constant of this process [28]. The major problem with the predictions is the height of the energy barrier, i.e., the free energy of the transition state, which was underestimated by ~6 to 8 kcal/mol. Therefore, when applying the previous protocol in prospective studies, one can expect greater accuracy in the estimates for the free energy difference between conformations corresponding to PMF minima than for the associated transition free energy barriers.

The average PMF for ETE wt shown in Figure 2C (blue PMF profile) displays two energy minima at $\xi = -54°$ and $\xi = 136°$, corresponding to conformations A and I, respectively, with the latter being slightly more stable than the former by (0.4 ± 0.3) kcal/mol. According to the previous result, these conformations would occur with probabilities of 23 to 45% and 55 to 77%, respectively, in agreement with the two P186(O) occupancies observed in the crystal structure of ETE (Figure 2A). The PMF also displays two energy maxima, T_1 and T_2, which define the energy barriers to transition from conformation A to I and vice versa. According to the heights of the energy barriers (~6 and ~19 kcal/mol), the interconversion between both conformations is more likely to occur through T_1 (energy barrier of 6 kcal/mol), which corresponds to the P186(O) atom pointing in the same direction as the P186 side chain (Figure 2C).

In addition, we calculated the PMF for V8 protease (PDBs: 2O8L), which belongs to a second type of GEP produced by *S. aureus* [18]. This enzyme bears a properly formed oxyanion hole, lacks the N-terminal α-helix and is far more active against the substrate Boc-L-Glu-↓OPh than ETs [19,29]. In agreement with the aforementioned structural features, our calculations show that conformation A of G166(O) (*G192*) is significantly more stable than conformation I (3.2 ± 0.8 kcal/mol) in V8 protease (Figure 2C, purple profile). Interestingly, we also realized, by inspecting previously reported multiple sequence alignments (MSAs), that Gly was always found at position *192* in other *staphylococcal* V8-like GEPs [30,31]. Hence, we generated a model of ETE/P186G mutant through in silico mutagenesis using Pymol [32] and calculated the PMF for the rotation of G186 dihedral (Figure 2C, orange profile). As can be observed, this PMF resembles, to a larger extent, that of V8 protease, with conformation A of G186 being more stable than conformation I by 2.4 ± 0.7 kcal/mol (Figure 2C). This suggests in turn that, if the enzymatic activity of ETE is regulated by the P186 flip, one must expect a significant increase in activity for ETE/P186G. This hypothesis will be evaluated in the next section.

2.3. Esterolytic Activity of ETE Wild-Type and ETE/P186G Mutant

The esterolytic activity of ETE and ETE/P186G against the synthetic substrate Boc-L-Glu↓-OPh was determined as described in Materials and Methods. In both cases, the initial velocities of the enzymatic reaction at different substrate concentrations follow the characteristic Michaelis–Menten hyperbola (Figure 3). However, due to the relatively high concentrations of Boc-L-Glu-↓OPh used in the assays and the solubility limit of this molecule in 1-4 dioxane, substrate concentrations greater than 20 mM could not be tested and, consequently, enzyme saturation was not reached (Figure 3). In light of these issues, we employed Hanes–Woolf plots [33] to estimate the kinetic parameters of the analyzed enzymes through linear regression (Figure 3).



Kinetic parameters

Enzyme	K_M (mM)	V_{max} (mM·min^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M ($mM^{-1}·s^{-1}$)
ETE	36.0±7.0	0.049±0.007	23.8±3.4	0.67±0.03
ETE/P186G	14.7±1.6	0.025±0.002	12.1±1.0	0.79±0.05

Figure 3. Determination of ETE and ETE/P186G kinetic parameters. The upper panels show the Michaelis–Menten and Hanes–Woolf plots for both enzymes. The table in the lower panel contains the values for the kinetic parameters estimated from the Hanes–Woolf plots through linear regression. V_o, [S], K_M, V_{max} and k_{cat} stand for initial velocity, substrate concentration, Michaelis constant, maximal velocity and catalytic constant, respectively. The k_{cat}/K_M ratio quantifies the catalytic efficiency of each enzyme. Mean values ± standard errors of the mean are shown in all cases.

The estimated kinetic parameters show that the P186G mutation caused a two-fold increase in the association constant of the enzyme and the substrate ($1/K_M$) and a two-fold decrease in the maximal velocity (V_{max}). Due to these opposing effects, the catalytic efficiencies (k_{cat}/K_M) of ETE and ETE/P186G are very similar, with that of the latter enzyme being slightly higher (Figure 3). Overall, the previous results demonstrate the negligible impact of the P186G mutation on the catalytic activity of ETE, regardless of the expected stabilization of the functional oxyanion hole conformation.

2.4. Oligomeric State of ETE in Solution

Initial evidence that ETE is a homodimer in solution was obtained based on size-exclusion chromatography (SEC) profiles. Therefore, we decided to further evaluate the oligomeric state of ETE in several solutions comprising different buffers, NaCl concentrations and pH values. As can be observed in Figure 4A, all the chromatograms showed the ETE elution peak at a volume of ~11 mL, thus indicating that the protein occurs in the same oligomeric state in all

the tested conditions. Moreover, we obtained the chromatograms for serum albumin (BSA) (MW ~66 kDa) and β-trypsin from bovine pancreas (MW ~24 kDa) and used them as controls to check whether the ETE elution peak corresponds to a dimeric (MW ~60 kDa) or a monomeric protein (MW ~30 kDa). It is evident from the overlap of the different chromatograms that the ETE elution profile is closer to that of BSA (Figure 4A), which confirms that the toxin forms a homodimer in solution.

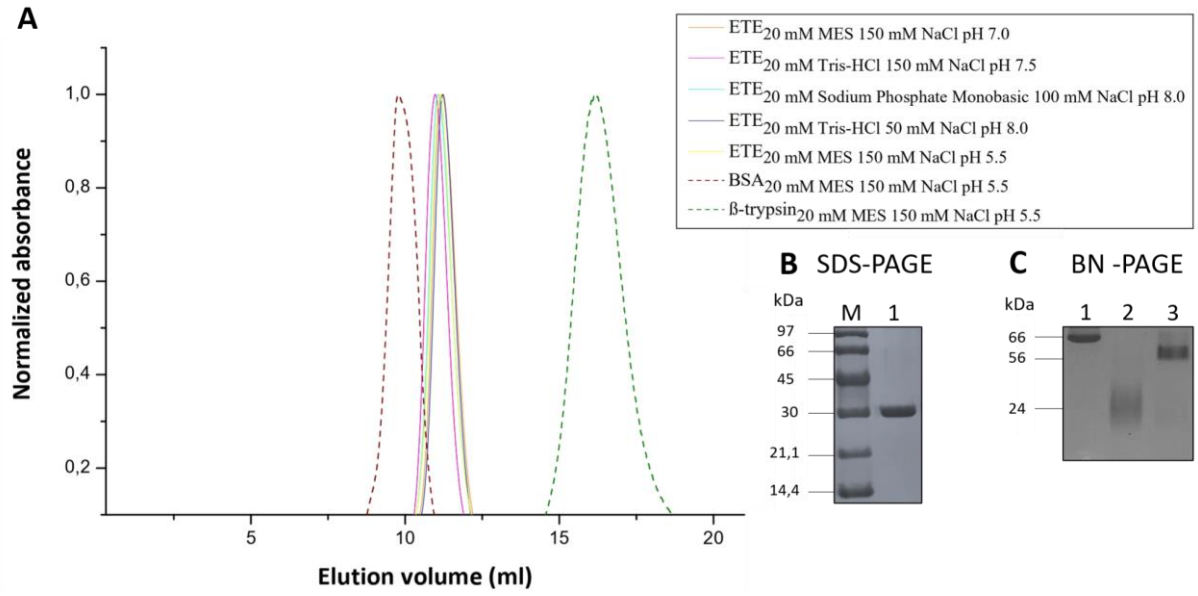


Figure 4. ETE is a dimer in different solution conditions. (**A**) SEC profiles for ETE under different buffering conditions (solid line). BSA and β-trypsin were employed as controls (dash line). Absorbance was measured at 230 nm in all cases, and normalized values are plotted. The overlap of peaks at ~12 mL indicates that ETE elutes in the same oligomeric state in all the tested conditions. (**B**) SDS-PAGE of ETE (1) using the protein marker LMW-SDS Marker KIT—GE Healthcare (M). (**C**) BN-PAGE of ETE (3) using BSA (1) and β-trypsin from bovine pancreas (2) as molecular weight markers. A running buffer containing 25 mM Tris pH 8.3 and 192 mM Glycine was employed for BN-PAGE.

The previous results were complemented by performing SDS-PAGE and BN-PAGE electrophoresis. As expected, the first method showed that, under denaturing conditions, ETE migrates as a monomer, producing a band at ~30 kDa (Figure 4B). On the other hand, BN-PAGE experiments (Figure 4C) showed the ETE band close to that of BSA monomer, whereas no ETE band corresponding to the molecular weight of β-trypsin was observed. The previous findings are in agreement with the SEC experiments and confirm the formation of ETE homodimers. Similar results are exhibited by ExhC from *S. sciuri*, already described as a homodimer in solution [34], which was used as a control in this study (Text S1 and Figure S4).

We also found that ETE/P186G forms homodimers in solution (Figure S5), thus indicating that this mutation does not affect the protein oligomeric state.

The oligomeric state of ETE in solution was further studied using SAXS. Figure 5A shows the intensity curve for the toxin in 20 mM MES pH 7.0 and 150 mM NaCl buffer (similar to physiological conditions). Assuming that the protein sample was monodispersed, the indirect Fourier transform (IFT) operation was performed, which provided the distribution of pairs of distances (ρ(r)) (Figure 5B) within the investigated particle [35]. The IFT analysis indicated that ETE has a radius of gyration of (29.8 ± 0.1) Å and a maximum size of ~97 Å. Furthermore, the shape of the ρ(r) distribution obtained is typical for dimeric structures, due to the presence of the shoulder around 80 Å (Figure 5B). These results, in addition to reinforcing the SEC and BN-PAGE data (Figure 4), were also used to support the prediction of the dimeric conformation of ETE in solution.

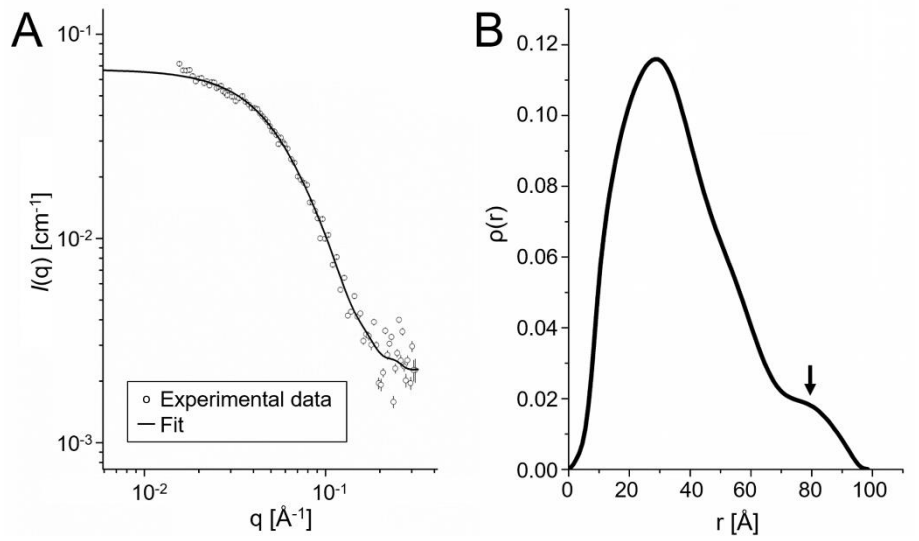


Figure 5. SAXS data for ETE in the buffer solution MES 20 mM, pH 7.0 and 150 mM NaCl. (A) Intensity SAXS profile. (B) Pair distance distribution function (ρ(r)). *I* and q stand for the intensity and the scattering vector, respectively. The arrow indicates the shoulder in the ρ(r) distribution, which is interpreted as a signature of dimerization.

2.5. Prediction of the ETE Homodimer Structure in Solution

As discussed earlier, the contacts formed by the interacting chains in the AUs of the two available crystal structures of ETE were predicted to be merely crystallographic. Therefore, to find the most stable conformation of ETE homodimers in solution we employed the workflow shown in Figure S6, which combined several protein–protein docking algorithms, free energy calculations and SAXS-profile fits (Figure S7). In parallel, we assessed the poor quality of an

ETE homodimer model generated by AlphaFold-multimer (Figure S8). The applied computational approaches ultimately led to the identification of a single pose (hereinafter referred to as ETE_2-CP-0, where CP-0 stands for ClusPro [36] pose 0) as the most stable conformation of ETE homodimers in solution (Table S2). Moreover, the proposed conformation passed subsequent tests aiming to assess the stability of the complex on the basis of its predicted PPI (Table S3). Furthermore, SPPIDER [37] predicted interface-forming residues (IFRs) at the complex interface (Table S3). Altogether, these results strongly suggest that ETE_2-CP-0 is the most likely conformation of ETE homodimers in solution.

The ETE_2-CP-0 conformation remained relatively stable along five replicate 1 μs MD simulations, as can be inferred from the backbone RMSD profiles shown in Figure 6A. Moreover, the mean ΔG_{eff} value for the homodimer, calculated by averaging the results from the replicate MD simulations, is −57 ± 2 kcal/mol. This value is consistent with that initially estimated from a single 1 μs MD simulation (−56.2 kcal/mol, Table S2), which reinforces the stability of the homodimer during all the independent MD simulations. Hotspot residues at homodimer PPI, i.e., those largely contributing to the binding process, were also identified by performing MM-GBSA per-residue free energy decomposition (Figure 6B). Two of these residues, R19 and E185, can form intermolecular H-bonds and salt bridges, which explains their favorable energy contribution. On the other hand, L91, Y158 and N159 possess large van der Waals free energy contribution, thus indicating the importance of hydrophobic interactions at the predicted PPI. Interestingly, L91 was identified as the main hotspot of ETE_2-CP-0, in agreement with previous results pointing out that it is an IFR (Figure 6B and Table S3).

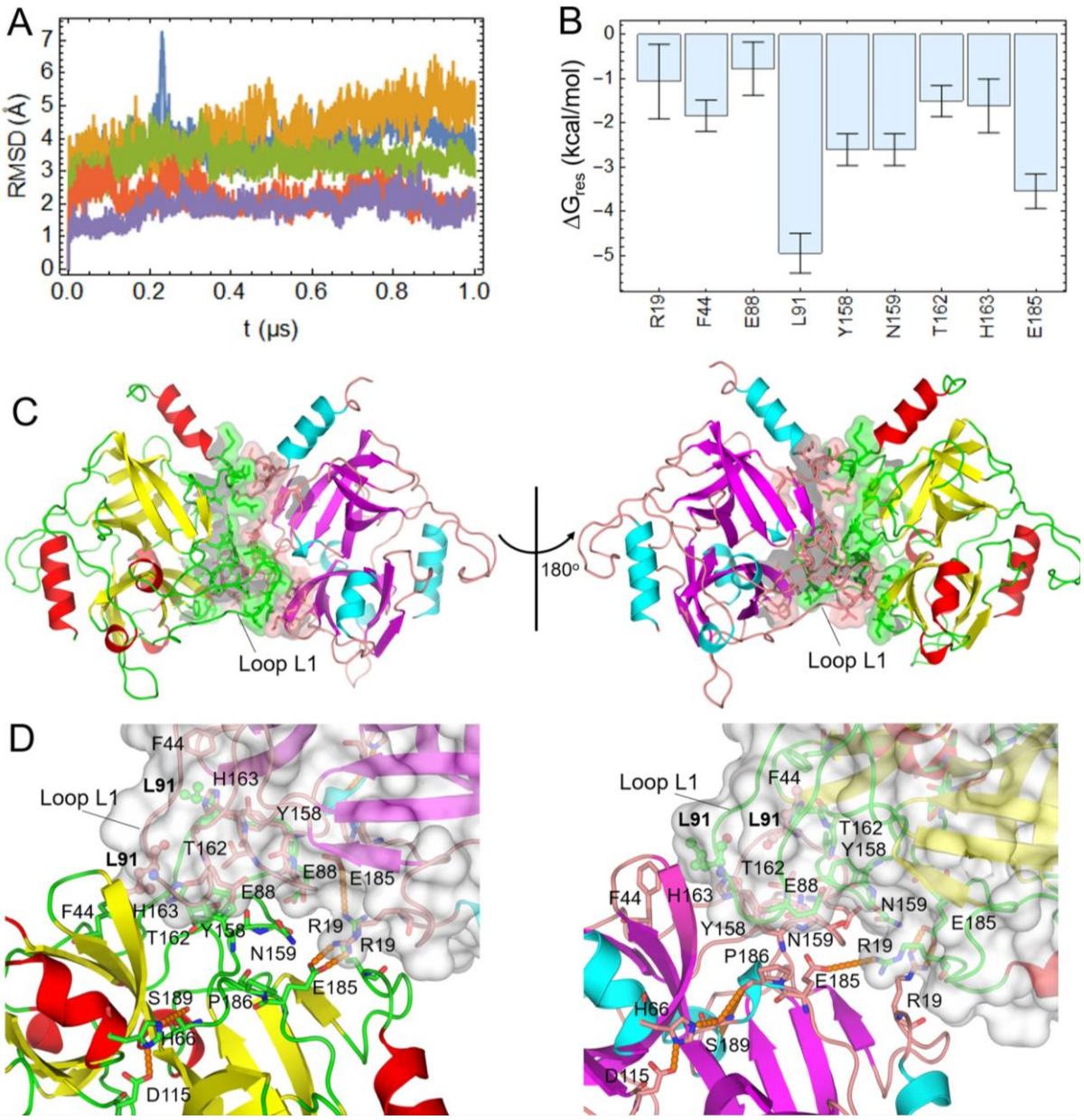


Figure 6. MD-based and structural analyses of the predicted structure of ETE homodimers (ETE_2-CP-0). (A) Backbone root mean squared deviation (RMSD) profiles for the five replicate MD simulations of the ETE homodimer. The initial structure (t = 0) was taken as reference for the RMSD calculations. (B) Per-residue free energies for the main hotspots of the homodimer interface calculated using MM-GBSA. (C) Two views of ETE homodimer upon 180° rotation around the vertical axis. The protein structure is represented as cartoons, with secondary structure elements being colored differently, i.e., red/cyan for α-helices, yellow/magenta for β-strands and green/salmon for loops in chains A/B. Interface residues are shown in surface and stick representations. (D) Two views of the complex interface displaying interface hotspot and active site residues as sticks. The main hotspot L91 of the homodimer interface is highlighted in sticks and spheres. H-bonds are represented as orange dashed lines.

Figure 6C shows a structural representation of the homodimer conformation ETE_2-CP-0. Like the vast majority of protein homodimers in nature, the proposed structure displays C_2 symmetry

[38], i.e., it is symmetric upon 180° rotation with respect to the indicated vertical axis (Figure 6C). It is worth noting that the proposed model significantly diverges from the crystallographic homodimers present in the AU of PDB structures: 8DAX and 5C2Z, as confirmed by the different composition of interface residues and heavy atom RMSD values (Table S4). Furthermore, the predicted PPI involves residues like Y158 and E185, which are close to P186 and were identified as hotspots of the homodimer interface (Figure 6B,D). An interesting structural feature of the ETE_2-CP-0 structure is the involvement in the PPI formation of a protruding loop, termed L1 (Figure 6C,D). Of note, this loop contains L91, the main interface hotspot residue (Figure 6B,D). To study the conservation of loop L1 among CLSPs, a structural alignment of different ETs and other proteins of the S1 family was carried out (Figures S9 and S10). It became clear that loop L1 corresponds to an insertion of several amino acids that were only observed in ETs. The equivalent regions in the other CLSPs form significantly shorter loops. A second insertion corresponding to another loop, termed L2, was also found in ETE, ETB and EXI, but it is not part of the predicted homodimer PPI. These results suggest that loop L1 might be a distinctive structural feature of ETs that allows the formation of homodimers.

Finally, we used template-based modeling, in silico mutagenesis and MD simulations to predict the structure of monomeric ETE in complex with a peptide matching the sequence of Dsg1 EC3-EC4 linker (Figure S11). Then, by superimposing this complex onto ETE_2-CP-0, we realized that loop L1 encroaches into the Sn' side of the neighboring chain's active site, thus sterically hindering the full accommodation of the substrate's peptide. Therefore, it is likely that ETE must undergo monomerization in order to become activated.

3. Discussion

ETs possess several unique features among CLSPs, i.e., a relatively high specificity for Dsg1, a densely charged N-terminal α-helix and a nonfunctional oxyanion hole due to a flipped conformation of residue *192* in most of the available crystal structures [18,19,20,21]. To date, the activation mechanism of these toxins remains largely unknown. Moreover, the direct role of the unusual conformation of residue *192* in the activation process cannot be inferred based on the controversial findings reported in the literature, e.g., the existence of crystal structures of ETs with functional oxyanion holes (PDBs: 1DT2 and 6E0U) [21]. Here, we provide evidence that sheds light on some of these issues by establishing connections between the oligomeric state and the oxyanion hole conformations with the action mechanism of ETE [14,25].

We found by inspecting a newly reported crystal structure of ETE (PDB: 8DAX) that the P186 residue of chain A occurs in two alternate conformations, I and A, corresponding to P186(O) oriented toward the catalytic Ser or rotated 180°. Furthermore, crystallographic evidence that P186(O) can coexist in alternate positions was supported by free energy calculations, which demonstrated that conformation I is marginally more stable than conformation A (~0.4 kcal/mol) and that the associated energy barrier separating them is only 6 kcal/mol high. For comparison purposes, we calculated the free energy barrier associated with the *cis-trans* isomerization of the peptide bond of L-alanyl-L-proline in bulk water [28], which turned out to be two times higher. Overall, our results show that the P186 flip is a dynamic event that can occur, in principle, without being promoted by an effector molecule and hence is unlikely to be the main structural feature controlling the ETE enzymatic activity.

More evidence in this regard was obtained from the analysis of the ETE/P186G mutant. Based on the presence of Gly at position *192* of staphylococcal V8-like GEPs [30,31], which possess functional oxyanion holes and higher esterolytic activities than ETs [18,29], we assessed whether this ETE variant would acquire V8-like properties. The free energy calculations performed in our work confirmed that the P186G mutation significantly stabilizes conformation A (2.4 ± 0.7 kcal/mol more stable than conformation I). Moreover, the esterolytic assays showed that it causes a two-fold increase in the association constant of the enzyme and the substrate. This modest affinity increase is likely the result of having the entire population of ETE/P186G with a preformed oxyanion hole. Hence, there is no energetic penalty associated with shifting the equilibrium from conformation I to A during substrate binding, as occurs for the wild-type enzyme. However, it was also found that the catalytic constant of the mutant is almost half that of ETE. Consequently, the catalytic efficiencies against Boc-L-Glu-↓OPh of ETE and ETE/P186G are approximately equal, thus reinforcing the slight impact of residue 186 flip on the enzymatic activity.

Furthermore, it was determined that ETE and ETE/P186G form homodimers in aqueous solution, a conclusion that was derived by using different experimental approaches, i.e., SEC, BN-PAGE and SAXS. This was not a surprising finding per se, since ExhC, an ET from *S. sciuri*, had already been reported as a homodimer in solution using some of the previous approaches [34]. Nonetheless, the fact that at least two different toxins form homodimers suggests that this might be a general property of ETs, probably playing some functional role.

Interestingly, although the two available crystal structures of ETE possess two contacting protein chains in the AU, several computational analyses performed here revealed that none of

them corresponds to a stable biological assembly. Therefore, we combined different computational approaches to predict the most favorable conformation of ETE homodimers. Of note, the proposed structure displays partially occluded active sites at the Sn' region, a feature that suggests that homodimers may show reduced or no catalytic activity and that monomerization would be a necessary step for ETE to become fully activated. This last event could be favored, for example, by interacting with Dsg1. The esterolytic of ETE against Boc-Glu↓-OPh poses additional questions on the ability of homodimers to directly cleave small molecules that do not bind extensively to the occluded Sn' side of the active site. Furthermore, the crystallization of ETE in assemblies different from the one found to be stable in solution and the relatively low affinity estimated by PISA on the basis of the predicted complex structure ($\Delta G_{bind} = -3.1$ kcal/mol) indicate the formation of a labile homodimer that could readily dissociate under certain conditions. This evidence reinforces that monomerization is a plausible step that could initiate the activation of this enzyme.

The partial occlusion of the ETE active site in the predicted homodimer is caused by loop L1. Moreover, the main hotspot of the PPI, L91, lies in the aforementioned loop, thus suggesting that the latter is essential for the homodimer formation. This hypothesis can be explored using site-directed mutagenesis and deletion of loop residues, which will help assess the influence of the loop on the dimerization of ETE. Of note, our results show that the protruding loop L1 seems to be a characteristic structural element of several ETs, not found in other CLSPs with available crystal structures. Therefore, it is likely to play a key role in the homodimerization of other ETs. Future studies must be carried out to evaluate the oligomeric state of all known ETs and the involvement of their loops L1 in the formation of the PPIs.

Hanakawa et al. proposed the existence of an exosite in the ETs responsible for their specific interaction with Dsg1 EC3 and subsequent activation [4]. Our results pose new questions about the interplay between homodimerization of ETs and the recognition of EC3 via exosites and whether the latter overlap, to some extent, with the homodimer interface. Should that be the case, monomerization would be necessary not only to expose the active site but to enable the interaction of ETs with their natural substrate. Further experiments are required to establish the oligomeric state of catalytically inactive ETs (e.g., ETE/S189A) in the presence of Dsg1 and to locate their exosite.

4. Materials and Methods

4.1. Protein Expression and Purification

Escherichia coli BL21(DE3)-T1R competent cells were transformed with pD441-ETE and pET28a(+)-ETE/P186G (GenScript) expression vectors. The selected bacterial colonies were grown overnight at 37 °C in Lysogeny Broth (LB) medium supplemented with kanamycin (50 mg/mL). Subsequently, the cultures were diluted 100-fold with a fresh LB medium containing kanamycin (50 mg/mL) and incubated at 37 °C. When the optical density measured at 600 nm (OD_{600}) reached 0.5, expression was induced with 0.4 mM IPTG for 5 h at 30 °C. The cells were collected by centrifugation at 2600× *g* for 10 min at 4 °C and homogenized by sonication in lysis buffer containing 5 mM $NaHPO_4$, pH 7.7, 400 mM NaCl, 10 mM imidazole and 10% (*v*/*v*) glycerol. The lysed cells were centrifuged at 15,000× *g* for 30 min at 4 °C, and both supernatants with a large amount of ETE and ETE/P186G were subjected to affinity chromatography using a Ni2+-Sepharose column (Bio-Rad), according to the manufacturer's instructions. The protein concentration was determined spectrophotometrically at approximately 10 mg/mL (Biomate 3S—Thermo scientific spectrophotometer), and additional purification steps were performed for ETE and ETE/P186G in 20 mM MES pH 7.0 and 150 mM NaCl buffer (and other buffering conditions for ETE) on an AKTA system purifier by SEC using a Superdex G75 10/300 GL column (GE—Healthcare Life Sciences) at a flow rate of 0.7 mL/min. The peak fractions were collected and analyzed by Western blotting, SDS-PAGE and BN-PAGE 15%.

4.2. SDS-PAGE and BN-PAGE 15%

The analysis of 30 µg of proteins by BN-PAGE and SDS-PAGE 15% was performed according to standard protocols [39]. BN-PAGE measured the difference in electrophoretic migration and size between ETE, β-trypsin from bovine pancreas (Sigma-Aldrich) and bovine serum albumin monomer (Sigma-Aldrich), while a molecular marker Amersham ECL Rainbow (GE-Healthcare) was used in SDS-PAGE. The samples were fixed and stained in a mixture containing methanol, water, acetic acid (50:50:10 *v*/*v*) and Coomassie Brilliant Blue 0.25%.

4.3. Enzymatic Assays

The esterolytic activity of ETE and ETE/P186G against the synthetic substrate Boc-L-Glu-↓OPh was determined as previously reported for ETA and ETB [19,23,29]. Briefly, solutions were prepared with a total volume of 800 µL, containing ETE or ETE/P186G, each at a final concentration of 1 µg/mL, 1-4 dioxane (Sigma) 2% (*v*/*v*) and variable Boc-L-Glu-↓OPh

concentrations (1.0, 5.0, 10.0, 15.0 and 20.0 mM) dissolved in HEPES 50 mM pH 7.4 buffer. The reaction progress was followed by measuring the absorbance at λ = 270 nm and *T* = 37 °C every 15 s for 20 min with a Cary UV-Vis Compact Peltier (Agilent) spectrophotometer. Each condition was assayed in triplicate. Initial velocities were estimated by finding the slopes of absorbance vs. time data through linear regression. Then, the kinetic parameters for ETE and ETE/P186G were estimated using the Hanes–Woolf linearization method [33]. Before the assays, Boc-L-Glu-↓OPh was synthesized as described in Text S2 and Figure S12.

4.4. Crystallization, Data Collection, Processing and Structure Determination

Crystals were obtained by hanging vapor diffusion from a buffer solution containing 30 mg/mL ETE, 20 mM Tris-HCl, pH 7.0 and 100 mM NaCl, which was equilibrated against a reservoir solution (4 M sodium formate, pH 6.0). Protein crystals were then frozen in a nitrogen gas stream at 100 K and diffracted at the W01B-MX2 beamline at the Brazilian Synchrotron Laboratory (LNLS, Campinas, Brazil). The wavelength of the radiation source was set to 1.458 Å and a Pilatus 2M detector was used to record the diffraction intensities. The crystal was exposed for 2 s per 0.1 degree of rotation and a total of 3600 images were collected. The data was indexed using the XDS package [40], and the structure was resolved by molecular replacement employing the atomic coordinates of another crystal structure of ETE (=ETD-like) (PDB: 5C2Z). Model refinement was performed through cycles of REFMAC5 [41] in the CCP4 software, followed by visual inspection of the electron density maps and manual reconstruction with WinCoot [42]. The validation of the new structure was carried out at the MolProbity server [43] and the atomic coordinates and properties of the crystal were deposited in the PDB with the code 8DAX.

4.5. Small-Angle X-ray Scattering Experiments

SAXS experiments were performed to evaluate the structural properties of ETE in solution close to the physiological conditions in terms of salt concentration and pH. The tests were conducted using the Xeuss 2.0 benchtop SAXS system from the XENOCS company, equipped with a Xenocs Genix 3D X-ray source with a copper anode that produces a beam with wavelength λ = 1.5419 Å. The collimation system, composed of a monochromator mirror and a set of scatterless slits, produces a beam of approximately 0.7 × 0.7 mm^2. The system uses a Pilatus 300K Dectris detector for all measurements performed with samples under vacuum conditions. The sample–detector distance was calibrated through measurements of a silver

behenate (AgBeh) standard that has peaks in the SAXS region providing a sample–detector distance of 1193 ± 8 mm. For measurement, 100 µL volume of ETE was used at a concentration of 5 mg/mL in 20 mM MES pH 7.0 and 150 mM NaCl. SAXS data were obtained from 7 cycles with run exposure times of 1800 s at a constant temperature of 25 °C. The 2D SAXS images were integrated with the Fit2D program [44]. Data integrations were processed and normalized in absolute scale with the SUPERSAXS program [45]. The treated curves were analyzed with the ATSAS software package [46], allowing us to obtain distribution curves of paired distances and to analyze the presence of oligomeric states.

SAXS can provide additional information that allows us to know the conformations of the ETE homodimers that better fit the experimental intensity profiles. As shown in Figure S6, the ability of 22 models of ETE homodimers to reproduce the experimental SAXS curves was assessed. For each model, we separated the monomer and homodimer structures. To evaluate the presence of monomers, homodimers or even mixtures in the system, the program Oligomer was used [46], which assessed the goodness of fit through chi-squared (χ^2) values and estimated the percentage of homodimers associated with each analyzed model. It was allowed to add a constant during data fitting. In principle, the homodimer models yielding theoretical intensity profiles with lower χ^2 values with respect to the experimental curve are more likely to be close to the actual ETE conformation in solution [35].

4.6. Computational Approaches to Predict the Conformation of ETE Homodimers

Each monomer of the reported ETE crystal structure (PDB: 8DAX) was submitted for ab initio docking at the web servers of five different protein–protein docking algorithms, i.e., ClusPro (https://cluspro.bu.edu/, access: date 10 August 2021) [36], HDOCK (http://hdock.phys.hust.edu.cn/, access date: 13 August 2021) [47], HADDOCK2.4 (https://bianca.science.uu.nl/haddock2.4/, access date: 20 August 2021) [48], GalaxyTongDock (GalaxyWeb: http://galaxy.seoklab.org/, access date: 18 August 2021) [49,50] and LZerD (https://lzerd.kiharalab.org/, access date: 21 August 2021) [51], in the above order and employing in each case the default parameters. Of note, HADDOCK2.4 docking simulations were restricted to a protein region that was predicted to contain interface-forming residues according to the CPORT (https://alcazar.science.uu.nl/services/CPORT/, access date: 25 August 2021) [52]. LZerD docking poses were generated using two different approaches, one of them involving C_2-symmetric docking and the other one not assuming a specific symmetry. Moreover, homology-based models of the homodimer structure were obtained at the

GalaxyWeb site [50,53]. The best nonredundant poses, i.e., those with pairwise RMSDs relative to their heavy atoms > 4 Å among the ten top-ranked solutions generated by each docking algorithm, were selected for refinement steps consisting of MD simulations, as well as MM-GBSA and US free energy calculations, in order to predict the most likely conformation of ETE dimers (Figure S6). SAXS was employed as an additional criterion for pose selection. The models yielding lower χ^2 values and, at the same time, having favorable US free energy values were selected for the next steps, which consisted of 1 μs-long MD simulations and additional free energy calculations (Figure S6).

The central structure collected from the 1 μs MD simulation of the selected model was subjected to further bioinformatic analyses to assess whether it corresponds to a biological assembly or not, using PISA (https://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver, access date: 9 September 2021) [26] and PRODIGY-CRYSTAL (https://bianca.science.uu.nl/prodigy/cryst, access date: 9 September 2021) [27] web servers. Moreover, SPPIDER was employed to predict whether dimer conformations contained IFRs [37]. SPPIDER makes use of solvent accessibility, amino acid residue conservation, charge conservation, amino acid residue size conservation, contacts and hydrophobicity in order to feed a neural network to classify residues on the interface or free surface. The model finally proposed as the most reliable conformation of ETE in solution was the one that passed all the steps of the devised protocol.

During the execution of this work, some reports showed that AlphaFold2-based approaches can predict the structure of protein–protein complexes with higher accuracy than protein–protein docking algorithms [54,55]. Therefore, we decided to assess in parallel the performance of these new approaches when applied to the prediction of the ETE homodimer structure. We employed for that purpose the AlphaFold-multimer [54] version implemented within the ColabFold open-source software [56]. As before, the top-ranked model was subjected to 100 ns MD simulations and subsequent free energy calculations. Moreover, the ability of this model to fit the SAXS data was assessed.

4.7. Molecular Dynamics Simulations

ETs were protonated at the H++ web server (http://newbiophysics.cs.vt.edu/H++/, access date: 16 July 2021) by setting the pH = 7.4 [57]. Each complex was then solvated in an octahedral box with edges placed at least 10 Å away from the protein surface using *tleap* of Amber20 [58]. The simulation box was then filled with TIP3P waters, and sufficient counterions (Na^+ or Cl^-) were added to neutralize the net charge. The protein parameters were derived from Amber

ff14SB force-field [59]. All MD simulations were run with *pmemd.cuda* of Amber 20 [58,60]. More details about the MD simulation setup can be found elsewhere [61].

Each solvated system was subjected to energy minimization (EM) followed by an NVT heating and NPT equilibration, each run for 500 ps in the presence of harmonic restraints acting on the protein heavy atoms (k = 10 kcal·mol^{-1}·Å^{-2}), to reach a temperature of 298 K and a pressure of 1 bar. Then, the harmonic restraints were gradually released from 8 to 2 kcal·mol^{-1}·Å^{-2}, with a 2 kcal·mol^{-1}·Å^{-2} stride, in four 500 ps NPT simulations. All production runs were conducted using hydrogen mass repartitioning (HMR) in order to increase the time step from 2 to 4 fs [62].

4.8. MM-GBSA Free Energy Calculations

MM-GBSA free energy calculations were calculated for all the simulated ETE homodimers with the MMPBSA.py program of Amber 20 [58,63]. The generated 100 ns/1 μs trajectories were desolvated, and the frames collected after 10 ns/100 ns were used to calculate the corresponding ΔG_{eff} values, which do not contain the configurational entropy component. The trajectories of the free monomers were extracted from those of the simulated complexes following the so-called single trajectory approach. The GB^{OBC} (*igb* = 2) implicit solvation model was employed to estimate the polar solvation free energies [58,64]. The internal and external dielectric constants were taken as default and the salt concentration was set to 0.1 M. The atomic radii were derived from the mbondi2 set, as recommended [58]. Moreover, per-residue free energy decomposition was carried out for the ETE homodimer with the GB^{OBC} model using MMPBSA.py [63] in order to predict the hotspot residues at the PPI.

4.9. Umbrella Sampling Free Energy Calculations

US is a robust approach to estimating binding free energies (ΔG_{bind}) [65] and thus was employed to predict more accurately the most favorable ETE homodimer conformation. As a result, PMFs associated with the separation of the ETE monomers from different initial ETE homodimer conformations along the *z* Cartesian axis were determined. Briefly, different ETE homodimer conformations were carefully oriented, and the monomers were separated along the *z* axis by displacing the second chain from −2 to 20 Å relative to its equilibrium position in the starting conformation. In total, 23 windows separated by 1 Å were generated and the proteins were attached by means of a special set of restraints to three dummy atoms fixed within the simulation box. This prevented the translation and rotation motions of the ETE chains relative to their centers of mass during the MD simulations carried out for each window [66]. Harmonic

potentials were also applied to keep the second protein chain at the corresponding restraint equilibrium position within each window.

Similarly, US free energy calculations were performed to determine the PMF associated with the rotation around the N-Cα-C-O dihedral (ξ) of residue P186 of ETE. Windows were created every 5° from −180° to 175° to span the full rotation of the dihedral, and harmonic restraints were employed to keep the dihedral fluctuating around each pre-established equilibrium position. In all cases, PMFs were obtained by combining the results from all the simulated windows using the weighted histogram analysis (WHAM) [67]. Details of the US protocols employed here are included in Text S3. We followed identical procedures to calculate the PMFs associated with the rotation of N-Cα-C-O dihedrals of residue G186 of ETE/P186G (modeled from wild-type structure PDB 8DAX using Pymol mutagenesis plugging [32]) and residue G166 of V8 protease from *S. aureus* (PDB: 2O8L). Moreover, as a control, we obtained PMF for the rotation around the peptide bond (CA-C-N-CA dihedral) of the dipeptide L-alanyl-L-proline in solution at 313.15 K for which experimental data is available [28]. This allowed us to assess the accuracy of the US free energy calculation protocol. Details of the US protocols employed here are included in Text S3.

4.10. Trajectory Analyses

RMSDs and interatomic distances along the trajectories were calculated using the *rms* and *distance* commands, respectively, of *cpptraj* module of Amber20. Trajectory clustering was performed with the *cluster* command of *cpptraj* using the average linkage algorithm [58,68]. The RMSD with respect to the heavy atoms of the interface residues, defined by a 4 Å cut-off from each interacting chain, was set as a metric for the clustering analysis. As a rule of thumb, five central structures were generated per trajectory, and the one corresponding to the largest cluster was chosen for structural representation or as the starting structure for US free energy calculations. H-bonds were determined using *hbond* of *cpptraj* by setting a donor–acceptor upper cut-off distance of 3.5 Å and a donor–hydrogen–acceptor lower cut-off angle of 120°.

5. Conclusions

In the present work, we have presented results suggesting that the formation of homodimers may play a key role in controlling the enzymatic activity of ETE and, potentially, the remaining ETs. Moreover, we found that the flipped conformation of residue *192*, which has long been

considered as the key structural feature of the ETs underlying their fine-tuned activity, does not seem to be as relevant to this process, at least for ETE. A better understanding of the exfoliative mechanism of the ETs will be valuable for the design of molecules with the capacity to modulate the enzymatic activity of these enzymes and thus prevent the degradation of their natural substrate Dsg1.

6. Author Contributions

Conceptualization, C.G. and J.E.H.G.; data curation, C.G., J.E.H.G., R.K.A. and R.B.M.; funding acquisition, P.G.P. and R.K.A.; investigation, C.G., J.E.H.G., A.R.N.S., R.K.A. and R.B.M.; methodology, C.G., J.E.H.G., A.R.N.S., A.F.Z.N., L.d.S.C., F.R.d.M., C.L.P.d.O., C.C.O., P.J.S.P. and R.B.M.; project administration, R.K.A. and R.B.M.; resources, C.L.P.d.O., C.C.O., P.G.P. and R.K.A.; supervision, R.B.M.; writing—original draft, C.G., J.E.H.G. and R.B.M.; writing—review and editing, J.E.H.G., P.G.P., R.K.A. and R.B.M. All authors have read and agreed to the published version of the manuscript.

7. Acknowledgments

We thank Yves Le Loir for the donation of pD441/ETE expression vector. J.E.H.G. and R.K.A. acknowledge the National Laboratory for Scientific Computing (LNCC/MCTI, Brazil) for providing HPC resources of the SDumont supercomputer, which have contributed to the research results reported within this work (http://sdumont.lncc.br, accessed date: 26 May 2022). This research used facilities of the Brazilian Synchrotron Light Laboratory (LNLS), part of the Brazilian Center for Research in Energy and Materials (CNPEM), a private nonprofit organization under the supervision of the Brazilian Ministry for Science, Technology, and Innovations (MCTI). The W01B-MX2 beamline staff is acknowledged for their assistance during experiments 20190086.

8. References

1. Bukowski M., Wladyka B., Dubin G. Exfoliative toxins of *Staphylococcus aureus*. *Toxins.* 2010;2:1148–1165. doi: 10.3390/toxins2051148. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

2. Mariutti R.B., Tartaglia N.R., Le Loir Y., Nishifuji K. Exfoliative toxins of *Staphylococcus aureus*. In: Enany S., Crotty Alexander L.E., editors. *The Rise of Virulence and Resistance in Staphylococcus aureus.* IntechOpen; London, UK: 2017. pp. 127–143. [Google Scholar]

3. Amagai M., Yamaguchi T., Hanakawa Y., Nishifuji K., Sugai M., Stanley J.R. Staphylococcal exfoliative toxin B specifically cleaves desmoglein 1. *J. Investig. Dermatol.* 2002;118:845–850. doi: 10.1046/j.1523-1747.2002.01751.x. [PubMed] [CrossRef] [Google Scholar]

4. Hanakawa Y., Schechter N.M., Lin C., Nishifuji K., Amagai M., Stanley J.R. Enzymatic and molecular characteristics of the efficiency and specificity of exfoliative toxin cleavage of desmoglein 1. *J. Biol. Chem.* 2004;279:5268–5277. doi: 10.1074/jbc.M311087200. [PubMed] [CrossRef] [Google Scholar]

5. Hanakawa Y., Selwood T., Woo D., Lin C., Schechter N.M., Stanley J.R. Calcium-dependent conformation of desmoglein 1 is required for its cleavage by exfoliative toxin. *J. Investig. Dermatol.* 2003;121:383–389. doi: 10.1046/j.1523-1747.2003.12362.x. [PubMed] [CrossRef] [Google Scholar]

6. O'Toole P.W., Foster T.J. Nucleotide sequence of the epidermolytic toxin A gene of *Staphylococcus aureus*. *J. Bacteriol.* 1987;169:3910–3915. doi: 10.1128/jb.169.9.3910-3915.1987. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

7. Lee C.Y., Schmidt J.J., Johnson-Winegar A.D., Spero L., Iandolo J.J. Sequence determination and comparison of the exfoliative toxin A and toxin B genes from *Staphylococcus aureus*. *J. Bacteriol.* 1987;169:3904–3909. doi: 10.1128/jb.169.9.3904-3909.1987. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

8. Yamaguchi T., Nishifuji K., Sasaki M., Fudaba Y., Aepfelbacher M., Takata T., Ohara M., Komatsuzawa H., Amagai M., Sugai M. Identification of the *Staphylococcus aureus* etd pathogenicity island which encodes a novel exfoliative toxin, ETD, and EDIN-B. *Infect. Immun.* 2002;70:5835–5845. doi: 10.1128/IAI.70.10.5835-5845.2002. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

9. Le Marechal C., Jardin J., Jan G., Even S., Pulido C., Guibert J.M., Hernandez D., Francois P., Schrenzel J., Demon D., et al. *Staphylococcus aureus* seroproteomes discriminate ruminant isolates causing mild or severe mastitis. *Vet. Res.* 2011;42:35. doi: 10.1186/1297-9716-42-35. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

10. Le Marechal C., Seyffert N., Jardin J., Hernandez D., Jan G., Rault L., Azevedo V., Francois P., Schrenzel J., van de Guchte M., et al. Molecular basis of virulence in *Staphylococcus aureus* mastitis. *PLoS ONE*. 2011;6:e27354. doi: 10.1371/journal.pone.0027354. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

11. Kapral F.A., Miller M.M. Product of *Staphylococcus aureus* responsible for the scalded-skin syndrome. *Infect. Immun.* 1971;4:541–545. doi: 10.1128/iai.4.5.541-545.1971. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

12. Arbuthnott J.P., Kent J., Lyell A., Gemmell C.G. Toxic epidermal necrolysis produced by an extracellular product of *Staphylococcus aureus*. *Br. J. Derm.* 1971;85:145–149. doi: 10.1111/j.1365-2133.1971.tb07200.x. [PubMed] [CrossRef] [Google Scholar]

13. Melish M.E., Glasgow L.A. The staphylococcal scalded-skin syndrome. *N. Engl. J. Med.* 1970;282:1114–1119. doi: 10.1056/NEJM197005142822002. [PubMed] [CrossRef] [Google Scholar]

14. Imanishi I., Nicolas A., Caetano A.B., Castro T.L.P., Tartaglia N.R., Mariutti R., Guedon E., Even S., Berkova N., Arni R.K., et al. Exfoliative toxin E, a new *Staphylococcus aureus* virulence factor with host-specific activity. *Sci. Rep.* 2019;9:16336. doi: 10.1038/s41598-019-52777-3. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

15. Wegener H.C., Andresen L.O., Bille-Hansen V. Staphylococcus hyicus virulence in relation to exudative epidermitis in pigs. *Can. J. Vet. Res.* 1993;57:119–125. [PMC free article] [PubMed] [Google Scholar]

16. Chen S., Wang Y., Chen F., Yang H., Gan M., Zheng S.J. A highly pathogenic strain of *Staphylococcus sciuri* caused fatal exudative epidermitis in piglets. *PLoS ONE.* 2007;2:e147. [PMC free article] [PubMed] [Google Scholar]

17. Li H., Wang Y., Ding L., Zheng S.J. *Staphylococcus sciuri* exfoliative toxin C (ExhC) is a necrosis-inducer for mammalian cells. *PLoS ONE.* 2011;6:e23145. doi: 10.1371/journal.pone.0023145. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

18. Demidyuk I.V., Chukhontseva K.N., Kostrov S.V. Glutamyl Endopeptidases: The Puzzle of Substrate Specificity. *Acta Nat.* 2017;9:17–33. doi: 10.32607/20758251-2017-9-2-17-33. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

19. Cavarelli J., Prevost G., Bourguet W., Moulinier L., Chevrier B., Delagoutte B., Bilwes A., Mourey L., Rifai S., Piemont Y., et al. The structure of *Staphylococcus aureus* epidermolytic toxin A, an atypic serine protease, at 1.7 A resolution. *Structure.* 1997;5:813–824. doi: 10.1016/S0969-2126(97)00235-9. [PubMed] [CrossRef] [Google Scholar]

20. Vath G.M., Earhart C.A., Rago J.V., Kim M.H., Bohach G.A., Schlievert P.M., Ohlendorf D.H. The structure of the superantigen exfoliative toxin A suggests a novel regulation as a serine protease. *Biochemistry.* 1997;36:1559–1566. doi: 10.1021/bi962614f. [PubMed] [CrossRef] [Google Scholar]

21. Papageorgiou A.C., Plano L.R., Collins C.M., Acharya K.R. Structural similarities and differences in *Staphylococcus aureus* exfoliative toxins A and B as revealed by their crystal structures. *Protein Sci.* 2000;9:610–618. doi: 10.1110/ps.9.3.610. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

22. Iyori K., Futagawa-Saito K., Hisatsune J., Yamamoto M., Sekiguchi M., Ide K., Son W.G., Olivry T., Sugai M., Fukuyasu T., et al. Staphylococcus pseudintermedius exfoliative toxin EXI selectively digests canine desmoglein 1 and causes subcorneal clefts in canine epidermis. *Vet. Derm.* 2011;22:319–326. [PubMed] [Google Scholar]

23. Bailey C.J., Redpath M.B. The esterolytic activity of epidermolytic toxins. Pt 1*Biochem. J.* 1992;284:177–180. [PMC free article] [PubMed] [Google Scholar]

24. Rago J.V., Vath G.M., Tripp T.J., Bohach G.A., Ohlendorf D.H., Schlievert P.M. Staphylococcal exfoliative toxins cleave alpha- and beta-melanocyte-stimulating hormones. *Infect. Immun.* 2000;68:2366–2368. [PMC free article] [PubMed] [Google Scholar]

25. Mariutti R.B., Souza T.A., Ullah A., Caruso I.P., de Moraes F.R., Zanphorlin L.M., Tartaglia N.R., Seyffert N., Azevedo V.A., Le Loir Y., et al. Crystal structure of Staphylococcus aureus exfoliative toxin D-like protein: Structural basis for the high specificity of exfoliative toxins. *Biochem Biophys Res Commun.* 2015;467:171–177. [PubMed] [Google Scholar]

26. Krissinel E., Henrick K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 2007;372:774–797. [PubMed] [Google Scholar]

27. Jimenez-Garcia B., Elez K., Koukos P.I., Bonvin A.M., Vangone A. PRODIGY-crystal: A web-tool for classification of biological interfaces in protein complexes. *Bioinformatics.* 2019;35:4821–4823. [PMC free article] [PubMed] [Google Scholar]

28. Shibukawa M., Miyake A., Eda S., Saito S. Determination of the cis-trans isomerization barriers of L-alanyl-L-proline in aqueous solutions and at water/hydrophobic interfaces by on-line temperature-jump relaxation HPLC and dynamic on-column reaction HPLC. *Anal. Chem.* 2015;87:9280–9287. [PubMed] [Google Scholar]

29. Rago J.V., Vath G.M., Bohach G.A., Ohlendorf D.H., Schlievert P.M. Mutational analysis of the superantigen staphylococcal exfoliative toxin A (ETA) *J. Immunol.* 2000;164:2207–2213. [PubMed] [Google Scholar]

30. Kawalec M., Potempa J., Moon J.L., Travis J., Murray B.E. Molecular diversity of a putative virulence factor: Purification and characterization of isoforms of an extracellular serine glutamyl endopeptidase of Enterococcus faecalis with different enzymatic activities. *J. Bacteriol.* 2005;187:266–275. [PMC free article] [PubMed] [Google Scholar]

31. Meijers R., Blagova E.V., Levdikov V.M., Rudenskaya G.N., Chestukhina G.G., Akimkina T.V., Kostrov S.V., Lamzin V.S., Kuranova I.P. The crystal structure of glutamyl endopeptidase from Bacillus intermedius reveals a structural link between zymogen activation and charge compensation. *Biochemistry.* 2004;43:2784–2791. [PubMed] [Google Scholar]

32. DeLano W.L. *PyMOL, 2.1.0.* DeLano Scientific; San Carlos, CA, USA: 2002. [Google Scholar]

33. Hanes C.S. Studies on plant amylases: The effect of starch concentration upon the velocity of hydrolysis by the amylase of germinated barley. *Biochem. J.* 1932;26:1406–1421. [PMC free article] [PubMed] [Google Scholar]

34. Li H., Li X., Lu Y., Wang X., Zheng S.J. *Staphylococcus sciuri* exfoliative toxin C is a dimer that modulates macrophage functions. *Can. J. Microbiol.* 2011;57:722–729. [PubMed] [Google Scholar]

35. Oliveira C.L.P. Investigating macromolecular complexes in solution by small angle X-Ray scattering. In: Chandasekaran A., editor. *Current Trends in X-Ray Crystallography.* IntechOpen; London, UK: 2011. pp. 367–392. [Google Scholar]

36. Kozakov D., Hall D.R., Xia B., Porter K.A., Padhorny D., Yueh C., Beglov D., Vajda S. The ClusPro web server for protein-protein docking. *Nat. Protoc.* 2017;12:255–278. [PMC free article] [PubMed] [Google Scholar]

37. Porollo A., Meller J. Prediction-based fingerprints of protein-protein interactions. *Proteins.* 2007;66:630–645. [PubMed] [Google Scholar]

38. Gaber A., Pavsic M. Modeling and structure determination of homo-oligomeric proteins: An overview of challenges and current approaches. *Int. J. Mol. Sci.* 2021;22:9081. [PMC free article] [PubMed] [Google Scholar]

39. Peters K., Braun H.P. Comparative analyses of protein complexes by blue native DIGE. In: Cramer R., Westermeier R., editors. *Difference Gel Electrophoresis (DIGE) Methods and Protocols.* Humana Press; New York, NY, USA: 2012. pp. 145–154. [PubMed] [Google Scholar]

40. Kabsch W. XDSS. *Acta Cryst. D Biol Cryst.* 2010;66:125–132. [PMC free article] [PubMed] [Google Scholar]

41. Murshudov G.N., Skubak P., Lebedev A.A., Pannu N.S., Steiner R.A., Nicholls R.A., Winn M.D., Long F., Vagin A.A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst. D Biol. Cryst.* 2011;67:355–367. [PMC free article] [PubMed] [Google Scholar]

42. Emsley P., Cowtan K. Coot: Model-building tools for molecular graphics. *Acta Cryst. D Biol. Cryst.* 2004;60:2126–2132. [PubMed] [Google Scholar]

43. Chen V.B., Arendall W.B., 3rd, Headd J.J., Keedy D.A., Immormino R.M., Kapral G.J., Murray L.W., Richardson J.S., Richardson D.C. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Cryst. D Biol. Cryst.* 2010;66:12–21. [PMC free article] [PubMed] [Google Scholar]
44. Hammersley A.P. FIT2D: A multi-purpose data reduction, analysis and visualization program. *J. Appl. Cryst.* 2016;49:646–652. [Google Scholar]
45. Oliveira C.L.P., Vorup-Jensen T., Andersen C.B.F., Andersen G.R., Pedersen J.S. Discovering new features of protein complexes structures by small-angle X-ray scattering. In: Gomez M., Nogales A., Garcia-Gutierrez M.C., Ezquerra T.A., editors. *Applications of Synchrotron Light to Scattering and Diffraction in Materials and Life Sciences.* Springer Berlin; Heidelberg, Germany: 2009. pp. 231–244. [Google Scholar]
46. Petoukhov M.V., Franke D., Shkumatov A.V., Tria G., Kikhney A.G., Gajda M., Gorba C., Mertens H.D., Konarev P.V., Svergun D.I. New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Cryst.* 2012;45:342–350. [PMC free article] [PubMed] [Google Scholar]
47. Yan Y., Tao H., He J., Huang S.Y. The HDOCK server for integrated protein-protein docking. *Nat. Protoc.* 2020;15:1829–1852. [PubMed] [Google Scholar]
48. Dominguez C., Boelens R., Bonvin A.M. HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* 2003;125:1731–1737. doi: 10.1021/ja026939x. [PubMed] [CrossRef] [Google Scholar]
49. Park T., Baek M., Lee H., Seok C. GalaxyTongDock: Symmetric and asymmetric ab initio protein-protein docking web server with improved energy parameters. *J. Comput. Chem.* 2019;40:2413–2417. [PubMed] [Google Scholar]
50. Ko J., Park H., Heo L., Seok C. GalaxyWEB server for protein structure prediction and refinement. *Nucleic. Acids Res.* 2012;40:W294–W297. doi: 10.1093/nar/gks493. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
51. Christoffer C., Bharadwaj V., Luu R., Kihara D. LZerD protein-protein docking webserver enhanced with de novo structure prediction. *Front. Mol. Biosci.* 2021;8:724947. [PMC free article] [PubMed] [Google Scholar]
52. de Vries S.J., Bonvin A.M. CPORT: A consensus interface predictor and its performance in prediction-driven docking with HADDOCK. *PLoS ONE.* 2011;6:e17695. [PMC free article] [PubMed] [Google Scholar]
53. Park T., Won J., Baek M., Seok C. GalaxyHeteromer: Protein heterodimer structure prediction by template-based and ab initio docking. *Nucleic Acids Res.* 2021;49:W237–W241. [PMC free article] [PubMed] [Google Scholar]
54. Evans R., O'Neill M., Pritzel A., Antropova N., Senior A., Green T. Protein complex prediction with AlphaFold-Multimer. *Biorxiv [Preprint]* 2021 doi: 10.1101/2021.10.04.463034. [CrossRef] [Google Scholar]
55. Ghani U., Desta I., Jindal A., Khan O., Jones G., Kotelnikov S., Padhorny D., Vajda S., Kozakov D. Improved docking of protein models by a combination of alphafold2 and cluspro. *Biorxiv [Preprint]* 2021 doi: 10.1101/2021.09.07.459290. [CrossRef] [Google Scholar]
56. Mirdita M., Schutze K., Moriwaki Y., Heo L., Ovchinnikov S., Steinegger M. ColabFold: Making protein folding accessible to all. *Nat. Methods.* 2022;19:679–682. doi: 10.1038/s41592-022-01488-1. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

57. Gordon J.C., Myers J.B., Folta T., Shoja V., Heath L.S., Onufriev A. H^{++}: A server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* 2005;33:W368–W371. doi: 10.1093/nar/gki464. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

58. Case D.A., Belfon K., Ben-Shalom I.Y., Brozell S.R., Cerutti D.S., Cheatham T.E., III, Cruzeiro V.W.D., Darden T.A., Duke R.E., Giambasu G., et al. *Amber 2020.* University of California; San Francisco, CA, USA: 2020. [Google Scholar]

59. Maier J.A., Martinez C., Kasavajhala K., Wickstrom L., Hauser K.E., Simmerling C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* 2015;11:3696–3713. doi: 10.1021/acs.jctc.5b00255. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

60. Salomon-Ferrer R., Gotz A.W., Poole D., Le Grand S., Walker R.C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh ewald. *J. Chem. Theory Comput.* 2013;9:3878–3888. doi: 10.1021/ct400314y. [PubMed] [CrossRef] [Google Scholar]

61. Hernandez Gonzalez J.E., Hernandez Alvarez L., Pascutti P.G., Valiente P.A. Predicting binding modes of reversible peptide-based inhibitors of falcipain-2 consistent with structure-activity relationships. *Proteins.* 2017;85:1666–1683. doi: 10.1002/prot.25322. [PubMed] [CrossRef] [Google Scholar]

62. Hopkins C.W., Le Grand S., Walker R.C., Roitberg A.E. Long-time-step molecular dynamics through hydrogen mass repartitioning. *J. Chem. Theory Comput.* 2015;11:1864–1874. doi: 10.1021/ct5010406. [PubMed] [CrossRef] [Google Scholar]

63. Miller B.R., 3rd, McGee T.D., Jr., Swails J.M., Homeyer N., Gohlke H., Roitberg A.E. MMPBSA.py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* 2012;8:3314–3321. doi: 10.1021/ct300418h. [PubMed] [CrossRef] [Google Scholar]

64. Onufriev A., Bashford D., Case D.A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins.* 2004;55:383–394. doi: 10.1002/prot.20033. [PubMed] [CrossRef] [Google Scholar]

65. Kästner J. Umbrella sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2011;1:932–942. doi: 10.1002/wcms.66. [CrossRef] [Google Scholar]

66. Heinzelmann G., Henriksen N.M., Gilson M.K. Attach-pull-release calculations of ligand binding and conformational changes on the first BRD4 bromodomain. *J. Chem. Theory Comput.* 2017;13:3260–3275. doi: 10.1021/acs.jctc.7b00275. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

67. Kumar S., Rosenberg J.M., Bouzida D., Swendsen R.H., Kollman P.A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* 1992;13:1011–1021. doi: 10.1002/jcc.540130812. [CrossRef] [Google Scholar]

68. Shao J., Tanner S.W., Thompson N., Cheatham T.E. Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. *J. Chem. Theory Comput.* 2007;3:2312–2334. doi: 10.1021/ct700119m. [PubMed] [CrossRef] [Google Scholar]

69. Hong P., Koza S., Bouvier E.S. Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates. *J. Liq. Chromatogr. Relat. Technol.* 2012;35:2923–2950. doi: 10.1080/10826076.2012.743724. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

70. Madhusudhan M.S., Webb B.M., Marti-Renom M.A., Eswar N., Sali A. Alignment of multiple protein structures based on sequence and structure features. *Protein Eng. Des. Sel.* 2009;22:569–574. doi: 10.1093/protein/gzp040. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

71. Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., et al. Clustal W and Clustal X version 2.0. *Bioinformatics.* 2007;23:2947–2948. doi: 10.1093/bioinformatics/btm404. [PubMed] [CrossRef] [Google Scholar]

72. Robin S., Zhu J., Galons H., Chuong P.-H., Claude J.R., Tomas A., Viossat B. A convenient asymmetric synthesis of thalidomide. *Tetrahedron Asymmetry.* 1995;6:1249–1252. doi: 10.1016/0957-4166(95)00155-I. [CrossRef] [Google Scholar]

73. Doudou S., Burton N.A., Henchman R.H. Standard Free Energy of Binding from a One-Dimensional Potential of Mean Force. *J Chem Theory Comput.* 2009;5:909–918. doi: 10.1021/ct8002354. [PubMed] [CrossRef] [Google Scholar]

74. Abraham M.J., Murtola T., Schulz R., Páll S., Smith J.C., Hess B., Lindahl E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 2015;1:19–25. doi: 10.1016/j.softx.2015.06.001. [CrossRef] [Google Scholar]

75. Hernandez Gonzalez J.E., Hernandez Alvarez L., Pascutti P.G., Leite V.B.P. Prediction of Noncompetitive Inhibitor Binding Mode Reveals Promising Site for Allosteric Modulation of Falcipain-2. *J. Phys. Chem. B.* 2019;123:7327–7342. doi: 10.1021/acs.jpcb.9b05021. [PubMed] [CrossRef] [Google Scholar]

Capítulo IV. Identification of potential partner protein of exfoliative toxin C from *Mammaliicoccus sciuri* and inhibitors/activators of its enzymatic activity

Identification of potential partner protein of exfoliative toxin C from *Mammaliicoccus sciuri* and inhibitors/activators of its enzymatic activity

Authors: Carolina Gismene [a, b], Mônika Aparecida Coronado[b], Raphael Eberle[b], Jorge Enrique Hernández González[a], Mohammadamin Mastalipour[c], Dieter Wilbold[b], Raghuvir Krishnaswamy Arni[a]

[a] Multiuser Center for Biomolecular Innovation, São Paulo State University - UNESP, São José do Rio Preto, SP, Brazil.

[b] Institute of Biological Information Processing (IBI-7: Structural Biochemistry), Forschungszentrum Jülich, 52428 Jülich, Germany.

[c] Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße, 40225 Düsseldorf, Germany.

Abstract

Exfoliative toxin C (ExhC), expressed by the bacterium *Mammaliicoccus sciuri*, is an atypical serine proteinase. In contrast to other exfoliative toxins (ETs), it has been classified as a moonlighting protein for its ability to induce necrosis in cell lines, inhibit macrophage phagocytic activity, in addition to causing skin exfoliation in pigs and mice. Previous experiments demonstrated that the latter function stems from the high specificity of ExhC for porcine and murine desmoglein-1 (Dsg1), a cadherin contributing to cell-cell adhesion in the epidermis. While the amino acid residues of ExhC responsible for cellular necrosis have been identified, information on the mode of action causing partial inhibition of the innate immune system is still unknown. Despite advances in understanding how this toxin operates, inhibitors for these diverse actions and potential molecular targets during the processes of cellular necrosis and macrophage phagocytic activity inhibition have not been described. Through the application of the phage display technique, we identified a potential partner protein from the murine host with high affinity for ExhC, alongside the discovery of inhibitors and activators that modulate the enzymatic activity of this toxin.

Keywords: exfoliative toxin C, *Mammaliicoccus sciuri*, esterolytic activity, phage display, peptides, inhibitors, potentiators, partner protein.

1. Introdution

Mammaliicoccus sciuri is a commensal and pathogenic bacterium of veterinary and clinical relevance (Nemeghaire; Argudin; Fessler; Hauschild et al., 2014). Recently, the species was described as the etiologic agent of exudative epidermitis (EE) in pigs, and the main virulence

factor involved in this clinical manifestation was exfoliative toxin C (ExhC) (Chen; Wang; Chen; Yang et al., 2007). This communicable skin disease, commonly caused by *S. hyicus*, manifests as an acute infection mainly in newly-weaned pigs and is characterized by exfoliation accompanied by epidermal cell separation, erythema and serous exudation (Ahrens; Andresen, 2004). The pig with generalized EE suffers from dehydration that can lead to death in a few days (Wegener; Andresen; Bille-Hansen, 1993).

ExhC belongs to a group of glutamyl endopeptidases of the chymotrypsin family collectively known as exfoliative toxins (ETs) (Bukowski; Wladyka; Dubin, 2010; Demidyuk; Dhukhontseva; Kostrov, 2017). Unlike most members of this family, ETs are inactive against a wide range of serine-proteinases substrates, showing particular specificity for desmoglein-1 (Dsg1), a multi-domain transmembrane protein that plays an important role in cell-cell adhesion in the epidermis (Bukowski; Wladyka; Dubin, 2010).

Pathogenic strains of *S. hyicus* secrete six ETs isoforms (ExhA, ExhB, ExhC, ExhD, ShetA and ShetB) that are associated with EE in swine (Ahrens; Andresen, 2004; Sato; Watanabe; Higuchi; Teruya et al., 2000), and two of them (ExhA and ExhC) also cause epidermal exfoliation in mice (Ahrens; Andresen, 2004). Furthermore, *S. aureus*, *S. chromogenes* and the recently-discovered *S. sciuri* – currently reclassified as *Mammaliicoccus sciuri* (Madhaiyan; Wirth; Saravanan, 2020) – and *S. pseudintermedius* produce ETs associated with severe skin infections, each with its unique features that define host specificity (Li; Wang; Ding; Zheng, 2011; Mariutti; Tartaglia; Seyffert; De Paula Castro et al., 2017; Imanishi; Nicolas; Caetano; Castro et al., 2019).

Through sequential alignment, the complete identity of *M. sciuri* – ExhC (GenBank ID: JF755400) and *S. hyicus* – ExhC (GenBank ID: AF515455) was observed. Apart from causing epidermal exfoliation in pigs and mice, recombinant ExhC of *M. sciuri* induced necrosis in multiple mammalian cell lines, including the renal fibroblasts of newborn hamsters (BHK-21) (Li; Wang; Ding; Zheng, 2011), and affects the host's innate immune system from the inhibition of phagocytosis by RAW264.7 macrophages (Li; Li; Lu; Wang et al., 2011). These properties, not observed for any other ET until now, suggests that, in addition to skin exfoliation, ExhC may be involved in other biochemical functions during *M. sciuri* and *S. hyicus* infections. Remarkably, the ability to display orthogonal functions is a characteristic found in certain proteins, a phenomenon that is termed moonlighting (www.moonlightingproteins.org/) (Jeffery, 2015).

Experiments conducted by Li et al. (2011a) revealed that $ExhC_{79-128}$ fragment (or $ExhC_{44–92}$ fragment in the absence of the signal peptide) can trigger necrosis and, subsequently, it was identified that the absence of specific residues R47, N49, Q51 and R89 abrogate the necrotic activity of ExhC (Gismene; Gonzalez; De Freitas Calmon; Nascimento et al., 2024). Of note, such mutations do not significantly modify the esterolytic activity of the protein, indicating that the necrotic and catalytic activities are independent functions involving distinct protein sites (Gismene; Gonzalez; De Freitas Calmon; Nascimento et al., 2024). The detection of a cavity containing all the residues responsible for necrotic activity is an interesting finding that provides evidence of the interaction of ExhC with biomolecules of the cells susceptible of undergoing necrosis (Gismene; Gonzalez; De Freitas Calmon; Nascimento et al., 2024). However, the specific molecular target(s) involved in this process have not yet been identified. Likewise, the molecules present in RAW264.7 macrophages that interact with ExhC and trigger damage to the host's innate immune system are unknown (Li; Li; Lu; Wang et al., 2011).

In this work, we used phage display to identify peptides exhibiting affinity for ExhC. To this end, we chose the TriCo-16 library, which contains $2.6x10^{10}$ peptides long enough to fold into secondary structural elements. This property is particularly useful to track interactions between the target and partner proteins possessing segments homolog to the identified structured peptides (Addepalli; Rao; Hunt, 2015; Petsalaki; Stark; García-Urdiales; Russell, 2009). In this study, we were able to identify through phage display several peptides showing high affinity for ExhC. Remarkably, some of these peptides act as inhibitors or potentiators of the toxin's esterolytic activity. Moreover, based on the sequence of the identified peptides, we proposed the interaction between ExhC and Pcyox1l from *M. musculus*, a protein involved in the maintenance of the cell membrane and responsible for bactericidal properties, which suggests that it could help mediate the effects of ExhC on macrophages and fibroblasts.

2. Materials and methods

2.1 Protein expression and purification

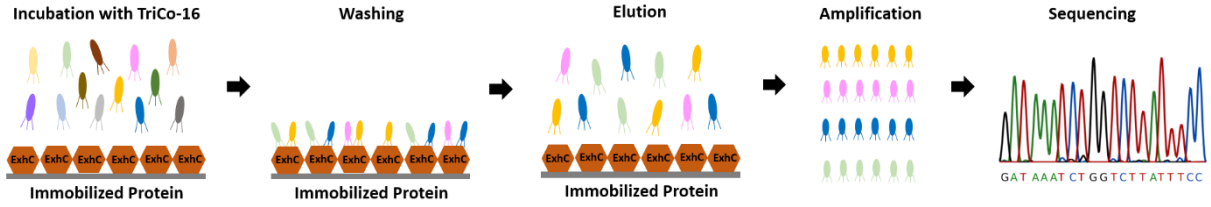
Escherichia coli BL21 (DE3) competent cells were transformed with the pET-28a (+)-ExhC expression vector (GenScript). A single colony of bacterial transformation was selected and grown overnight at 37 ºC in LB medium supplemented with kanamycin (50 mg/mL) (Rosano; Ceccarelli, 2014). Subsequently, the culture was diluted 100-fold with a fresh LB medium containing kanamycin (50 mg/mL) and incubated at 37 ºC. When the optical density measured at 600 nm (OD_{600}) reached 0.5, the expression was induced with 0.4 mM IPTG for 5 h at ºC.

The cells were collected by centrifugation at 2600 g for 10 min at 4 ºC and homogenized by sonication in a lysis buffer containing 5 mM $NaHPO_4$, pH 7.7, 400 mM NaCl, 10 mM imidazole and 10 % (v/v) glycerol (Gismene; Gonzalez; De Freitas Calmon; Nascimento et al., 2024). The lysed cells were centrifuged at 15,000 g for 30 min at 4° C, and the supernatant was subjected to affinity chromatography using a Ni^{2+}-Sepharose column (Bio-Rad) (Rodriguez; Poddar; Iftekhar; Suh et al., 2020). Protein concentration was determined to be 10 mg/mL by a spectrophotometer (Biomate 3S - Thermo scientific) and an additional purification step was performed in a 20 mM Tris-HCl buffer, pH 7.5 containing 100 mM NaCl using an AKTA system purifier by SEC (Burgess, 2018) on a Superdex G75 10/300 GL column (GE - Healthcare Life Sciences) at a flow rate of 0.7 mL/min. Western blot and SDS-PAGE evaluated protein purity by 15 % (Peters; Braun, 2012).

2.2 Phage display with TriCo-16 library

The phage display with the TriCo-16 library (Creative Biolabs, Shirley, USA) was performed using the ExhC (Fig. 1). Experimentally, three rounds of selection were performed to select only the high-binding affinity peptides. Briefly, for each selection step, wells of a high-binding plate – 3D-NHS (PolyAn – Molecular Surface Engineering) in assays with the protein were sensitized for 30 minutes with 30 pM of protein dissolved in 100 µL of 10 mM Sodium Acetate buffer pH 5.2. Then, the solution containing the target was removed and 100 µL of Quenching buffer (1 M Ethanolamine*HCl pH 8.5) was added for 30 minutes. Subsequently, this solution was discarded and 100 µL of Blocking buffer (25 mM HEPES, 150 mM NaCl, pH 7.5 and 1% BSA) was added for 30 minutes. Blocking buffer discard was followed by three washes using 200 µL of Washing buffer (25 mM HEPES, 150 mM NaCl, pH 7.5 and 0.1% BSA). Then, the target was incubated with 100 µL of $1.66.10^{11}$ from the TriCo-16 library (Creative Biolabs, Shirley, USA) for 30 minutes. After incubation, the biopanning step was started: five washings with a Washing buffer were performed for each well, in the first selection cycle, while in the second and third cycles, 15 washes were performed each. The objective of this step was to ensure the removal of phages that did not interact with the target and the highly interacting phages were eluted by adding 100 µL of 0.2 M Glycine-HCl pH 2.2 solution for 10 minutes. Neutralization of this low pH solution was performed with 25 µL of 1 M Tris pH 9.0, which ensured that 125 µL of the non-amplified eluate was obtained from target (TS). All steps were performed at room temperature (RT).

Figure 1. *In vitro* biopanning using peptide libraries to identify high-affinity binding to immobilized protein. The TriCo-16 random phage library contains a variety of clones with different peptides displayed on the viral capsid (indicated by phages with different colors). The library is incubated with immobilized target protein and some phages bind to the target protein through the peptides with high binding affinity. Several washes remove unbound phage and recombinant protein-binding phage are recovered by elution. The recovered phages are amplified by infecting bacterial culture. DNA sequencing of amplified phage clones is performed to determine the sequence of peptides with high affinity for the target protein.



Source: Prepared by the author.

During the first selection cycle, $1.66.10^{11}$ from the TriCo-16 library was used as an experimental control. This control was called Empty control (ES) and in the second and third cycles, this control was performed with the amplified phages of the Empty control of the first cycle (ES1) and with the amplified phages of the Empty control of the second cycle (ES2), respectively. In addition, a second control was added to the experimental steps from the second selection cycle, which was called Direct control (DC). This control was produced using the amplified phages that eluted from the wells containing the target, to evaluate, right after the sequencing step, which peptides interact only with the target, and which interact with the components of the high-binding plate.

2.3 Amplification of phages

Amplification of the selected phages (Fig.1) was performed from the *E. coli*-ER2738 in 20 ml of LB culture medium containing tetracycline (20 mg/ml), individually for each non-amplified eluate (TS1, TS2, TS3, ES1, ES2, ES3, DC2 and DC3). After reaching an OD_{600} of 0.1, 120 µL of each non-amplified eluate was added to the bacterial cultures and incubated for 4 hours under constant agitation at 37 ºC. Then, the cultures were centrifuged at 4°C at 5,000 rpm for 20 minutes and the supernatants were transferred to sterile tubes containing a 7 ml solution of 25% PEG 8,000 and 2.5 M NaCl (PEG/NaCl). The solutions were incubated overnight at 4°C to precipitate the amplified phages and then centrifuged at 5,000 rpm for 1 hour at 4°C to discard the supernatants.

The phages were then resuspended in 1 ml of selection buffer (25 mM HEPES and 150 mM NaCl, pH 7.5) and precipitated again with 200 µL of PEG/NaCl for 1 hour on ice. Subsequently, the samples were centrifuged at 5,000 rpm at 4°C and the supernatants were discarded. The precipitates were resuspended in 100 µL of the selection buffer, thus obtaining the amplified phages. The concentration of each amplified phage was measured with UV-VIS-Spektrophotometer UV-1800 (Shimadzu Deutschland GmbH) using absorbance 220 and 340 nm. Dilutions of the amplified phages were performed to ensure a concentration of $1.66.10^{11}$, to be used in the second and third cycle of selection with the target ExhC.

2.4 Phage titration

To assess whether peptides coupled to the viral capsids were interacting with the ExhC (TS) during the three phage display selection steps with the TriCo-16 library, after the elution step, 1 µL of each eluate obtained was used in titration experiments between 10^{-3} a 10^{-11}. Control experiments were also evaluated using this technique (ES and DC). Serial dilutions were incubated with 100 µl of *E. coli* - ER2738 in 0.6 OD_{600} and plated on LB medium containing IPTG (0.5 mM) and X-gal (40 µg/ml) along with 800 µL of Agarose Top. After incubation in an oven for 16 hours at 37 ºC, blue colonies were counted to evaluate experimental success.

2.5 ELISA

ELISA assays were performed with the amplified phages obtained after the elution of the first, second and third stages of selection against the target ExhC. The objective was to verify the interaction of the target with the selected peptides at each stage of the phage display selection. For this, eight wells were separated and named: TS1+, TS2+, TS3+, TS1-, TS2-, TS3-, C+ and C-. First, 30 pM of ExhC was diluted in 10 mM Sodium Acetate (pH 5.2) and added to wells with a positive sign (+), while only the diluents were added to the wells with a negative sign (-). These solutions were incubated for one hour at RT. Then, the solution containing the target was removed and 100 µL of Quenching buffer was added to each well and incubated for 30 minutes. Subsequently, these solutions were discarded and 100 µL of Blocking buffer (25 mM HEPES, 150 mM NaCl, pH 7.5 and 1% BSA) was added for 30 minutes.

Discarding the Blocking buffer was followed by three washes using 200 µL of Washing buffer (25 mM HEPES, 150 mM NaCl, pH 7.5 and 0.1% BSA) for each well. Subsequently, the wells named TS1, TS2 and TS3 were incubated for 1 hour with 100 µL of $1.66.10^{11}$ amplified phages obtained from the elution of TS1, TS2, and TS3 during the phage display. During this step,

control wells in the absence of phages (C+ and C-) were completed with 100 µL of Selection buffer. At the end of this process, five washes with 200 µL of Washing buffer were performed for each well, and then 100 µL of 1:2500 bacteriophage M13 antibody (11973-MMO5T-H) per well were incubated for 1 hour and subsequently adding 100 µL of a solution 0.1 mg/ml 3,3',5,5'-Tetramethylbenzidine (TMB), Phosphate-Citrate and 10% DMSO. Immediately, a change in color was observed due to the reaction interaction of the TMB with the M13 bacteriophage antibody. The continuity of the reaction was inhibited with the addition of 100 µL of 2M H_2SO_4. Finally, the absorbance was measured by a plate reader at 450 nm and the results were plotted using the GraphPad Prism 8 software.

2.6 DNA extraction and PCR

Ten microliters of the amplified phages (TS1, TS2, TS3, ES1, ES2, ES3, DC2 and DC3), obtained from phage display experiments with ExhC, were mixed with 90 µl of TBS 1X plus 200 µL of PEG/NaCl and incubated for 20 minutes at RT. Subsequently, the samples were centrifuged at 14,000 rpm for ten minutes at 4°C and the supernatant was completely removed. The pellet obtained from each sample was resuspended in 200 µL of 3M Sodium Acetate and 0.1 X TAE, then 500 µL of 99% ethanol was added and the samples were incubated for 15 minutes in this solution. Again, the samples were centrifuged at 14,000 rpm for ten minutes at 4°C and the supernatant was completely removed. After removing the supernatant, 250 µL of 70% ethanol was added per tube and succeeding centrifugation under the same conditions already described, it was removed. From the DNA extraction, the samples were dried at 30 °C and 40 µL of milliQ water was added to each sample to dilute the acid nucleic. The DNA concentration of each extraction was measured using the NanoPhotometer (IMPLEN).

For the amplification of the viral DNA obtained from the extraction, the PCR technique was performed. Each PCR reaction was prepared from: 12.5 µL Kapa HiFi HotStart ReadyMix (Roche), 5 µM primer reverse, 5 µM primer forward and 25 ng of DNA. All reactions were completed with milliQ water in a final volume of 25 µL. The thermal cycler used Thermal Cycler C1000 Touch (BIO-RAD) for the PCR reactions.

Subsequently performing the PCR, the concentration of each sample was measured by the NanoPhotometer (IMPLEN), and in all cases, concentrations greater than 500 ng/µL were obtained, being sufficient for performing the sequencing. Furthermore, samples were evaluated using a 1% agarose gel and next-generation sequencing (NSG) was performed by the Universität Düsseldorf (HHU).

2.7 Analysis of peptide sequences

DNA sequences resulting from NGS (Fig.1) were transcribed into peptide sequences, as previously described (Santur; Reinartz; Lien; Tusche et al., 2021). Transcribed sequences were filtered based on three parameters: their relative frequency increase within the TS (library < TS1 < TS2 < TS3) (i), the relative frequency of a sequence observed in a TS compared to its abundance within the respective DC (ii), and the relative frequency of a sequence observed in a TS compared to its abundance within the respective ES (iii). The sequences of the TS3 round that passed the filter were classified according to their enrichment factor (TS3/library = enrichment factor) and empty score (TS3/ES3 = empty score). Filtered sequences were used as inputs for the Hammock clustering software (Krejci; Hupp; Lexa; Vojtesek et al., 2016). Cluster motif logos for phage-display with ExhC were generated from initial clusters after greedy clustering using the WebLogo 3.4 application (Crooks; Hon; Chandonia; Brenner, 2004).

After analyses, seven peptides were selected for primary screening interaction assays with ExhC according to: empty score (1), enrichment factor (2) values, as well as their similarity to the original substrate porcine Dsg-1 (3) and presence of residues frequently observed in the cluster motif logos (4). Physicochemical parameters were also investigated, such as isoelectric point (5) and peptide solubility (6).

2.8 Enzimatic assays

The esterolytic activity of ExhC against the synthetic substrate Boc-L-Glu↓-OPh was determined in buffers 20 mM HEPS pH 7.4 (i) and 20 mM MES pH 6.5 (ii). First, 1 M Boc-L-Glu-↓OPhe was diluted in DMSO (Sigma) and used as a stock substrate solution. Esterolytic reactions occurred under conditions of 2% DMSO (v/v) using 10 µg/mL of toxin, varying the substrate concentration between 0 to 20 mM, with strides of 5 mM. All substrate concentrations were tested in triplicate. Phenol release (molar absorption coefficient = 1.5 mM/cm) from the synthetic substrate Boc-L-Glu-↓OPhe (λ = 270 nm) was measured for 30 minutes, every 30 seconds at a constant temperature of 37 ºC, using the Infinite 200 PRO plate reader (Tecan, Männedorf, Switzerland).

2.9 Peptide-ExhC interaction assays

The interaction assays of ExhC by the peptides were performed in two buffer conditions: 20 mM HEPES pH 7.4 (i) and 20 mM MES pH 6.5 (ii). For preliminary screening, 40 µM of peptides were used in reactions with 10 µg/mL of protein and 2.5 mM of Boc-L-Glu↓-OPh

diluted in buffer. Three incubation periods (10 minutes, 30 minutes, and 2 hours) between peptides and protein were tested. After incubation, the synthetic substrate was added to each solution and the absorbance at 270 nm was measured at 30 second intervals for 30 minutes using the Infinite 200 PRO plate reader (Tecan, Männedorf, Switzerland) at 37 °C. All inhibition assays were performed in triplicate.

2.10 BLAST analysis

Due to ExhC being a moonlight protein (Singh; Bhalla, 2020), with necrotic, exfoliative activity (Li; Wang; Ding; Zheng, 2011) and inhibition of the host's innate immune system (Li; Li; Lu; Wang et al., 2011), the objective of this in silico step was to search for potential partner proteins involved in these mechanisms of action. For this, the peptides with the highest empty score and enrichment factor values (Table 1) were evaluated using the BLAST protein tool (blast.ncbi.nlm.nih.gov/ Blast.cgi). In these analyses, only proteins from host organisms (*Sus scrofa domesticus*, *Mus musculus* and *Mesocricetus auratus*) of *M. sciuri* were selected. The proteins with the greatest identity with the peptides were investigated according to their functions and structure.

Table 1. Peptides with the highest empty score and enrichment factor values selected by phage-display performed with ExhC.

Sequence	Empty score	Enrichment factor
WAQMPQPFWKHSPFER	1253.91	1358.46
YTSPWWYYQMDEHMQG	353.55	47.88
DMQYWLGDSQQTRIS	277.53	300.67
VMPWSSNVYGYMPNYR	271.12	146.86
VSWSYWDPIHPWQKRDR	223.49	121.06

Source: Prepared by the author.

2.11 Prediction of binding modes and MD simulations

Among the peptides identified through phage display, one of them, YTSPWWYYQMDEHMQG, exhibited well-defined secondary structure based on 3D-structure predictions performed with AlphaFold2. In fact, it was determined that this peptide folds as an α-helix. Subsequently, we resorted to blind protein-peptide docking using three docking algorithms (Kozakov; Hall; Xia; Porter et al., 2017; Park; Baek; Lee; Seok, 2019; Yan; Zhang; Zhou; Li et al., 2017), i.e., ClusPro, GalaxyTongDock and HDOCK, to predict the binding modes of the peptide to ExhC. After determining the binding models by each docking algorithm, the stability of protein-peptide binding mode was investigated using MD simulations

with pmemd.cuda of Amber22, followed by MM-GBSA free-energy calculations (GBneck model, implemented in MMPBSA.py of Amber22). The same approach was followed to predict the structure of the putative ExhC-Pcyox1l complex. However, to reduce the computational demand associated with this prediction, only the docking algorithm (ClusPro) that led to the most stable binding mode of the peptide to ExhC was employed in this case. The stability of the predicted binding modes during the MD simulations of both the peptide and the partner protein were monitored through RMSD calculations with cpptraj of Amber22.

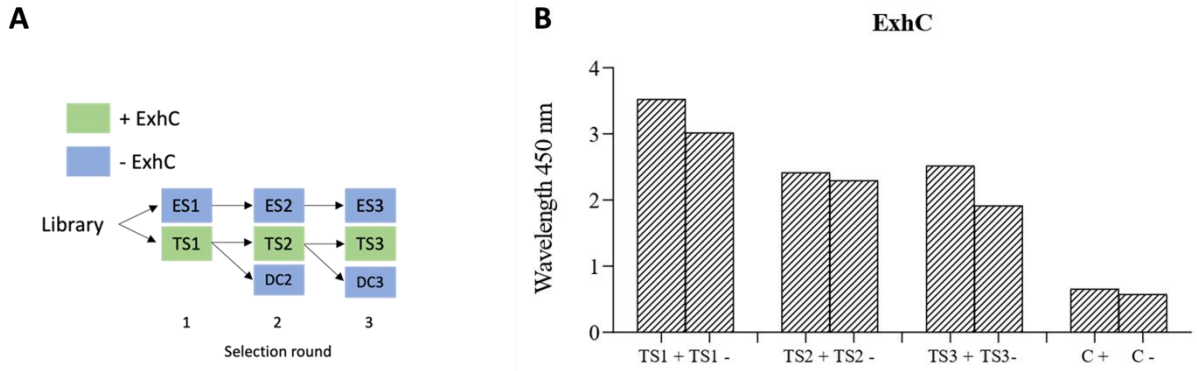
3. Results

3.1 Phage display identification of ExhC-interacting peptides

For the identification of peptides with high affinity to ExhC, phage display was performed against the target protein in three selection cycles (TS1, TS2 and TS3) (Fig 2.A). During the first selection cycle, $1.66.10^{11}$ from the TriCo-16 library was used as an experimental control. This control without ExhC, was called Empty control (ES) and in the second and third cycles, this control was performed with the amplified phages of the Empty control of the first cycle (ES1) and with the amplified phages of the Empty control of the second cycle (ES2), respectively (Fig 2.A). In addition, a second control was added to the experimental steps from the second selection cycle, which was called Direct control (DC) (Fig 2.A). This control was carried out using the amplified phages that eluted from the wells containing ExhC (Fig 2.A).

The success in the process of selecting peptides that interact specifically with ExhC was confirmed by the results obtained by ELISA (Fig 2.B). The control experiments (C+ and C-) showed a significant difference with the treated groups (Fig 2.B), indicating low interaction between the components of the high-binding plate (C-) and the targets (C+) with the M13 antibody bacteriophage, in the absence of phages. In addition, in the presence of targets (TS+) obtained by the first, second and third cycles, there was higher absorbance than in the absence of the evaluated targets (TS-), that is, there are peptides present in the viral capsids of the TriCO-16 library that interact specifically with ExhC in all the selection rounds (Fig. 2.B).

Figure 2. Phage display selection of the TriCO-16 library against ExhC in three consecutive selection rounds. **(A)** Selection scheme: library = TriCO-16 library; TS = target selection on ExhC; DC = direct control; ES = empty control. Samples were derived from each other in the direction of the arrows. For TS, ExhC was immobilized on a high-binding plate (green color), while this enzyme was not immobilized for ES and DC (blue color). **(B)** ELISA assays were performed with phages selected during the first (TS1), second (TS2) and third (TS3) selection cycles of ExhC phage display with the TriCO-16 Library (New England Biolabs). Results in the absence of phages were generated as controls (C). The signs (+) and (-) represent the presence and absence of ExhC, respectively, during the ELISA assays.

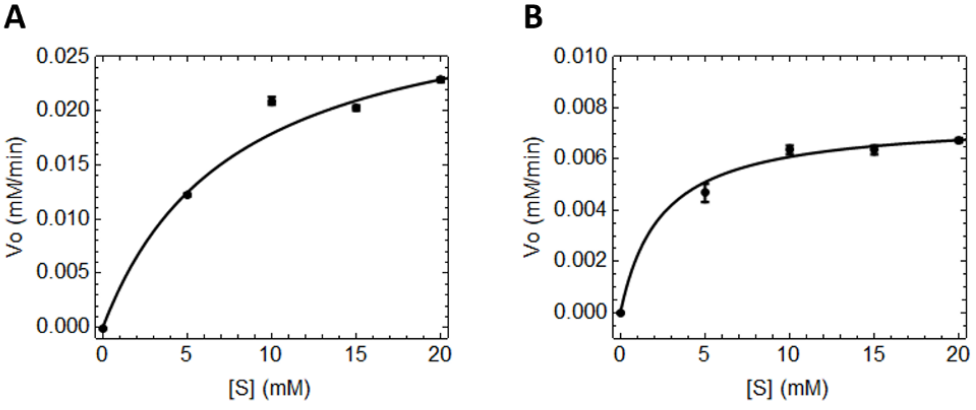


Source: Prepared by the author.

3.2 Peptide-ExhC interaction assays

The esterolytic activity of ExhC against the synthetic substrate Boc-L-Glu↓-OPh was determined in buffers 50 mM HEPES pH 7.4 (i) and 20 mM MES pH 6.5 (ii) with 2% DMSO (v/v). In both the cases, the curve of the enzymatic reaction at different substrate concentrations follows the characteristic Michaelis–Menten hyperbola (Fig. 3) and we employed the Michaelis-Menten equation to estimate the enzymatic kinetic parameters of ExhC in both buffers (Table 2).

Figure 3. Hydrolysis of Boc-L-Glu↓-OPh by ExhC in buffer 50 mM HEPES pH 7.4 **(A)** and 20 mM MES pH 6.5 **(B)** using the enzyme kinetics model by Michaelis-Menten.



Source: Prepared by the author.

Table 2. Kinetic parameters for the hydrolysis of Boc-L-Glu-↓OPhe by ExhC in buffer 50 mM HEPES pH 7.4, and 20 mM MES pH 6.5.

Buffer	KM (mM)	Vmax (mM/min)	Kcat/KM ($mM^{-1} \cdot s^{-1}$)
50 mM HEPES pH 7.4	7.7 (2.0)	0.032 (0.004)	2.0 (0.6)
20 mM MES pH 6.5	2.4 (0.8)	0.0075 (0.0004)	1.5 (0.5)

Source: Prepared by the author.

After determining the kinetic parameters of ExhC in buffers 20 mM HEPES pH 7.4 (i) and 20 mM MES pH 6.5 (ii), interaction assays in the presence of the peptides (Table 3) were performed as described in Materials and Methods. The results indicated that peptides 1, 2, 4, 6 and 7 did not cause significant inhibition of the enzimatic activity of ExhC, while 40 µM of peptides 3 and 5 caused approximately 30% inhibition in 20 mM MES buffer pH 6.5 (Table 4) and 50 mM HEPES pH 7.4, respectively (Table 5). In contrast, assays with 40 µM of peptides 2 and 4 (20 mM MES buffer pH 6.5) incubated with ExhC for two hours increased the enzymatic activity of the toxin by 34.43% and 25.30%, respectively (Table 4).

Table 3. Peptides selected for primary screening interaction assays with ExhC. Characteristics such as isoelectric point (pI), solubidity, empty score and enrichment factor were evaluated.

ID peptide	Sequence	Empty score	Enrichment score	pI	Solubility
1	WAQMPQPFWKHSPFER	1253.91	1358.46	9.99	Poor water solubility
2	YTSPWWYYQMDEHMQG	353.55	47.88	4.35	Poor water solubility
3	DMQYWLGDSQQTRIS	277.53	300.67	3.71	Good water solubility
4	VMPWSSNVYGYMPNYR	271.12	146.86	8.47	Poor water solubility
5	WENQHWEGQGRRWEWR	172.2	46.64	7.7	Good water solubility
6	MPHHYVRSHQWPPVGQ	130.06	70.45	9.59	Poor water solubility
7	MLSEAPHHKWHVKGHE	102.58	13.89	7.86	Good water solubility

Source: Prepared by the author.

Table 4. Results of primary screening interaction assays with ExhC in the presence of peptides selected by phage display. The assays were performed in 20 mM MES buffer, pH 6.5. The numbering of the peptides refers to the numbers reported in Table 3. The results are shown as mean ± standard deviation of the data obtained after a 10-min, 30-min and 2 hours incubation of peptides with the protein.

ID peptide	10 min of incubation	30 min of incubation	2 hours of incubation
1	116.53 ± 3.88	108.23 ± 5.52	102.85 ± 6.72
2	119.10 ± 34.61	111.50 ± 4.46	134.43 ± 9.62
3	84.31 ± 5.75	108.75 ± 5.81	69.63 ± 4.75
4	106.31 ± 10.61	97.35 ± 6.58	125.30 ± 6.04

5	101.48 ± 3.55	99.08± 3.95	82.90 ± 16.10
6	113.58 ± 3.64	105.97 ± 7.71	113.86 ± 5.77
7	99.17 ± 5.26	107.02 ± 5.64	82.07 ± 2.21

Source: Prepared by the author.

Table 5. Results of primary screening interaction assays with ExhC in the presence of peptides selected by phage display. The assays were performed in 20 mM HEPES buffer, pH 7.4. The numbering of the peptides refers to the numbers reported in Table 3. The results are shown as mean ± standard deviation of the data obtained after a 10-min, 30-min and 2 hours incubation of peptides with the protein.

ID peptide	10 min of incubation	30 min of incubation	2 hours of incubation
1	98.34 ± 0.51	82.98 ± 18.77	100 ± 2.2
2	115.14 ± 4.42	104.30 ± 6.60	86.12 ± 23.80
3	97.19 ± 2.88	86.84 ± 17.32	100 ± 30.34
4	113.48 ± 6.21	92.32 ± 5.76	104.11 ± 8.21
5	68.87 ± 13.80	94.37 ± 15.20	100 ± 8.60
6	79.71 ± 8.57	100 ± 0.71	100 ± 3.5
7	100 ± 6.08	84.5 ± 2.11	100 ± 2.2

Source: Prepared by the author.

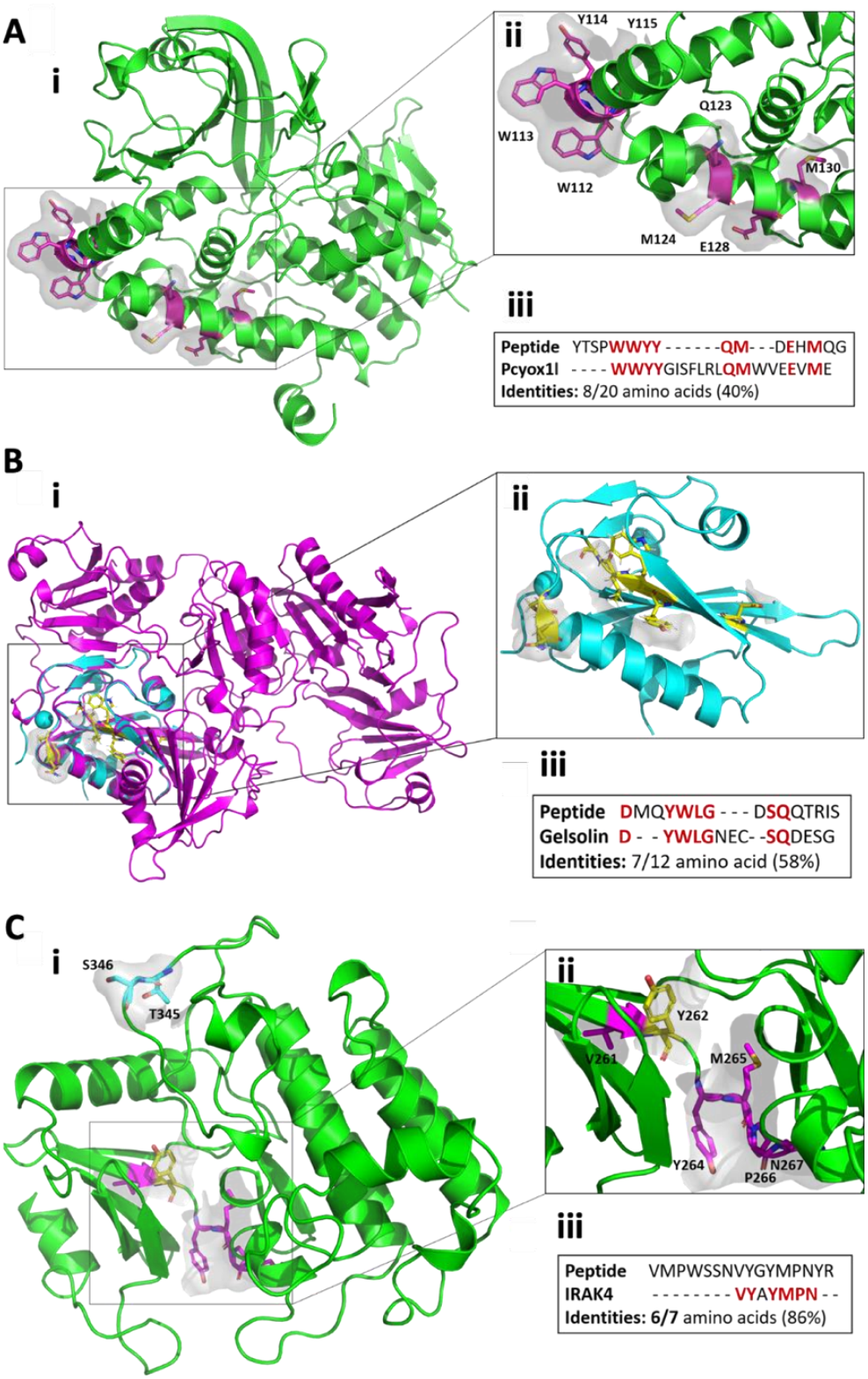
3.3 Search for ExhC partner proteins

Among the proteins identified by BLAST as having a higher degree of identity with the peptide targets (Table 1), three particularly distinguished themselves due to their cellular functions and three-dimensional structures: Prenylcysteine oxidase-1-like protein (Pcyox1l) (a), Gelsolin (b), and Interleukin-1 receptor-associated kinase 4 (c) from *Mus musculus* (Fig. 4). These proteins presented the amino acids of identity with peptides (Table 1) exposed to the solvent in the surface area (Fig. 4), which allows the interaction of these residues with ExhC secreted by *M. sciuri*.

Pcyox1l showed 40% identity with the peptide YTSPWWYYQMDEHMQG (Fig 3.A). Docking algorithms ClusPro (clpro), GalaxyTongDock (glx) and HDOCK (hdock) (Kozakov; Hall; Xia; Porter et al., 2017; Park; Baek; Lee; Deok, 2019; Yan; Zhang; Zhou; Li et al., 2017) predicted two poses with the highest scores for the binding of this peptide with ExhC (Fig 5.A). The pose predicted by the clpro and glx algorithms was denoted as site 1 in ExhC, showcasing a disparity from the hdock-predicted pose, which aligns with the N-terminal α-helix of the toxin (Fig 5.A). However, subsequent Molecular Dynamics (MD) simulations revealed the instability

of the hdock binding mode, whereas those predicted by clpro and glx exhibited relative stability (Figure 5.B). Ultimately, free energy calculations underscored the structural prediction by clpro as the most stable among the three configurations (Fig 5.C and 5.D).

Figure 4. (A) (i) 3D-structure of Pcyox1l from *M. musculus* generated by AlphaFold. **(ii)** Highlight of the region containing residues W112, W113, Y114, Y115, Q123, M124, E128 and M13 in sticks (magenta) and surface (gray). **(iii)** Blast analysis of murine Pcyox1l (Q8NBM8) and YTSPWWYYQMDEHMQG peptide sequence reveals 40% sequence identity. **(B) (i)** 3D structure of Gelsolin from *Mus musculus* generated by Alphafold (magenta) overlap with the structure of the G6 domain (PDB ID: 6I4D) (cyan). **(ii)** 3D structure of the G6 domain of Gelsolin from *Mus musculus* with highlights on sticks (yellow) and surface (gray) for residues D84, Y87, W88, L89, G90, S94 and Q95. The sphere represents a Ca^{2+} ion. **(iii)** Blast analysis of murine Gesolin (P13020) and DMQYWLGDSQQTRIS peptide sequence reveals 58% sequence identity. **(C) (i)** 3D structure of the central kinase domain of IRAK4 from *Mus musculus* generated by Alphafold. On sticks, the active site residues (cyan) and residues V261, Y262, Y264, M265, P266 and N267 (magenta and yellow). The surface area has been highlighted in gray. **(ii)** Highlight the region containing residues V261, Y262, Y264, M265, P266 and N267 (magenta and yellow) on sticks (magenta and yellow) and surface (gray). **(iii)** Blast analysis of murine IRAK4 (Q8R4K2) and VMPWSSNVYGYMPNYR peptide sequence reveals 86% sequence identity.



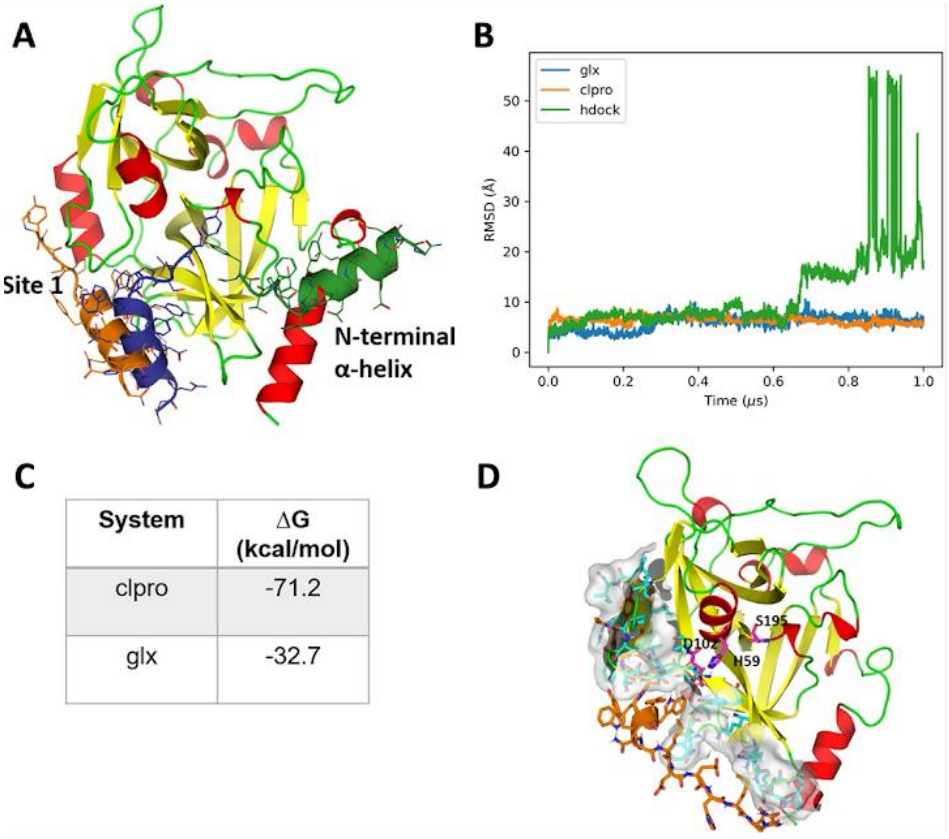
Source: Prepared by the author.

The binding mode of the putative partner protein to ExhC was determined by docking using the ClusPro algorithm. The best pose was subjected to 1 μs MD simulation, with the representative structure of the simulation being shown in Fig. 6.A. The segment of Pcyox1l matching the

peptide sequence interacts with the site 1 residues of ExhC, as predicted also for the peptide and the latter protein (Fig. 6.A). Fig. 6.B shows the evolution of RMSD values for the protein backbone atoms along the simulation time. The values remained relatively small, which indicates that the initial structure predicted by docking is stable.

The MM-GBSA calculations carried out for the ExhC-Pcyox1l complex yielded an estimated binding free energy of -66.4 kcal/mol, similar to the one calculated for the toxin and the matching peptide. Interestingly, the results of the per-residue decomposition analysis, which predicts the energy contribution of each residue to the binding, identified the residues W112 and W113 as the main hot spots of the protein-protein interface (Table 6). This result underscores the importance of two of the matching residues to the binding of the putative partner protein to ExhC.

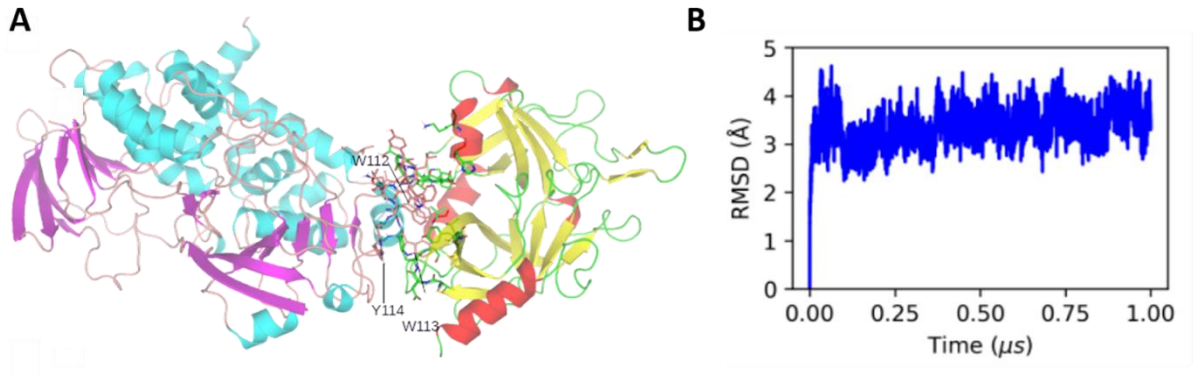
Figure 5. Prediction of the binding mode of the peptide YTSPWWYYQMDEHMQG to ExhC. **(A)** top-scoring poses predicted with clpro (orange), glx (blue) and hdock (green). **(B)** RMSD of the peptide's heavy atoms during the MD simulations, calculated after fitting all trajectory frames onto the backbone atoms of ExhC in the initial structure. **(C)** MM-GBSA free energy values for the indicated structures. **(D)** Representative structure of the peptide in complex with ExhC. The surface of the binding site is shown in gray and the interacting protein residues are depicted as cyan sticks.



System	ΔG (kcal/mol)
clpro	-71.2
glx	-32.7

Source: Prepared by the author.

Figure 6. Prediction of the binding mode of the Pcyox1l to ExhC. **(A)** Depiction of the representative structure of the complex obtained from the MD trajectory. Residues labeled correspond to those matching the peptide. **(B)** RMSD time profile for the backbone atoms.



Source: Prepared by the author.

Table 6. Per-reside free-energy contributions at the interface of the ExhC-Pcyox1l complex.

ExhC	**Pcyox-1l**	**$\Delta G_{res,tot}$[a] (kcal/mol)**	**$\Delta G_{res,sc}$[b] (kcal/mol)**
G115	-	-2.9	0
V116	-	-4.4	-3.9
F179	-	-6.5	-6.1
S216	-	-5.4	-3.4
-	L105	-3.3	-4.5
-	L106	-4.7	-3.2
-	L108	-4.1	-2.2
-	F109	-5.2	-4.6
-	R110	-2.4	-2.1
-	**W112**	-5.3	**-5.3**
-	**W113**	-6.3	**-6.3**
-	I117	-2.4	-2.4
-	L333	-4.5	-4.5

[a] Total per-residue free-energy contribution. Only residues contributing more than -2 kcal/mol are shown. [b]Side-chain per-residue free-energy contribution.

Source: Prepared by the author.

4. Discussion

ExhC, a protein that causes the exfoliation of the epidermis, is involved in the dissemination and pathogenicity of *M. sciuri* (Chen; Wang; Chen; Yang et al., 2007; Tanabe; Sato; Sato; Watanabe et al., 1996). It also holds additional importance due to its necrotic activity (Li; Wang; Ding; Zheng, 2011) and partial inhibition of the immune system (Li; Li; Lu; Wang et al., 2011). The discovery of these multiple functions, which confirm that ExhC behaves as a moonlighting protein, was extremely relevant in the clinical area. However, the search for potential inhibitors for ExhC and the cellular targets affected by this enzyme remains an unexplored endeavor.

Using the phage display technique, seven peptides were tested in the presence of the synthetic substrate and two of them partially inhibited the esterolytic activity of ExhC. These results indicate that the peptides DMQYWLGDSQQTRIS and WENQHWEGQGRRWEWR are potential inhibitors for the exfoliative activity of ExhC. In contrast, the peptides YTSPWWYYQMDEHMQG and VMPWSSNVYGYMPNYR cause an increase in the enzymatic activity of ExhC. This activation potential can also be explored to understand the effects of the interaction of the peptide with the enzyme and the influence of these conformational changes on the exfoliative activity.

Based on the sequence of one of the peptides identified through phage display, it was suggested the interaction between ExhC and Pcyox1l. Pcyox1l, produced by *Mus musculus*, is responsible for maintaining the cell membrane, and its absence can lead to a decrease in cell viability (Petenkova; Auger; Lamb; Quellier et al., 2023). Additionally, Pcyox1l CRISPRed-out neutrophils exhibit deficient bactericidal properties (Petenkova; Auger; Lamb; Quellier et al., 2023). Thus, the stability of the interaction observed between ExhC and Pcyox1l in *in silico* analyses could explain the ability of ExhC to induce cell death (Li; Wang; Ding; Zheng, 2011) and inhibit the phagocytosis of *M. sciuri* by RAW264.7 macrophages (Li; Li; Lu; Wang et al., 2011). However, in vitro tests are still necessary to conclusively support these hypotheses.

Finally, although three of the seven selected peptides do not cause an effect in the esterolytic activity of the enzyme, it is still possible that they influence the mechanisms of necrotic action and/or inhibition of the host's innate immune system. To evaluate these properties, assays with the cell lines BHK-21(Li; Wang; Ding; Zheng, 2011) and RAW264.7 (Li; Li; Lu; Wang et al., 2011) in the presence of peptides and ExhC need to be performed.

5. Conclusion

Through the discovery of peptides with high affinity to ExhC, together with *in vitro* enzymatic assays and *in silico* analyses, we discovered inhibitors and activators for the esterolytic activity of this toxin. Additionally, we uncovered proteins that can interact with ExhC, potentially triggering necrosis and/or inhibition of the immune system. This knowledge holds significance for optimizing inhibitory molecules and expanding our understanding of the role played by ExhC in hosts of *M. sciuri*.

6. References

ADDEPALLI, B.; RAO, S.; HUNT, A. G. Phage display library screening for identification of interacting protein partners. **Polyadenylation in Plants: Methods and Protocols**, p. 147-158, 2015.

AHRENS, P.; ANDRESEN, L. O. Cloning and sequence analysis of genes encoding *Staphylococcus hyicus* exfoliative toxin types A, B, C, and D. **J Bacteriol**, 186, n. 6, p. 1833-1837, Mar 2004.

BUKOWSKI, M.; WLADYKA, B.; DUBIN, G. Exfoliative toxins of *Staphylococcus aureus*. **Toxins (Basel)**, 2, n. 5, p. 1148-1165, May 2010.

BURGESS, R. R. A brief practical review of size exclusion chromatography: Rules of thumb, limitations, and troubleshooting. **Protein Expr Purif**, 150, p. 81-85, Oct 2018.

CHEN, S.; WANG, Y.; CHEN, F.; YANG, H. *et al.* A highly pathogenic strain of *Staphylococcus sciuri* caused fatal exudative epidermitis in piglets. **PLoS One**, 2, n. 1, p. e147, Jan 10 2007.

CROOKS, G. E.; HON, G.; CHANDONIA, J. M.; BRENNER, S. E. WebLogo: a sequence logo generator. **Genome research**, v. 14, n. 6, p. 1188-1190, 2004.

DEMIDYUK, I. V.; CHUKHONTSEVA, K. N.; KOSTROV, S. V. Glutamyl Endopeptidases: The Puzzle of Substrate Specificity. **Acta Naturae**, 9, n. 2, p. 17-33, Apr-Jun 2017.

GISMENE, C.; GONZALEZ, J. E. H.; DE FREITAS CALMON, M.; NASCIMENTO, A. F. Z. *et al.* Necrotic activity of ExhC from *Mammaliicoccus sciuri* is mediated by specific amino acid residues. **Int J Biol Macromol**, 254, n. Pt 1, p. 127741, Jan 2024.

IMANISHI, I.; NICOLAS, A.; CAETANO, A. B.; CASTRO, T. L. P. *et al.* Exfoliative toxin E, a new *Staphylococcus aureus* virulence factor with host-specific activity. **Sci Rep**, 9, n. 1, p. 16336, Nov 8 2019.

JEFFERY, C. J. Why study moonlighting proteins? **Front Genet**, 6, p. 211, 2015.

KOZAKOV, D.; HALL, D. R.; XIA, B.; PORTER, K. A. *et al.* The ClusPro web server for protein-protein docking. **Nat Protoc**, 12, n. 2, p. 255-278, Feb 2017.

KREJCI, A.; HUPP, T. R.; LEXA, M.; VOJTESEK, B. *et al.* Hammock: a hidden Markov model-based peptide clustering algorithm to identify protein-interaction consensus motifs in large datasets. **Bioinformatics**, 32, n. 1, p. 9-16, Jan 1 2016.

LI, H.; LI, X.; LU, Y.; WANG, X. *et al. Staphylococcus sciuri* exfoliative toxin C is a dimer that modulates macrophage functions. **Can J Microbiol**, 57, n. 9, p. 722-729, Sep 2011.

LI, H.; WANG, Y.; DING, L.; ZHENG, S. J. *Staphylococcus sciuri* exfoliative toxin C (ExhC) is a necrosis-inducer for mammalian cells. **PLoS One**, 6, n. 7, p. e23145, 2011.

MADHAIYAN, M.; WIRTH, J. S.; SARAVANAN, V. S. Phylogenomic analyses of the Staphylococcaceae family suggest the reclassification of five species within the genus Staphylococcus as heterotypic synonyms, the promotion of five subspecies to novel species, the taxonomic reassignment of five Staphylococcus species to Mammaliicoccus gen. nov., and the formal assignment of Nosocomiicoccus to the family Staphylococcaceae. **Int J Syst Evol Microbiol**, 70, n. 11, p. 5926-5936, Nov 2020.

MARIUTTI, R. B.; TARTAGLIA, N. R.; SEYFFERT, N.; DE PAULA CASTRO, T. L. *et al.* Exfoliative toxins of *Staphylococcus aureus*. **J The Rise of Virulence Antibiotic Resistance in Staphylococcus aureus**, 2017.

NEMEGHAIRE, S.; ARGUDIN, M. A.; FESSLER, A. T.; HAUSCHILD, T. *et al.* The ecological importance of the *Staphylococcus sciuri* species group as a reservoir for resistance and virulence genes. **Vet Microbiol**, 171, n. 3-4, p. 342-356, Jul 16 2014.

PARK, T.; BAEK, M.; LEE, H.; SEOK, C. GalaxyTongDock: Symmetric and asymmetric ab initio protein-protein docking web server with improved energy parameters. **J Comput Chem**, 40, n. 27, p. 2413-2417, Oct 15 2019.

PETENKOVA, A.; AUGER, S. A.; LAMB, J.; QUELLIER, D. *et al.* Prenylcysteine oxidase 1 like protein is required for neutrophil bactericidal activities. **Nat Commun**, 14, n. 1, p. 2761, May 13 2023.

PETSALAKI, E.; Stark, A.; García-Urdiales, E.; Russell, R. B. Accurate prediction of peptide binding sites on protein surfaces. **PLoS computational biology**, v. 5, n. 3, p. e1000335, 2009.

PETERS, K.; BRAUN, H. P. Comparative analyses of protein complexes by blue native DIGE. **Methods Mol Biol**, 854, p. 145-154, 2012.

RODRIGUEZ, E. L.; PODDAR, S.; IFTEKHAR, S.; SUH, K. *et al.* Affinity chromatography: A review of trends and developments over the past 50 years. **J Chromatogr B Analyt Technol Biomed Life Sci**, 1157, p. 122332, Nov 10 2020.

ROSANO, G. L.; CECCARELLI, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. **J Front Microbiol,**, 5, p. 172, 2014.

SANTUR, K.; REINARTZ, E.; LIEN, Y.; TUSCHE, M. *et al.* Ligand-Induced Stabilization of the Native Human Superoxide Dismutase 1. **ACS Chem Neurosci**, 12, n. 13, p. 2520-2528, Jul 7 2021.

SATO, H.; WATANABE, T.; HIGUCHI, K.; TERUYA, K. *et al.* Chromosomal and extrachromosomal synthesis of exfoliative toxin from *Staphylococcus hyicus*. **J Bacteriol**, 182, n. 14, p. 4096-4100, Jul 2000.

SINGH, N.; BHALLA, N. Moonlighting Proteins. **Annu Rev Genet**, 54, p. 265-285, Nov 23 2020.

TANABE, T.; SATO, H.; SATO, H.; WATANABE, K. *et al.* Correlation between occurrence of exudative epidermitis and exfoliative toxin-producing ability of *Staphylococcus hyicus*. **Vet Microbiol**, 48, n. 1-2, p. 9-17, Jan 1996.

WEGENER, H. C.; ANDRESEN, L. O.; BILLE-HANSEN, V. *Staphylococcus hyicus* virulence in relation to exudative epidermitis in pigs. **Can J Vet Res**, 57, n. 2, p. 119-125, Apr 1993.

YAN, Y.; ZHANG, D.; ZHOU, P.; LI, B. *et al.* HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. **Nucleic Acids Res**, 45, n. W1, p. W365-W373, Jul 3 2017.

Capítulo V. Conclusões gerais

Neste trabalho foram exploradas técnicas experimentais e computacionais com o objetivo de compreender as propriedades estereoquímicas da ExhC responsáveis pelo potencial necrótico dessa enzima. O alinhamento primário da ExhC com as demais ETs já descritas na literatura, em conjunto com o modelo 3D da toxina, possibilitou que identificássemos quatro resíduos de aminoácidos potencialmente envolvidos nesse mecanismo de ação (Gismene; Gonzalez; De Freitas Calmon; Nascimento et al., 2024).

A mutação dos resíduos R47, N49, Q51 e R89 da ExhC ($ExhC_{mut4}$) garantiu a expressão de uma proteína sem potencial necrótico que manteve níveis de atividade enzimática similares ao da ExhC selvagem ($ExhC_{wt}$) (Gismene; Gonzalez; De Freitas Calmon; Nascimento et al., 2024). Esses dados são indicativos de que as atividades necrótica e catalítica são independentes e envolvem sítios ativos distintos. Mutações pontuais independentes desses quatro resíduos de aminoácidos indicaram que os resíduos R47, N49 e Q51 são individualmente cruciais para a atividade necrótica da ExhC de *M. sciuri*, enquanto o resíduo R87 é menos relevante para a morte celular da BHK-21, mas ainda interfere significativamente na eficácia desse mecanismo de ação (Gismene; Gonzalez; De Freitas Calmon; Nascimento et al., 2024).

Além disso, a resolução das estruturas cristalográficas de $ExhC_{wt}$ e $ExhC_{mut4}$ revelou que os quatro resíduos principais submetidos à mutação se encontram na superfície formando uma cavidade com alta densidade de cargas positivas (Gismene; Gonzalez; De Freitas Calmon; Nascimento et al., 2024). Ademais, a variante mutada apresentou diferenças marcantes em termos de potencial eletrostático e hidrofobicidade dos aminoácidos da região implicada na necrose, sugerindo que as mutações podem interferir no reconhecimento e ligação de uma molécula efetora responsável pela morte celular (Gismene; Gonzalez; De Freitas Calmon; Nascimento et al., 2024).

A propriedade homodimérica da ExhC descoberta por Li et al. (2011b), também foi explorada durante a realização deste trabalho. Observou-se novamente que a ExhC é um dímero em solução, propriedade atípica em serino proteinases convencionais, tais como a ß-tripsina, que comumente atuam como monômeros na natureza (Palmer; Bonner, 2007). Curiosamente, as mutações aqui conduzidas para esta enzima, que afetaram drasticamente o seu potencial necrótico, não alteraram o seu estado oligomérico (Gismene; Gonzalez; De Freitas Calmon; Nascimento et al., 2024).

Os resultados referentes ao homodimerização da ExhC de *M. sciuri* tornaram-se *insights* para avaliação do estado oligomérico das demais toxinas esfoliativas (ETs) produzidas por bactérias do gênero Staphylococcus. Portanto, realizou-se experimentos de BN-PAGE, SAXS

e SEC que confirmaram que a ETE (Gismene; Gonzalez; Rocio Niño Santisteban; Nascimento et al., 2022) e ETA (dados não mostrados) de *S. aureus* também formam dímeros em solução. O conjunto de dados obtidos a partir da ExhC, ETA e ETE tornaram-se indicativos de que a homodimerização poderia ser uma propriedade generalista entre as ETs, provavelmente desempenhando papel funcional (Gismene; Gonzalez; Rocio Niño Santisteban; Nascimento et al., 2022).

As ETs são enzimas da família das serino proteinases semelhantes à quimotripsina (CLSPs), todavia são consideradas atípicas pois não exibem atividade proteolítica contra substratos de amplo espectro das CLPs, sendo altamente específicas a clivagem da Dsg-1 (Demidyuk; Chukhontseva; Kostrov, 2017). Essa especificidade foi atribuída ao longo de anos por duas características estruturais únicas: (i) a presença de α-hélice N-terminal carregada e (ii) ligação peptídica P192-G193 invertida em 180° (Cavarelli; Prevost; Bourguet; Moulinier et al., 1997; Demidyuk; Chukhontseva; Kostrov, 2017; Papageorgiou; Plano; Collins; Acharya, 2000; Vath; Earhart; Rago; Kim et al., 1997). Em teoria, essa última característica dificultaria o acesso do substrato ao *oxyanion hole* ° (Vath; Earhart; Rago; Kim et al., 1997). A partir da construção de variante mutante ETE/P186G com atividade catalítica similar a ETE selvagem, amenizamos a influência dessa propriedade na atividade enzimática da ETE, e atribuímos maior relevância funcional ao estado dimérico dessa toxina (Gismene; Gonzalez; Rocio Niño Santisteban; Nascimento et al., 2022).

A análise computacional para estrutura homodimérica prevista da ETE de *S. aureus* revelou a participação de uma alça (L1), não conservada em outras CLPs (Gismene; Gonzalez; Rocio Niño Santisteban; Nascimento et al., 2022). Essa alça é aparentemente crucial para a formação da interface proteína-proteína, além de causar a oclusão parcial do sítio ativo da ETE, sugerindo que a monomerização seria uma etapa necessária para sua atividade catalítica (Gismene; Gonzalez; Rocio Niño Santisteban; Nascimento et al., 2022). Ainda, os dados reforçam que a conformação dimérica aqui proposta poderia se dissociar facilmente sob determinadas condições, evidenciando novamente que a monomeriazação pode influenciar na ativação enzimática e clivagem da Dsg-1 (Gismene; Gonzalez; Rocio Niño Santisteban; Nascimento et al., 2022).

Apesar dos achados sobre o estado oligomérico das ETs, e sua potencial influência da atividade esfoliativa dessas enzimas, os resultados obtidos em avaliação da atividade necrótica da ExhC, indicaram que os resíduos R47, N49, Q51 e R89 mutados não são cruciais para formação do dímero e sugerem que a presença da forma monomérica não é essencial para a

interação com o(s) alvo(s) celular(es) que causa(m) necrose (Gismene; Gonzalez; De Freitas Calmon; Nascimento et al., 2024).

Apesar dos avanços na compreensão de como ExhC atua, os alvos celulares envolvidos nesse processo ainda não foram identificados, assim como não há descrição de potenciais inibidores para o mecanismo de ação necrótico e esfoliativo. Através da aplicação da técnica de phage display, além de inibidores e ativadores que modulam a atividade enzimática da ExhC, identificamos a Pcyox1l como potencial proteína parceira do hospedeiro murino com alta afinidade por essa toxina. A provável interação entre ExhC-Pcyox1l explicaria o potencial necrótico da ExhC, visto que a Pcyox1l está envolvida na manutenção da membrana celular de células de *Mus musculus* (Petenkova; Auger; Lamb; Quellier et al., 2023). Entretanto, ensaios de interação *in vitro* entre ExhC-Pcyox1l ainda deverão ser concluídos para confirmar essa hipótese.

Em conclusão, os resultados apresentados na tese para obtenção do título de Doutorado pela aluna Carolina Gismene, contribuíram para a compreensão do mecanismo de ação necrótico e esfoliativo da ExhC de *M. sciuri*. Além disso, os dados estruturais obtidos da ETE de *S. aureus* proporcionaram conhecimento para a elaboração novas hipóteses sobre a atividade esfoliativa das ETs. Desvendar detalhadamente os eventos moleculares envolvidos nos processos necróticos e esfoliativos em nível atômico tornou-se uma etapa importante para projetar inibidores altamente específicos que possam ser usados clinicamente e para melhorar o diagnóstico envolvendo os patógenos que codificam essas toxinas.

Referências

CAVARELLI, J.; PREVOST, G.; BOURGUET, W.; MOULINIER, L. *et al.* The structure of *Staphylococcus aureus* epidermolytic toxin A, an atypic serine protease, at 1.7 A resolution. **Structure**, 5, n. 6, p. 813-824, Jun 15 1997.

DEMIDYUK, I. V.; CHUKHONTSEVA, K. N.; KOSTROV, S. V. Glutamyl Endopeptidases: The Puzzle of Substrate Specificity. **Acta Naturae**, 9, n. 2, p. 17-33, Apr-Jun 2017.

GISMENE, C.; GONZALEZ, J. E. H.; DE FREITAS CALMON, M.; NASCIMENTO, A. F. Z. *et al.* Necrotic activity of ExhC from *Mammaliicoccus sciuri* is mediated by specific amino acid residues. **Int J Biol Macromol**, 254, n. Pt 1, p. 127741, Jan 2024.

GISMENE, C.; HERNANDEZ GONZALEZ, J. E.; SANTISTEBAN, A. R. N.; ZIEM NASCIMENTO, A. F.; DOS SANTOS CUNHA, L.; DE MORAES, F. R.; DE OLIVEIRA, C. L. P.; OLIVEIRA, C. C; JOCELAN SCARIN PROVAZZI, P.; PASCUTTI, G.; ARNI, R.K.; BARROS MARIUTTI, R. *Staphylococcus aureus* Exfoliative Toxin E, Oligomeric State and Flip of P186: Implications for Its Action Mechanism. **Int J Mol Sci**, v. 23, n. 17, p. 9857, 2022.

LI, H.; LI, X.; LU, Y.; WANG, X. *et al. Staphylococcus sciuri* exfoliative toxin C is a dimer that modulates macrophage functions. **Can J Microbiol**, 57, n. 9, p. 722-729, Sep 2011.

PAPAGEORGIOU, A. C.; PLANO, L. R.; COLLINS, C. M.; ACHARYA, K. R. Structural similarities and differences in *Staphylococcus aureus* exfoliative toxins A and B as revealed by their crystal structures. **Protein Sci**, 9, n. 3, p. 610-618, Mar 2000.

PETENKOVA, A.; AUGER, S. A.; LAMB, J.; QUELLIER, D. *et al.* Prenylcysteine oxidase 1 like protein is required for neutrophil bactericidal activities. **Nat Commun**, 14, n. 1, p. 2761, May 13 2023.

VATH, G. M.; EARHART, C. A.; RAGO, J. V.; KIM, M. H. *et al.* The structure of the superantigen exfoliative toxin A suggests a novel regulation as a serine protease. **J Biochemistry**, 36, n. 7, p. 1559-1566, 1997.

.

Apêndices

Appendix A - Supplementary material for the scientific article entitled Necrotic activity of ExhC from *Mammaliicoccus sciuri* is mediated by specific amino acid residues

Necrotic activity of ExhC from *Mammaliicoccus sciuri* is mediated by specific amino acid residues

Carolina Gismene[a]; Jorge Enrique Hernández González[a]; Marília de Freitas Calmon[b]; Andrey Fabricio Ziem Nascimento[c]; Angela Rocio Niño Santisteban[a]; Felipe Antunes Calil[a]; Alana Della Torre da Silva[d]; Paula Rahal[b], Rejane Maira Góes[d]; Raghuvir Krishnaswamy Arni[a]; Ricardo Barros Mariutti[a]

[a]Multiuser Center for Biomolecular Innovation, São Paulo State University - UNESP, São José do Rio Preto, SP, Brazil.

[b]Laboratory of Genomic Studies, São Paulo State University - UNESP, São José do Rio Preto, SP, Brazil.

[c]Brazilian Synchrotron Light Laboratory (LNLS), Brazilian Center for Research in Energy and Materials (CNPEM), Campinas, SP, Brazil.

[d]Departament of Biological Sciences, São Paulo State University - UNESP, São José do Rio Preto, SP, Brazil.

Supplementary material

Table S1. Effects of progressively increasing concentration of $ExhC_{wt}$ and its mutant variants on the viability of the cell line BHK-21 by MTT assays. Results represent three independent experiments presented as mean ± SD.

Cell Viability (%)						
Protein concentration	$ExhC_{wt}$	$ExhC_{mut4}$	$ExhC_{R47A}$	$ExhC_{N49F}$	$ExhC_{Q51K}$	$ExhC_{R89A}$
60 µM	28.6 ± 3.1	92.8 ± 17	79.5 ± 6,2	92.6 ± 12.2	77.9±1.8	66.1±9.1
30 µM	42.5 ±12.8	94.7 ± 14	85.4 ± 6.2	97.4 ± 10.4	102.5±13.9	80.4±16.0
15 µM	58.8 ± 9.4	95.7 ± 8	101.2 ±3.7	92.5 ± 15.5	96.3±5.7	81.4±11.5
PBS 1X	96.4 ± 4.8	95.0 ± 1.4	98.8 ± 7.2	85.3± 6.5	94.3±11.4	109.3±10.1

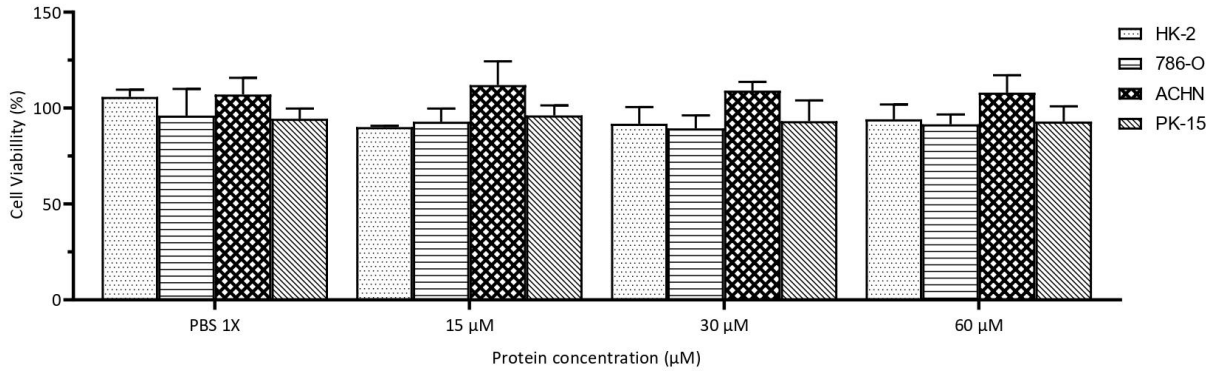


Figure S1. Treatment with $ExhC_{wt}$ at concentrations of 15, 30 and 60 µM in the cell lines HK-2, 786-O, ACHN and PK-15. The plotted results represent three independent MTT experiments shown as mean ± SD. Statistically significant differences between the control experiment and treated samples were calculated using Dunnett's test. In all cases not significant results were obtained.

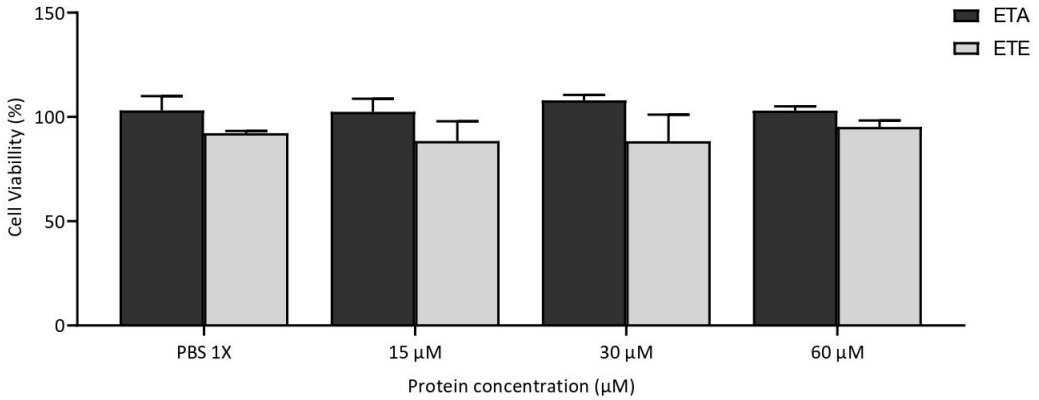


Figure S2. Treatment with ETA and ETE at concentrations of 15, 30 and 60 µM in the cell line BHK-21. The plotted results represent three independent MTT experiments shown as mean ± SD. Statistically significant differences between the control experiment and treated samples were calculated using Dunnett's test. In all cases not significant results were obtained.

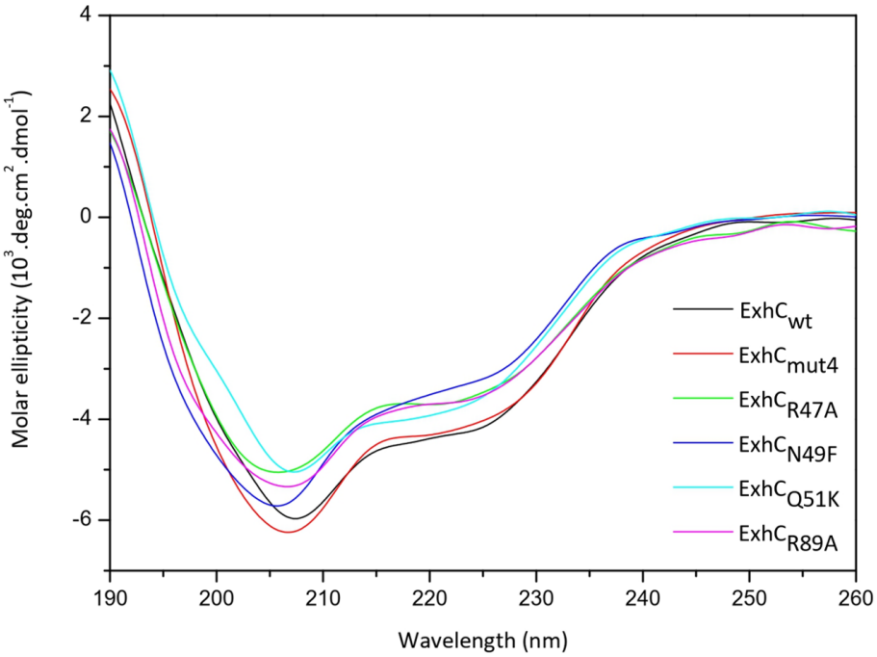


Figure S3. Characteristic CD-UV spectra of ExhC$_{wt}$ and its mutant variants at 25°C in 5 mM NaH_2PO_4 pH 7.4 and 50 mM NaF buffer solution.

Table S2. Percentages of secondary structure fractions of ExhC$_{wt}$ and its mutant variants, ETA and ETE obtained through UV-CD experiments. Deconvolution and calculation of secondary structure percentages were performed by the CONTINLL software from the CDPro software package using the SP29 protein reference set.

Structure	α-helix	β-sheet	Turns	Random Coil
ExhC$_{wt}$	9.4%	33.4%	21.4%	35.7%
ExhC$_{mut4}$	10.1%	34.7%	20.8%	34.4%
ExhC$_{R47A}$	8.4%	36.8%	21.1%	33.7%
ExhC$_{N49F}$	4.1%	47.5%	19.8%	28.6%
ExhC$_{Q51K}$	10.3%	32.3%	22.6%	34.8%
ExhC$_{R89A}$	8.6%	35.8%	21.4%	34.1%
ETA	17.5%	23.1%	12.3%	47.2%
ETE	12.7 %	19.5 %	12.7 %	55.0%

Table S3. Conformational change of the cavity containing the necrotic-linked region in the presence ($ExhC_{wt}$) and absence ($ExhC_{mut4}$) of residues R47, N49, Q51 and R89 from the molecular properties calculated by the ProteinPlus server [1].

	$ExhC_{wt}$	**$ExhC_{mut4}$**
Volume	100.43 A^3	381.76 A^3
Surface	177.20 A^2	464.47 A^2

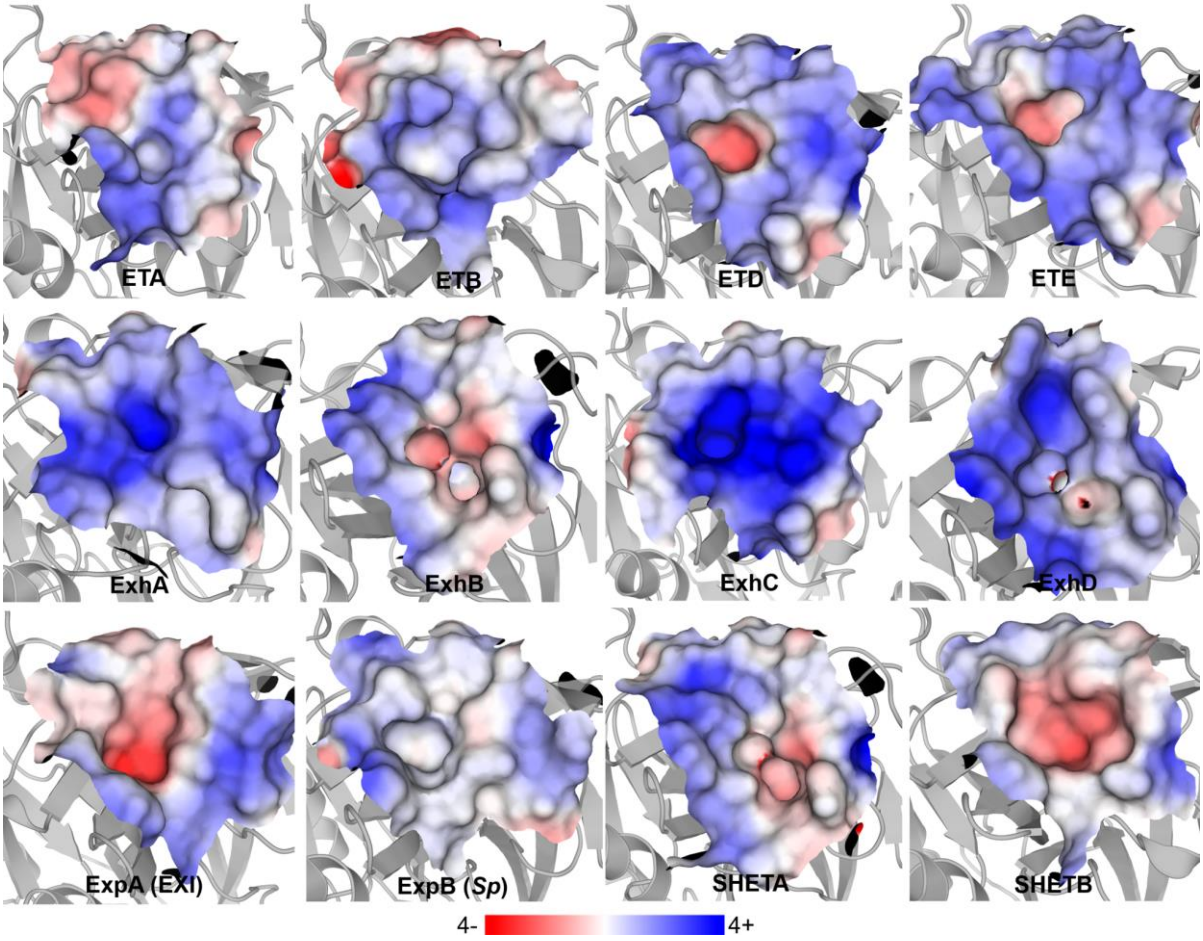


Figure S4. Electrostatic potential surface for the necrotic-linked region of $ExhC_{wt}$ and equivalent regions of other staphylococcal ETs. The protein surfaces are colored according to the electrostatic charge density, expressed in k_bT/e units, where k_b, T and e stand for the Boltzmann's constant, the temperature (298.15 K) and the electron charge, respectively.

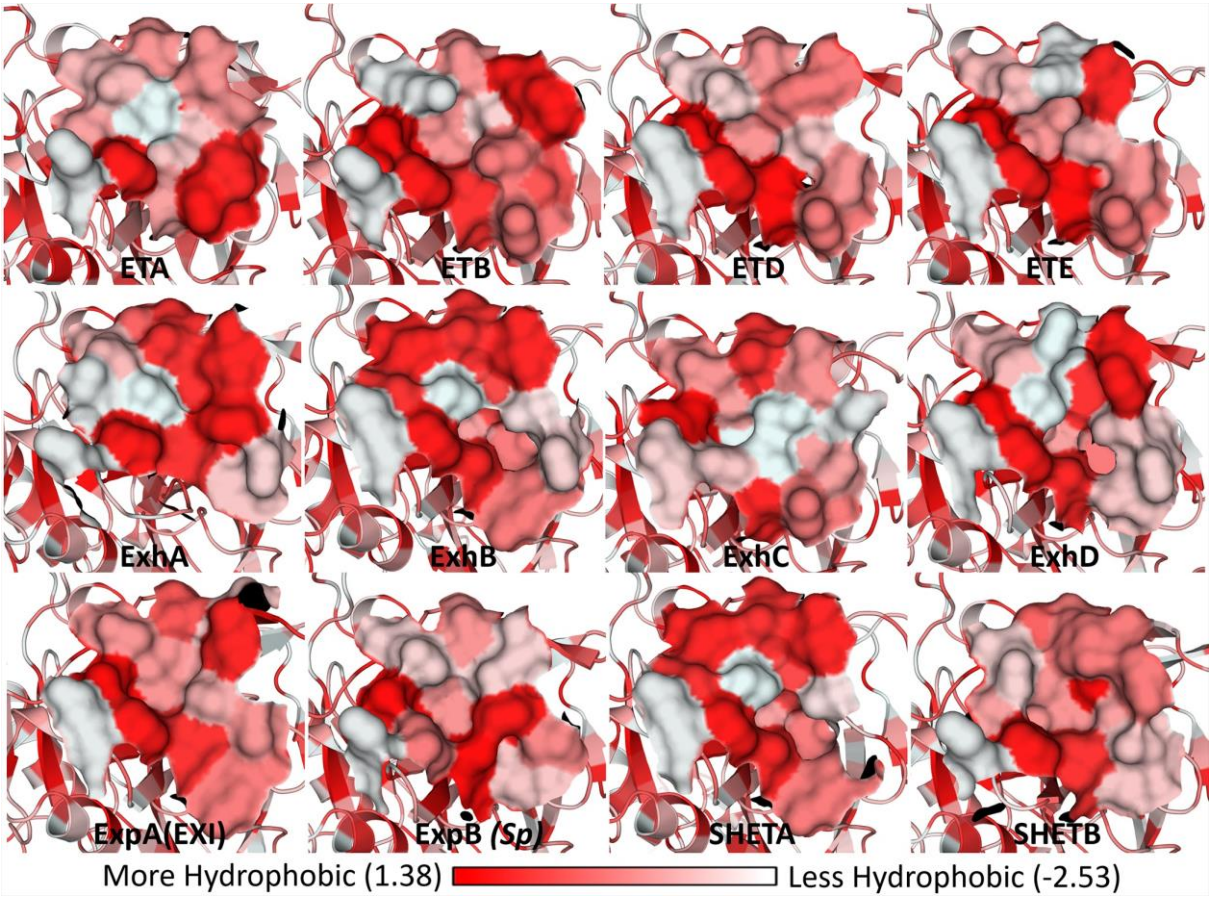


Figure S5. Structural representation of the necrosis-linked region ExhC and equivalent regions in other staphylococcl ETs, with the surface colored by residue hydrophobicity, according to the gradient below the figure. The normalized consensus hydrophobicity scale for each protein residue was employed [2].

References

[1] K. Schöning-Stierand, K. Diedrich, R. Fährrolfes, F. Flachsenberg, A. Meyder, E. Nittinger, R. Steinegger, M.J.N.a.r. Rarey, Proteins Plus: interactive analysis of protein–ligand binding interfaces, 48(W1) (2020) W48-W53.
[2] D. Eisenberg, E. Schwarz, M. Komaromy, R. Wall, Analysis of membrane and surface protein sequences with the hydrophobic moment plot, J Mol Biol 179(1) (1984) 125-42.

Appendix B - Supplementary material for the scientific article entitled *Staphylococcus aureus* Exfoliative Toxin E, Oligomeric State and Flip of P186: Implications for Its Action Mechanism

Carolina Gismene[a]; Jorge Enrique Hernández González[a]; Marília de Freitas Calmon[b]; Andrey Fabricio Ziem Nascimento[c]; Angela Rocio Niño Santisteban[a]; Felipe Antunes Calil[a]; Alana Della Torre da Silva[d]; Paula Rahal[b], Rejane Maira Góes[d]; Raghuvir Krishnaswamy Arni[a]; Ricardo Barros Mariutti[a]

[a]Multiuser Center for Biomolecular Innovation, São Paulo State University - UNESP, São José do Rio Preto, SP, Brazil.

[b]Laboratory of Genomic Studies, São Paulo State University - UNESP, São José do Rio Preto, SP, Brazil.

[c]Brazilian Synchrotron Light Laboratory (LNLS), Brazilian Center for Research in Energy and Materials (CNPEM), Campinas, SP, Brazil.

[d]Department of Biological Sciences, São Paulo State University - UNESP, São José do Rio Preto, SP, Brazil.

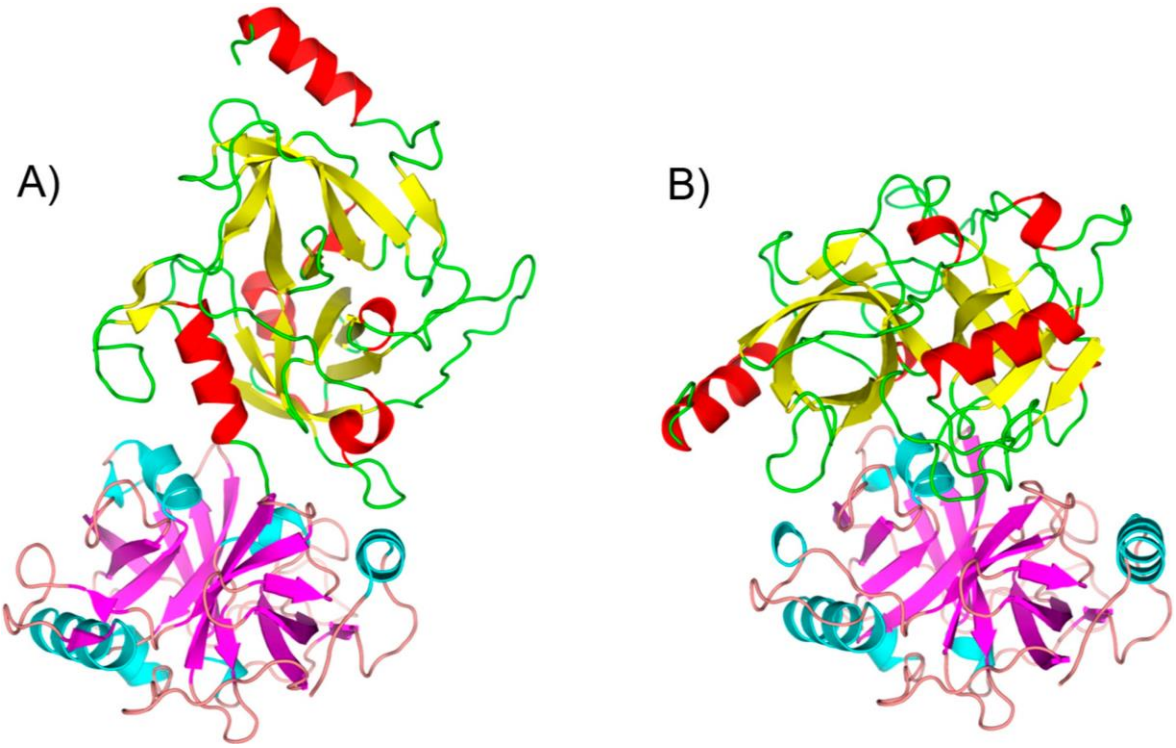


Figure S1. Representation of the two chains in the asymmetric units in the available crystal structures of ETE. A) PDB: 5C2Z. B) New crystal structure. Protein backbone is represented as cartoon and secondary structure elements are colored differently in each chain, i.e., α-helices, yellow/magenta for β-strands and green/salmon for loops in chains A/B. To facilitate the visual comparison of the structures, chain B is represented in the same orientation in both cases.

Table S1. Analysis of the stability of putative ETE homodimers in the available crystal structures.

Structure	Homodimer stability PISA	Homodimer stability PRODIGY-CRYSTAL	ΔGeff,100ns (kcal/mol)	ΔGPMF,1 (kcal/mol)
5C2Z	Not stable	100% crystallographic	-7.2	-
8DAX	Not stable	65.6% crystallographic	-30.2	-0.3(0.1)

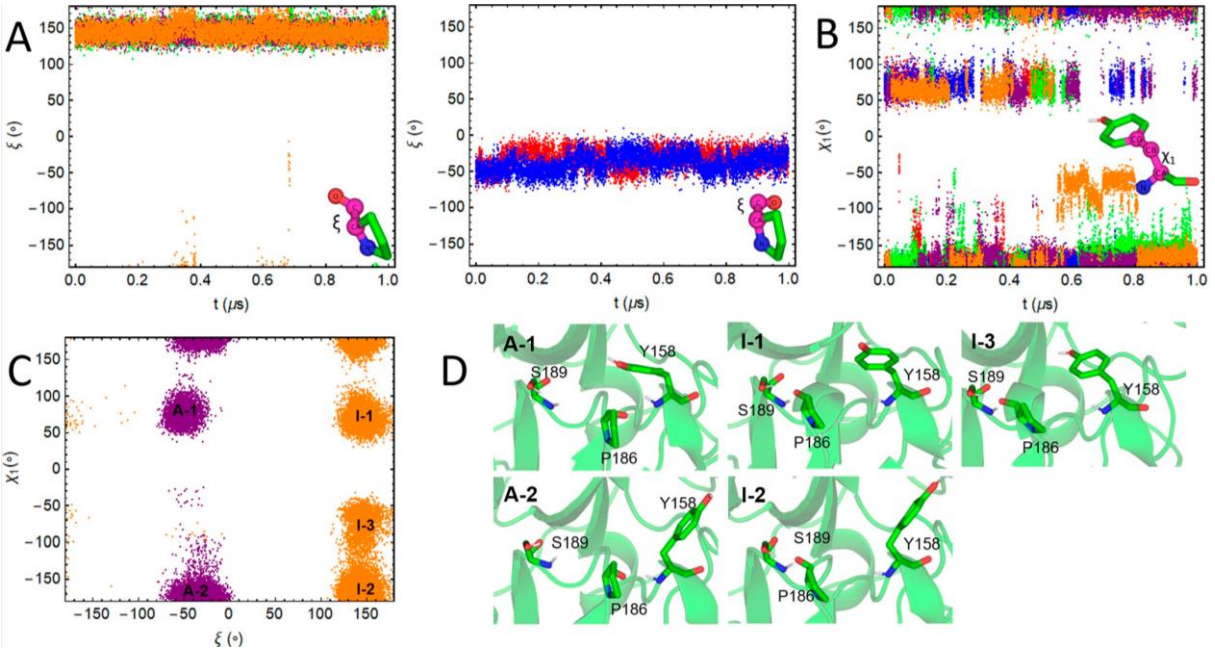


Figure S2. Fluctuation of P186 N-CA-C-O (ξ) and Y158 N-CA-CB-CG (χ1) dihedral angles throughout 1 μs replicate MD simulations of ETE monomers in water. A) Time profiles of P186 ξ dihedral during five replicate MD simulations, three started from an ETE structure with an inactive conformation of P186 (graph on the left) and two, from an ETE structure with an active conformation of P186 (graph on the right). **B)** Time profiles of Y158 χ1 dihedral along the five replicate MD simulations of ETE. **C)** Y158 χ1 vs. P186 ξ distribution along the replicate MD simulations. The five main populations of dihedral pairs are labeled. Letter A and I indicate whether the P186 ξ dihedra corresponds to an active or inactive conformation. The accompanying numbers after the hyphen are employed to discriminate the different populations of Y158 χ1 dihedral. **D)** Frames extracted from the MD trajectories representing each of the dihedral populations identified in C). In all cases, the profiles belonging to different replicate MD simulations are colored differently.

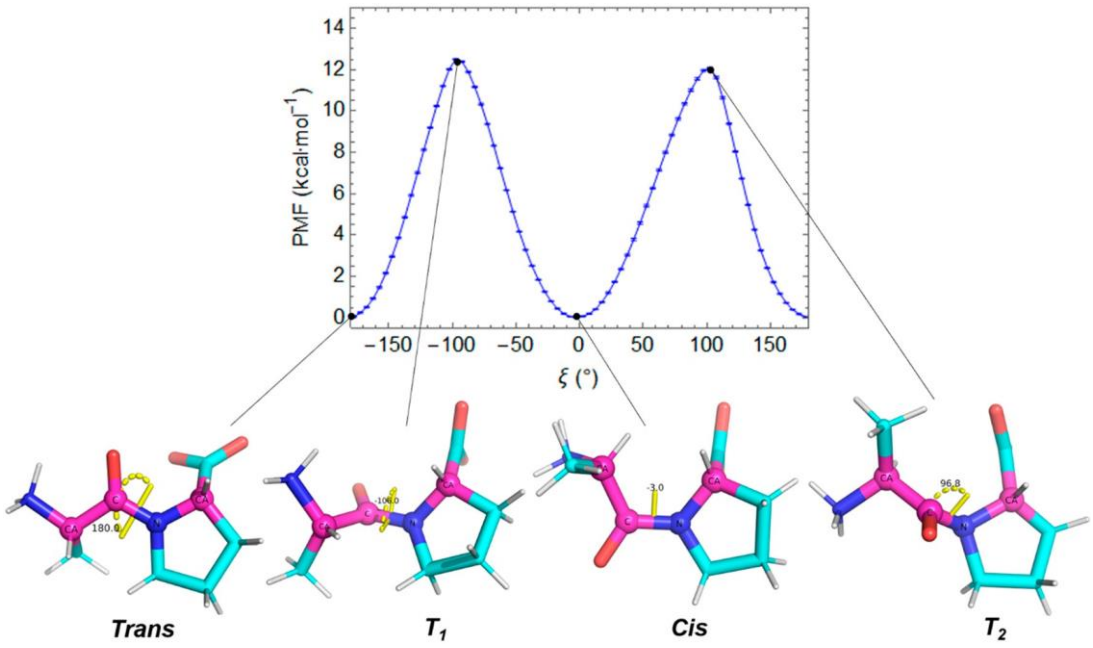


Figure S3. PMF for the rotation of L-alanyl-L-proline peptide bond. The structures of Ala-Pro conformations corresponding to free energy minima and transition states identified in the PMF are represented in the lower part of the figure. The local minima are the *trans* (CA-C-N-CA dihedral, ξ=±180º) and *cis* (ξ=0º) dipeptide isomers. The calculated free energy for the *cis-trans* isomerization ($\Delta G_{c\text{-}t}$) is 0.03±0.03 kcal/mol, which means that both states are isoenergetic, in agreement with the experimental $\Delta G_{c\text{-}t}$ value of -0.12±0.03 kcal/mol ($K_{c\text{-}t}$=1.21±0.06 at 313 K) [28]. The predicted free energies barriers associated with the *cis-trans* ($\Delta G^{\ddagger}_{c-t}$) and *trans-cis* ($\Delta G^{\ddagger}_{t-c}$) isomerization are 12.08±0.03 and 12.51±0.02 kcal/mol, respectively. The experimental energy barrier values for $\Delta G^{\ddagger}_{c-t}$ range from 18.2 to 20.8 kcal/mol, and for $\Delta G^{\ddagger}_{t-c}$, from 19.6 to 21.0 kcal/mol [28]. Therefore, the employed free energy method tends to underestimate the free energies of the T_1 and T_2 by 6 to 8 kcal/mol, probably as a result of inaccuracies inherent to modeling the molecular system with a classical force field.

Text S1. Size-exclusion chromatograms for ETE and ExhC

For the comparative analysis between ETE from *S. aureus* and ExhC from *S. sciuri*, experiments to obtain the latter protein were performed. *Escherichia coli* BL21(DE3) competent cells were transformed with pET28a(+)-ExhC expression vector. Then, we followed the same expression and purification protocol established for ETE from *S. aureus* (description in Protein expression and purification). ExhC was produced on a large-scale and high-purity (Figure S4B).

The chromatographic profiles of the ETE showed the main absorbance peak at elution volumes 10-12 ml, similar to the results observed for ExhC using the Superdex G75 10/300 GL column (GE) (Figure S4A). The large overlap between ExhC and ETE main elution peaks can be interpreted as a consequence of the high sequential and structural homology, as well as the identical oligomeric state of both proteins. However, relatively small discrepancies in their

elution volumes (Figure S4A) are expected because of the particular physicochemical characteristics of each ET [69]. In accordance with the SEC results, SDS-PAGE (Figure S4B) and BN-PAGE (Figure S4C) assays indicated that ETE and ExhC have similar sizes under denaturing (MW ~30 kDa) and native (MW ~56 kDa) conditions, respectively. This information reinforces that ETE is a dimer in solution, as previously proposed for ExhC [34].

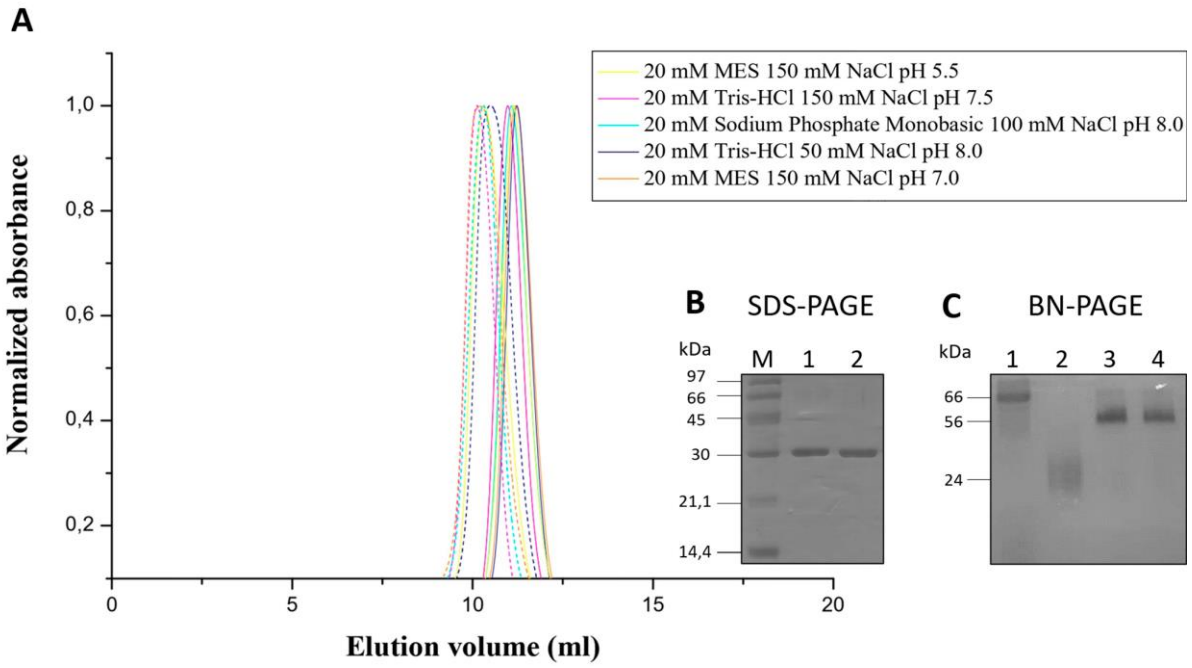


Figure S4. Comparison between ETE of *S. aureus* and ExhC of *M. sciuri*. A) SEC chromatograms for ETE (solid line) and ExhC (dash line) under different buffering conditions. Absorbance was measured at 230 nm in all cases, and normalized values are plotted. The overlap of peaks at ~10-12 ml indicates that ETE and ExhC elute in the same oligomeric state in all the tested conditions. B) SDS-PAGE of ETE (1) and ExhC (2) using the protein marker LMW-SDS Marker KIT – GE Healthcare (M). C) BN-PAGE of ETE (3) and ExhC (4) using BSA (1) and β- trypsin from bovine pancreas (2) as molecular weight markers. A running buffer containing 25 mM Tris pH 8.3 and 192 mM Glycine was employed for BN-PAGE.

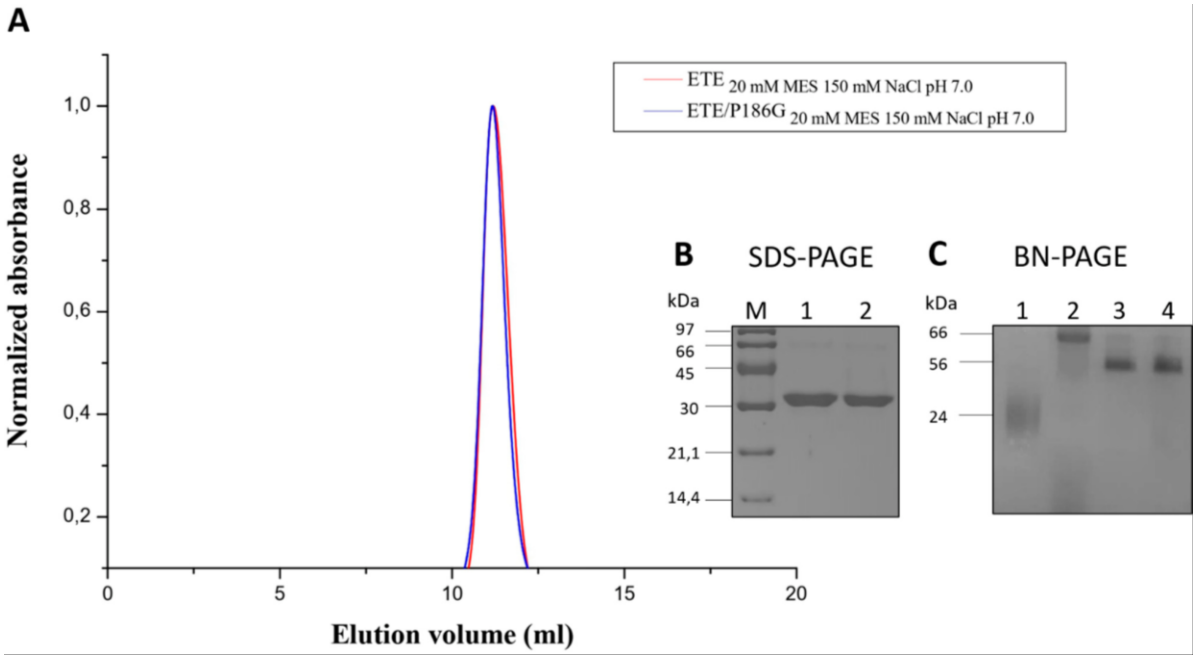


Figure S5. Comparison between SEC chromatograms of ETE and ETE/P186G of *S. aureus*. A) SEC chromatograms for ETE (red line) and ETE/P192G (blue line) in 20 mM MES pH 7.0 and 150 mM NaCl buffer. Absorbance was measured at 230 nm and normalized values are plotted. The overlap of peaks at ~10-12 ml indicates that ETE and ETE/P192G elute in the same oligomeric state in the tested condition. B) SDS-PAGE of ETE (1) and ETE/P192G (2) using the protein marker LMW-SDS Marker KIT — GE Healthcare (M). C) BN-PAGE of ETE (3) and ETE/P192G (4) using β-trypsin from bovine pancreas (1) and BSA (2) as molecular weight markers. A running buffer containing 25 mM Tris pH 8.3 and 192 mM Glycine was employed for BN-PAGE.

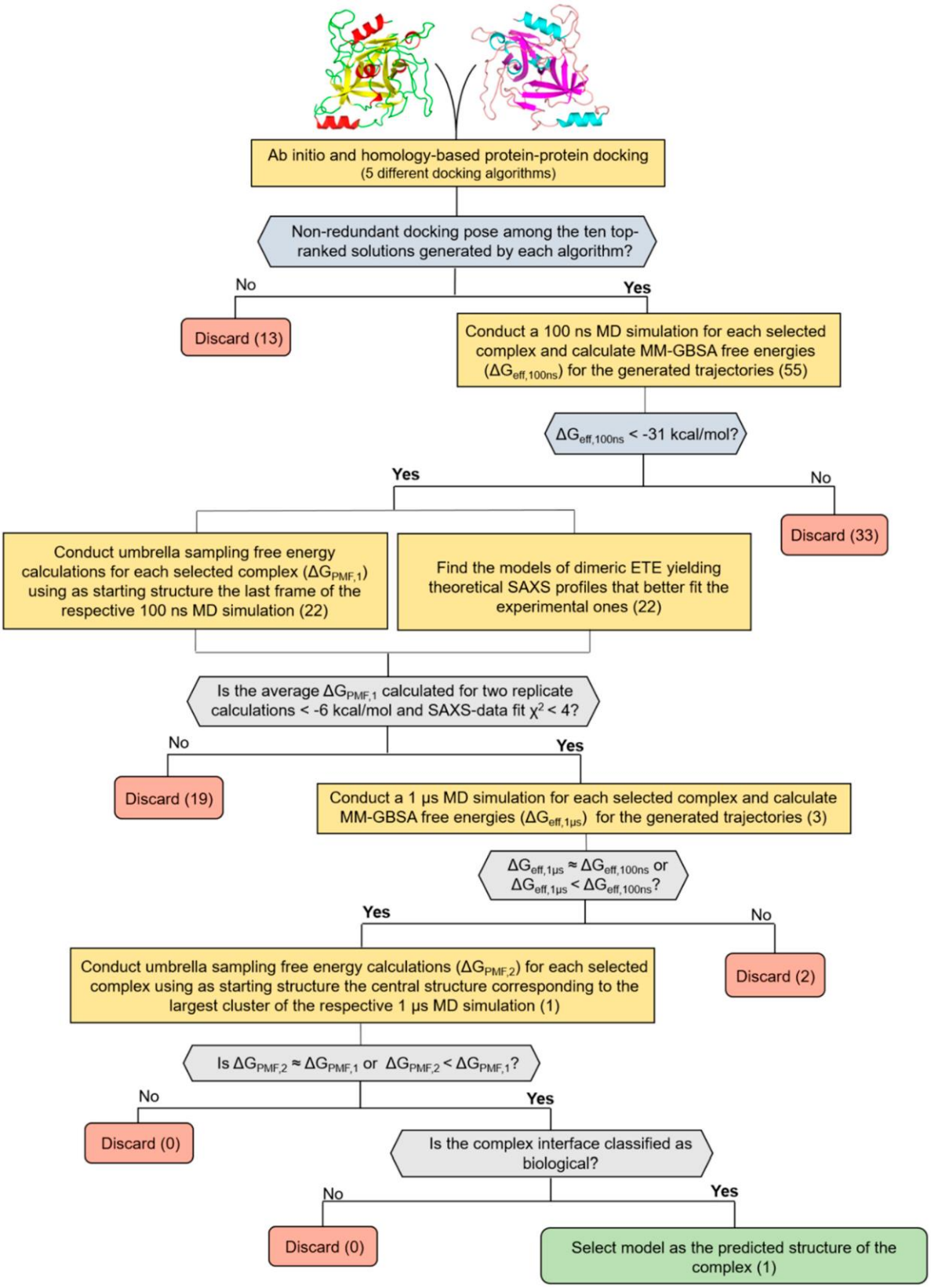


Figure S6. Workflow for the prediction of the most stable ETE dimers in solution. The number of models analyzed at every step is shown in parentheses. More details in Materials and Methods.

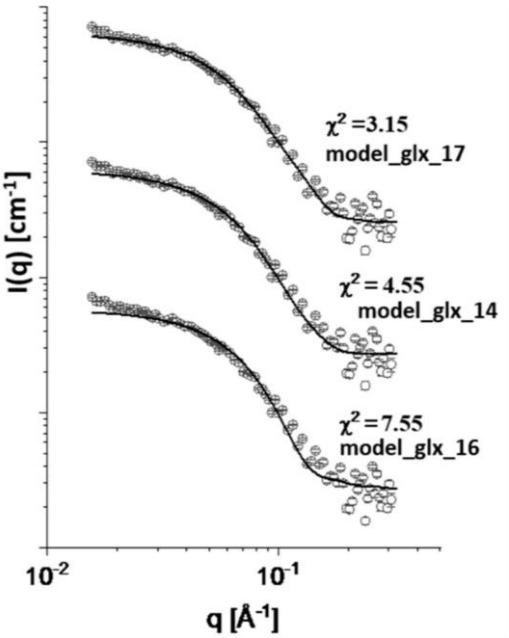
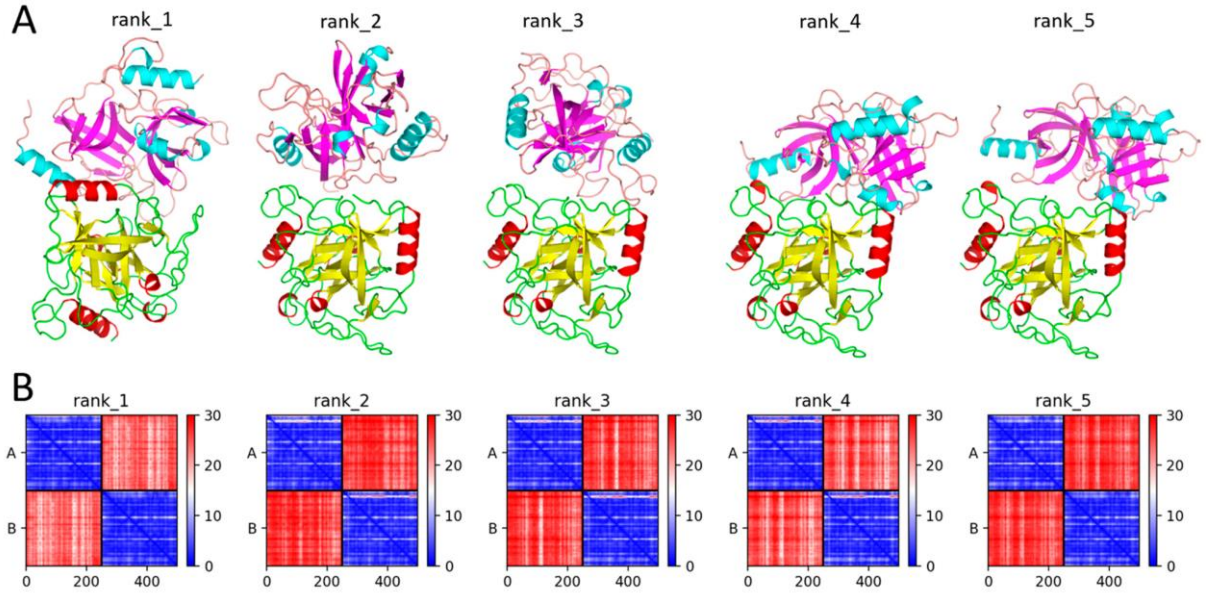


Figure S7. Examples of SAXS intensity profile fits using three different homodimer models. *I* and *q* stand or the intensity, expressed in arbitrary units, and the scattering vector, respectively. Models selected as examples will be discussed in the section below.



C

Model	$\Delta G_{eff,100ns}$ (kcal/mol)	SAXS χ^2	Dimer (%)
rank_1	-8.50	8.96	98.6

Figure S8. Models of ETE homodimer using ColabFold's AlphaFold-multimer version. A) Cartoon representation of five ETE homodimers. Each chain is colored differently. Models are ranked according to the average values of pLDDT metric. B) Predicted aligned error (PAE) matrices for the generated models. PAE is measured in Å. Low values indicate higher confidence in the positions of residues *x* and *y* are expected to be modeled correctly. The matrices indicate that AlphaFold2 correctly predicted the structures of the two monomers

(blue squares), whereas low confidence is associated with the prediction of the homodimer conformation (red squares). C) Assessment of the top-ranked model accuracy through MM-GBSA free energy calculations based on 100 ns MD trajectories and SAXS-data fits. Both $\Delta G_{eff,100ns}$ and χ^2 values compared poorly with those estimated from several docking poses (Table S3), thus reinforcing that AlphaFold2-multimer did not performed well when predicting the homodimer conformation.

Table S2. Analyses conducted for the selected docking poses (to be continued).

Pose	Docking Algorithm	$\Delta G_{eff,100ns}$ (kcal/mol)	$\Delta G_{PMF,1}$ (kcal/mol)	SAXS χ^2	Dimer (%)	$\Delta G_{eff,1\mu s}$ (kcal/mol)	$\Delta G_{PMF,2}$ (kcal/mol)
0	ClusPro	-45.6	-10.4(1.1)	3.32	94.8	-56.2	-9.9(0.4)
1		-27.9					
2		-31.8	-7.1(0.8)	4.99	98.3		
3		-9.0					
4		-21.9					
5		-17.4					
6		-16.3					
7		-26.0					
8		-30.1					
9		-45.9	-7.4(1.0)	4.03	100		
1	HDOCK	-55.5	-10.3(0.4)	5.30	97.9		
2		-23.1					
3		-28.4					
4		-11.0					
5		-30.5					
6		-18.8					
7		-32.7	-2.9(1.0)	7.12	100		
8		-32.7	-5.2(1.8)	3.98	89.6		
9		-38.0	-6.7(0.2)	5.37	95.8		
10		-33.7	-6.4(0.3)	3.17	84.2	-2.76	
1	HADDOCK	-22.9					
2		-23.5					
3		-15.4					
4		-46.7	-6.1(2.7)	5.50	100		
5		-36.7	-9.0(1.3)	4.08	100		
6		-42.9	-3.6(0.6)	4.24	97.9		
7		-27.2					
8		-27.0					
9		-6.2					
10		-19.2					

Table S2. Analyses conducted for the selected docking poses.

Pose*	**Docking Algorithm**	**Template PDB**	**$\Delta G_{eff,100ns}$ (kcal/mol)**	**$\Delta G_{PMF,1}$ (kcal/mol)**	**SAXS χ^2**	**Dimer (%)**	**$\Delta G_{eff,1\mu s}$ (kcal/mol)**
1	GalaxyWeb Homodimer Homology Modeling	1P3E	-8.5				
2		4IW4	-1.4				
3		2HLC	-3.2				
4		1LVM	-21.1				
5		1FUJ	-11.8				
6		4WVP	-26.7				
7		4WVP	-52.4	-1.7(0.3)	3.33	84.5	
8		4NSV	0.24				
9	Galaxy TongDock		-37.8	-4.6(0.5)	5.54	100	
10			-46.0	-9.9(0.8)	4.66	100	
11			-18.8				
12			-29.0				
13			-31.3	-5.40(0.03)	5.00	100	
14			-35.6	-4.9(1.3)	4.55	96.9	
16			-33.1	-4.8(1.1)	7.55	100	
17			-58.9	-9.9(0.9)	3.15	91.7	-25.3
18			-10.9				
19			-33.5	-6.1(2.0)	5.30	95.8	
7	LZerD+C_2[a]		-13.0				
8			-21.6				
9			-7.8				
3	LZerD No C_2[b]		-39.3	-6.6(2.6)	5.32	100	
8			-34.9	-4.9(0.1)	5.00	99.0	
9			-35.5	-1.9(0.7)	5.00	89.6	
10			-20.1				

[a]Poses are numbered according to the ranking established by each docking algorithm. The lower the pose number, the higher the associated docking score. Poses labeled in bold correspond to the five models that yielded the lower chi-squared values when fitting the SAXS profiles.
[b]LZerD can be conducted by forcing the complexes to display C_2 symmetry or not.

Table S3. Bioinformatic analysis of the stability of the predicted ETE homodimer.

Structure	PISA	PRODIGY CRYSTAL	SPPIDER IFR[a]
ETE_2-CP-0	Stable (ΔG_{diss}=3.1 kcal/mol)[b]	60.4% biological (ΔG_{bind}=-8 kcal/mol)[c]	G89, L91, V92 (chain A) L91, V92 (chain B)

[a]Interface forming residues predicted by SPPIDER.
[b]Free energy of dissociation of the homodimer. A positive value indicates a stable oligomer.
[c]Free energy of binding of the homodimer. In this case, a negative value corresponds to a stable oligomer. In principle, $\Delta G_{bind} = -\Delta G_{diss}$

Table S4. Comparison of homodimer interfaces in the available crystal structures and the proposed model.

Structure	Residues chain A[a]	Residues chain B[a]	Buried area ($Å^2$)[b]	RMSD (Å)[c]
5C2Z	K17, T18, K48, H66, R69, L70, E107, P136, S189, S205, G206, K207	E103, E104, I105, Y126, G127, D132, K231, T246, E248, Q249	542.5	0
8DAX	K48, T49, H66, I67, R69, L70, E72, E107, A108, G112, G113, Y158, N159, T177, P186, S189, G206, K207, G208, G209, Q210, F215, S226, Y227	K48, H66, R69, L70, E72, A108, G111, G112, G113, T177, P186, S189, S205, G206, K207, G209, Q210, F215, S226, Y227	1073.3	13.08
ETE_2-CP-0	K17, T18, R19, S21, F44, K48, T49, I50, R86, D87, G89, S90, L91, V92, Y158, N159, T160, S161, T162, Y183	K17, T18, R19, S21, F44, G47, K48, I50, R86, E88, G89, S90, L91, V92, Y158, N159, T160, S161, T162, H163, Y183, E185	1064.8	18.26, 22.21

[a]For comparison purposes, all residues appear in sequential order, as in PDB: 8DAX. A 4 Å cutoff was used to define the interface residues.
[b]Buried area of the complex interface was calculated with PISA.
[c]Heavy atom RMSD values between the homodimer conformations. The ETE structure 5C2Z was taken as reference for the calculation of all RMSD values, except for the second value of ETE_2-CP-0, which was determined using the structure 8DAX as reference.

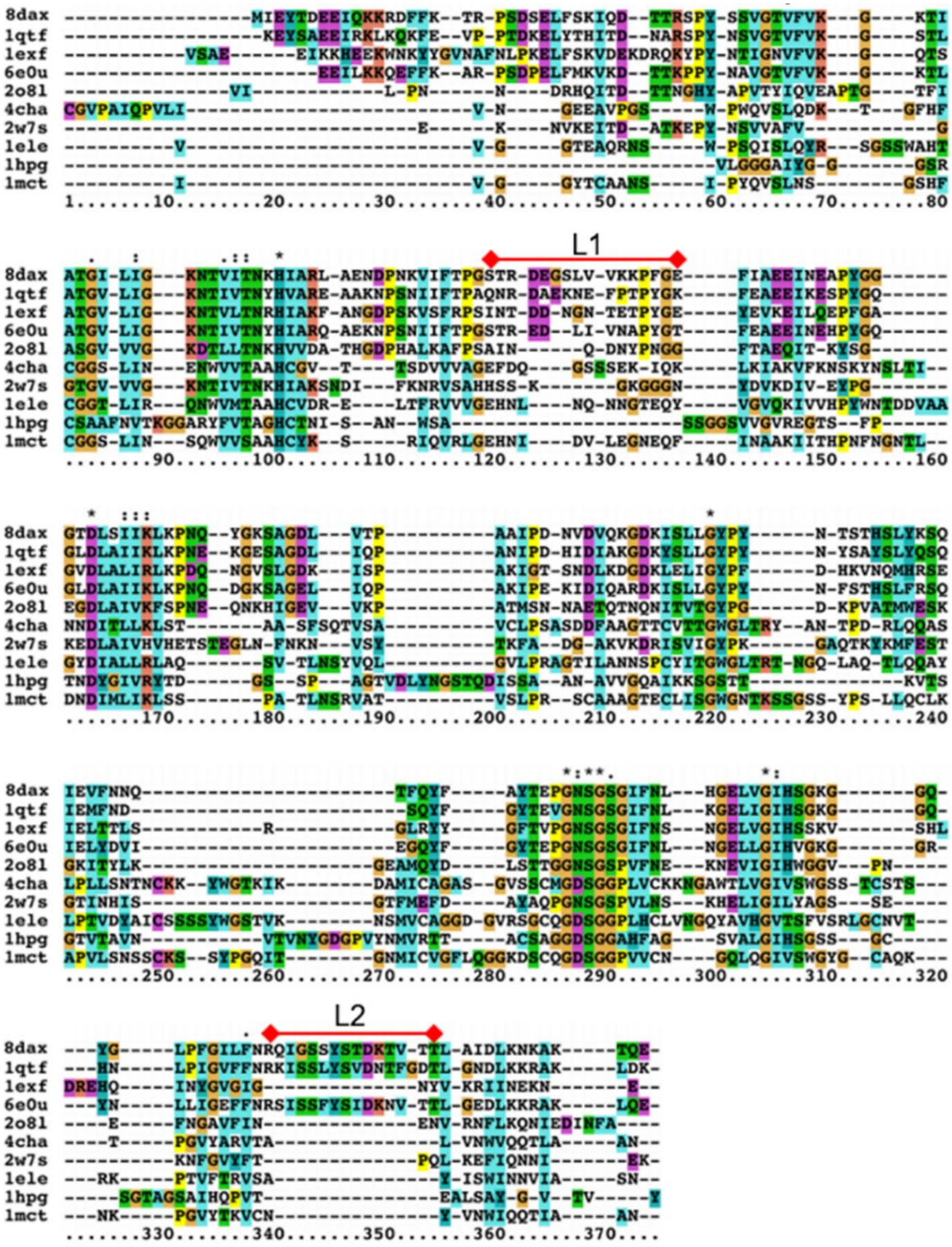


Figure S9. Structural alignment of different serine proteases shown at the level of their primary sequences. ETE (8DAX), ETB (1QTF), ETA (1EXF), EXI (6E0U), V8 protease (2O8L), Bovine chymotrypsin (4CHA), SP1A (2W7S), porcine pancreatic elastase (1ELE), Glu-SGP (1HPG) and bovine trypsin (1MCT) were structurally aligned at https://modbase.compbio.ucsf.edu/salign/ [70]. ClustalX2 [71] was used to visualize the alignment. Loops L1 and L2 are highlighted.

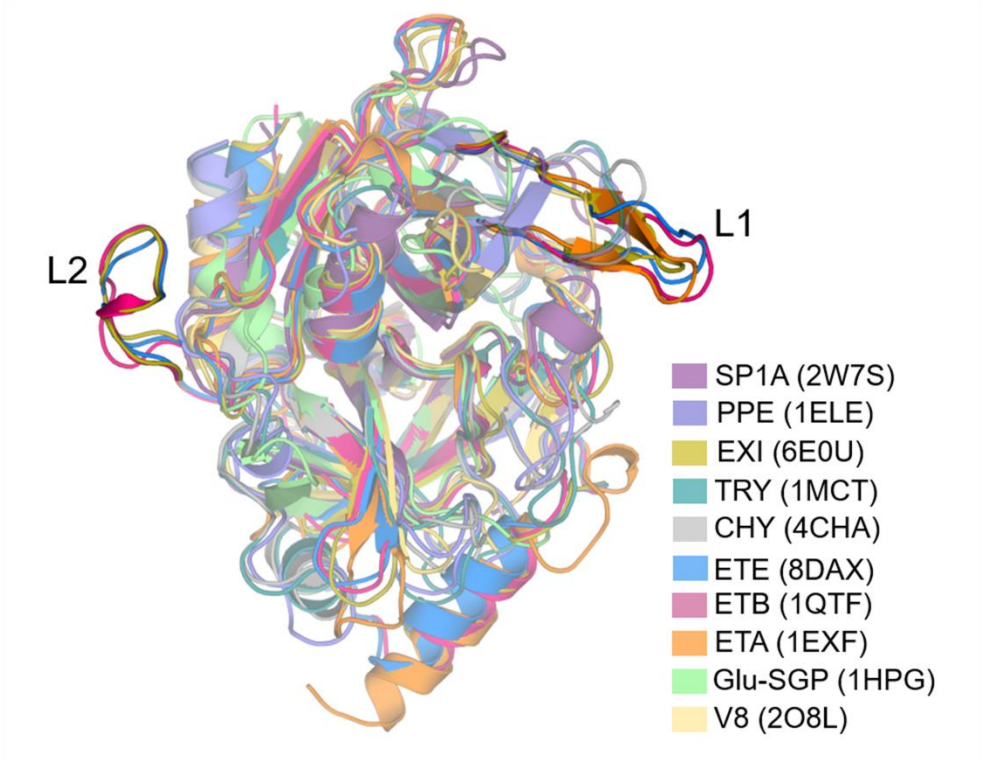


Figure S10. Structural alignment of different serine proteases. Loops L1 and L2 are highlighted. TRY, CHY and PPE stand for trypsin, chymotrypsin and porcine pancreatic elastase, respectively. The PDB ID of each structure is shown in parentheses. Proteins were structurally aligned at https://modbase.compbio.ucsf.edu/salign/.

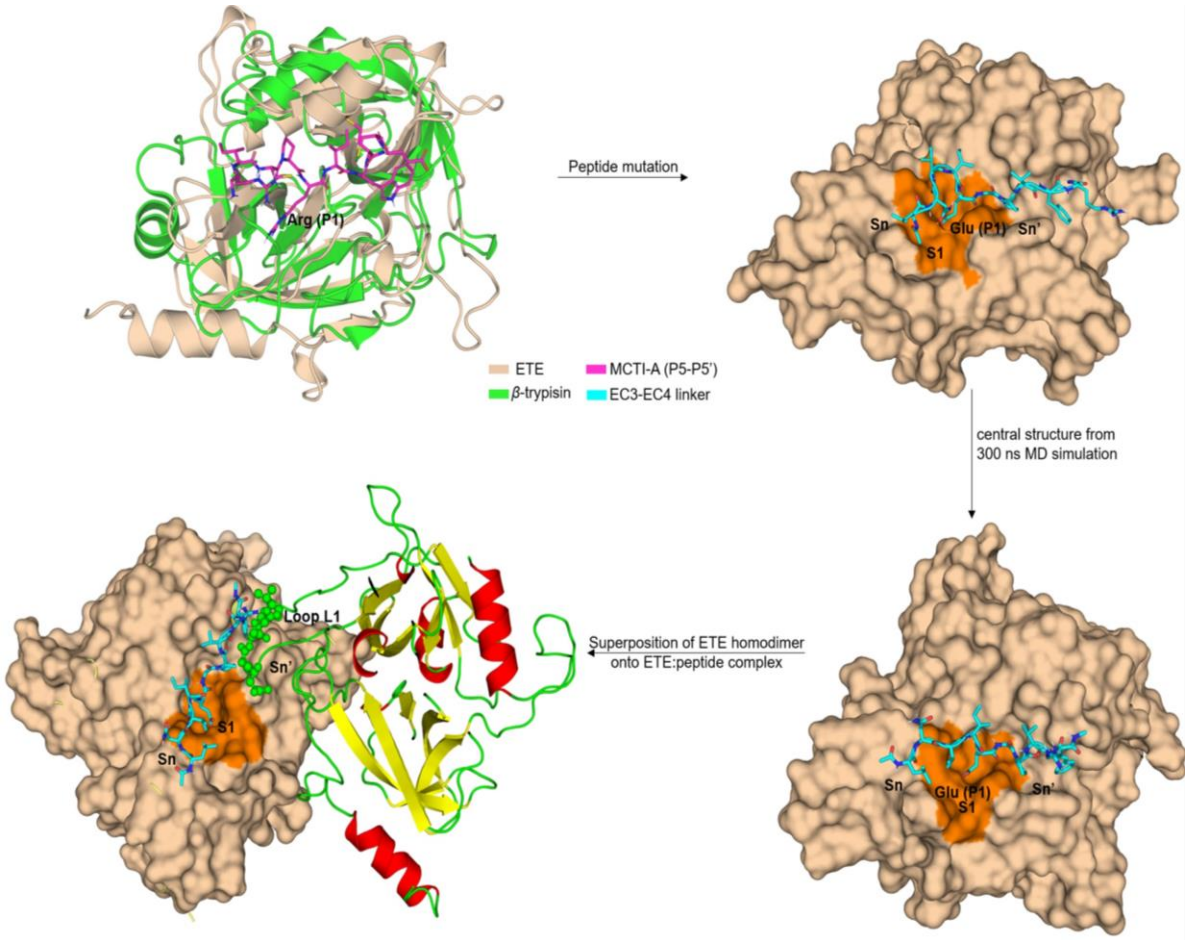


Figure S11. Model of ETE in complex with Dsg1 EC3-EC4 linker and superposition with the predicted structure of ETE homodimer. The crystal structures of ETE with an active oxyanion hole and porcine *β*-trypsin in complex with the peptide inhibitor MCTI-A (PDB 1MCT) were superimposed. The MCTI-A residues from P5 to P5' were then mutated to match the sequence of the EC3-EC4 linker (LNVIE↓GSVFR) with the mutagenesis plugin of Pymol 2.1.0 [32]. The central structure of the ETE:peptide complex from a 300 ns MD simulation was obtained and the ETE homodimer (ETE_2-CP-0) was superimposed onto the former structure. Notice that loop L1 (residues shown in spheres) inserts into the Sn' side of ETE active site, which sterically hinders the linker accommodation at the Sn' side of the active site when ETE is in the dimeric form. Due to the C2 symmetry of the ETE_2-CP-0 structure, the Sn' side of the second active site (not represented for brevity's sake) is also occupied by loop L1.

Text S2. Preparation of *N*-Tertbutyioxycarbonyl glutamic acid *O*-phenyl ester (Boc-L-Glu-OPh)

*Synthesis ester **2***. A solution of Boc-L-GIu-(OBn)-OH **1** (2.01 g, 5.95 mmol), phenol (557 mg, 5.92 mmol), dicyclohexylcarbodiimide (DCC) (1.219 g, 5.91mmol) and pyridine (491 mg, 6.2 mmol) in 30ml of ethyl acetate, was stirred at room temperature for 14 hours (step 1, Figure S12). The precipitated dicyclohexylurea was filtered. The filtrate was washed with saturated $NaHCO_3$ solution, 10% citric acid solution and water respectively. The organic layer dried over Na_2SO_4, filtered, and the solvent removed under reduced pressure. The crude mixture was

submitted to flash chromatography (SiO_2, 5:1 hexane-ethyl acetate) gave (750 mg, 1.81 mmol) (30%) of **2**. Spectral data were identical to the reported data [28].

^{1}H NMR (400MHz) ($CDCl_3$) : 1.45 (s, 9H), 2.09-2.18 (m, 1H), 2.38-2.42 (m, 1H), 2.50-2.64 (m, 2H), 4.57 (m, 1H), 5.14 (s, 2H), 5.13-5.20 (m, 1H), 7.08-7.34 (m, 10H)

*Synthesis of Boc-L-Glu-OPh (**3**).* The product **2** was dissolved in 12 ml of methanol, followed by the addition of 75 mg of Pd/C (10% w/w). The flask was closed with a rubber septum and a balloon with hydrogen was attached. This mixture was stirred at room temperature for 3 hours (step 2, Figure S12). After this time, the catalyst was removed by filtration under CELITE and the solvent evaporated under reduced pressure gave (286 mg, 0.9 mmol) (50%) of **3**. Spectral data were identical to the reported data [72].

^{1}H NMR (250MHz) ($CDCl_3$) MHz ^{1}H NMR ($CDCl_3$) 1.46 (s, 9H), 2.13-2.16 (m, 1H), 2.36 (s, 1H), 2.5-2.57 (m, 2H), 4.57 (d, 1H), 5.22-5.25 (m, 1H), 7.08-7.38 (m, 5H).

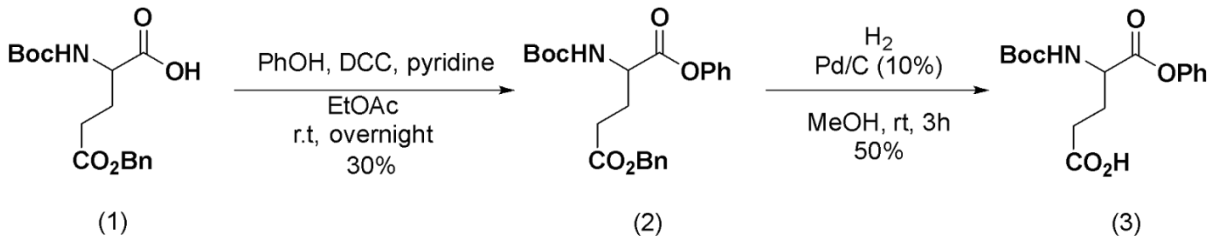


Figure S12. Synthesis of Boc-L-Glu-OPh. Steps are described in Text S2.

Text S3. Details of the Umbrella sampling free energy calculation protocols

The last frames of the 100 ns trajectories of the complexes with ΔG_{eff} < -31 kcal/mol and the central structures obtained from 1 μs MD simulations were chosen to conduct US free energy calculations (Figure 2) in order to predict the lowest-energy pose [65, 73]. The proteins were pulled along the *z* axis after careful manual orientation to prevent clashes. In total, 23 windows separated by 1 Å were created by displacing the second protein chain using the program *gmx editconf* of Gromacs 5.1.4 [74]. Two of the previous windows were obtained by approaching both proteins 2 and 1 Å, respectively, relative to the equilibrium position, i.e., that of the initial structure, while the remaining windows spanned distances between the reference atoms (see below) ranging from 0 to 20 Å along *z*. The six external degrees of freedom (DOFs) of each protein chain, i.e., the rotational and translational motions, were prevented by attaching the proteins to three dummy lead-like particles (D1, D2 and D3) that were kept fixed in the simulation box by means of tight position restraints (k= 100 kcal·mol^{-1}·Å^{-2}) [66, 75]. The attachment of each chain to the fixed dummy particles involved a distance restraint (D1-

G192(Cα), k= 10 kcal·mol^{-1}·Å^{-2}), two angular restraints (D2-D1-G192(Cα) and D1- G192(Cα)-I151(Cα), k= 200 kcal·mol^{-1}·rad^{-2}) and three dihedral restraints (D3-D2-D1- G192(Cα), D2-D1-G192(Cα)-I151(Cα) and D1- G191(Cα)-I151(Cα)-S168(Cα), k= 200 kcal·mol^{-1}·rad^{-2}), with the restraint equilibrium values being set to those initially calculated for every window of the different systems. Please, note that the previous k values must be halved when set in the restraint inputs of Amber 20 [58].

Each window was solvated in a cuboid box with edges along the z direction being placed, at least, 12 Å away from the protein surface to prevent interaction with copies in neighboring cells as the partners were separated within the simulation box. In the orthogonal axes, the edges were placed 8 Å away from the proteins. HMR was employed to increase the integration timestep from 2 to 4 fs during the production runs [62]. All windows were simulated in duplicate for 50 ns after appropriate EM and equilibration steps, which were identical to those reported for the conventional MD simulations, except for the gradual release of the restraints on the protein heavy atoms, which was not conducted in this case.

The distance profile between D1 and G192(Cα) of the second chain, i.e., the one pulled along the z axis, was collected for every window and potentials of mean force (PMFs) were obtained from these distance distributions using weighted histogram analysis (WHAM) with the wham program available at Grossfield lab's website (http://membrane.urmc.rochester.edu/?page_id=126). The first 10 ns of each production run were discarded prior to these calculations.

US free energy calculations were performed to predict the PMFs associated with the rotation around the N-Cα-C-O dihedral (ξ) of residue P186 of ETE. The N-Cα-C-O dihedral was manually rotated from -180 to 175° in 5° strides, totalizing 72 windows, and was restrained to each equilibrium value by means of a dihedral restraint with k=200 kcal·mol^{-1}·rad^{-2}. As before, EM and equilibration steps were carried out for each window, followed by a 50 ns production run. Hydrogen mass repartitioning was used to increase the performance. PMFs were calculated as explained in the previous paragraph, but this time on the basis of the collected N-Cα-C-O dihedral distributions for every window. All MD simulations were conducted five times and SEMs associated with each PMF point are reported as error bars in the graphs.

References

The references described in this appendix are referenced in chapter III.