



**UNICAMP**

**UNIVERSIDADE ESTADUAL DE CAMPINAS**

**INSTITUTO DE QUÍMICA**

**GONZALO GARCIA DELGADO**

**ESTUDOS SOBRE A OXIDAÇÃO DO INTERMEDIÁRIO LINEAR DO  
POLIÉTER IONÓFORO SALINOMICINA: CARACTERIZAÇÃO DA FLAVINA  
MONOOXIGENASE SaIC**

**CAMPINAS  
2021**

**GONZALO GARCIA DELGADO**

**ESTUDOS SOBRE A OXIDAÇÃO DO INTERMEDIÁRIO LINEAR DO  
POLIÉTER IONÓFORO SALINOMICINA: CARACTERIZAÇÃO DA FLAVINA  
MONOOXIGENASE SaIC**

Dissertação de Mestrado apresentada ao Instituto de Química da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Química na área de Química Orgânica.

**Orientador: Profa. Dra. Luciana Gonzaga de Oliveira**

**O arquivo digital corresponde à versão final da Dissertação defendida pelo aluno Gonzalo Garcia Delgado e orientada pela Profa. Dra. Luciana Gonzaga de Oliveira.**

**CAMPINAS  
2021**

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Química  
Simone Luiz Alves - CRB 8/9094

G165e Garcia Delgado, Gonzalo, 1992-  
Estudos sobre a oxidação do intermediário linear do poliéter ionóforo salinomicina : caracterização da flavina monooxigenase SalC / Gonzalo Garcia Delgado. – Campinas, SP : [s.n.], 2021.

Orientador: Luciana Gonzaga de Oliveira.  
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Química.

1. Flavina monooxigenase. 2. Produtos biológicos. 3. Biossíntese. 4. Streptomyces. I. Oliveira, Luciana Gonzaga de, 1975-. II. Universidade Estadual de Campinas. Instituto de Química. III. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Studies of the linear intermediary oxidation of ionophore polyether salinomicyn : flavine monooxygenase SalC characterization

**Palavras-chave em inglês:**

Flavin monooxygenase

Biological products

Biosynthesis

Streptomyces

**Área de concentração:** Química Orgânica

**Titulação:** Mestre em Química na área de Química Orgânica

**Banca examinadora:**

Luciana Gonzaga de Oliveira [Orientador]

Ljubica Tasic

Daniella Barbosa Trivella

**Data de defesa:** 10-02-2021

**Programa de Pós-Graduação:** Química

**Identificação e informações acadêmicas do(a) aluno(a)**

- ORCID do autor: <https://orcid.org/0000-0002-0225-7729>

- Currículo Lattes do autor: <http://lattes.cnpq.br/7756202041719632>

## **BANCA EXAMINADORA**

Profa. Dra. Luciana Gonzaga de Oliveira (Orientadora)

Profa. Dra. Ljubica Tasic (IQ - UNICAMP)

Dra. Daniela Barreto Barbosa Tivella (CNPEM)

A Ata da defesa assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

Este exemplar corresponde à redação final da Dissertação de Mestrado defendida pelo aluno **Gonzalo Garcia Delgado**, aprovada pela Comissão Julgadora em 10 de março de 2021.

*A ellos, que ya no están a mi lado, pero siempre vivirán en mi corazón*

## **Agradecimentos**

- Agradeço imensamente à Professora Luciana, quem além de orientar este projeto de pesquisa, sempre mostrou preocupação constante por meu bem-estar acadêmico e pessoal desde o primeiro dia na UNICAMP. Obrigado pela paciência, dedicação e apoio durante toda a pesquisa.
- À banca examinadora composta por Profa. Dra. Ljubica Tasic, Dra. Daniela Barbosa Trivella, Prof. Dr. André Damásio, Prof. Dr. Márcio Bertacine Dias pela disponibilidade de avaliar a dissertação, assim como à Profa. Dra. Anita Jocelyne e o Prof. Dr. Fabio Squina pelos aportes feitos nesta pesquisa.
- Agradeço ao pessoal do LABBICOMB; Ana, Bruno, Douglas, Renata, Julio, assim como Ana Tereza, Xisto e Bruno que contribuíram com ideias durante a pesquisa e me apoiaram durante o início da minha vida no Brasil.
- Às pessoas que me receberam no Brasil, Julie, Katherine, Josaphat, também ao Edison que sempre tem me apoiado incondicionalmente. Agradeço também ao Jonathan Janampa, Jonathan Yepes e Karen Villegas que tem me acompanhado em diferentes momentos no Brasil.
- Às pessoas que me apoiaram na Colômbia, Miriam, Luis Carlos Vesga, Luis Carlos Gomez e Bryant Castellanos.
- A minha família, meu Tio Oscar, meu irmão José, minha cunhada Wendy, Martha Probst, Felipe e Nicolas Navarro, meu primo Julian, os quais sem dúvida nenhuma, apoiaram-me durante a vida toda e sem eles, não teria alcançado muitas conquistas.
- Agradeço imensamente à Maria Fernanda e ao Juan Fernando, que chegaram a minha vida e mudaram tudo em mim para o bem.
- Agradeço à memória dos meus entes queridos que já partiram, o vestígio da vida deles no meu coração foi suficiente para continuar.
- Ao Instituto de Química pela estrutura e suporte à pesquisa realizada e aos professores com alta qualificação e trajetória cientista que fazem do IQ e da UNICAMP uma instituição de alta qualidade na pesquisa.
- À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo apoio a pesquisa no estado de São Paulo, Brasil (FAPESP 2019/10564-5, 2014/12727-5, e 2014/50249-8).

- Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio à pesquisa no Brasil.
- O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

## Resumo

O uso dos produtos naturais se intensificou na indústria farmacêutica nos últimos anos. Atualmente, estudos para aprimorar alvos aplicáveis em terapia podem ser feitos alterando-se a estrutura da uma molécula sem alterar a relação estrutura-atividade através da bioengenharia. A salinomicina é uma molécula biossintetizada por organismos do gênero *Streptomyces* e possui atividade terapêutica tanto na saúde veterinária quanto na saúde humana. Estudos diversos mostraram que a biossíntese da salinomicina não é realizada quando os genes *salC* e *salBIII* são deletados. Os resultados desta dissertação de mestrado apresentam a caracterização *in-silico* e *in-vitro* da enzima SalC que junto mostram que a enzima é uma monooxygenase dependente de FAD. Sua função está relacionada a reação de epoxidação em sítios de insaturação e com a capacidade de oxidar NADH para restaurar a flavina. Realizaram-se medições de consumo de NADH em sistemas de reações compostos de enzima (SalC) obtida por expressão heteróloga em *Escherichia coli* BL21(DE3), substratos com presença de insaturação e o intermediário linear obtido da cepa de *Streptomyces albus*  $\Delta$ *salC* buscando-se obter possíveis produtos oxidados dos materiais de partida. Também foram feitos estudos estruturais usando dicroísmo circular que mostraram uma enzima altamente flexível e junto aos modelos tridimensionais propostos pela ferramenta I-TASSER propõe-se uma estrutura tridimensional da enzima SalC que foi usada para fazer estudos *in-silico* destacando-se as interações das etapas distintas do ciclo de oxido-redução do FADH<sub>2</sub>. Finalmente, os resultados obtidos foram juntados com pesquisas prévias para elucidar uma etapa adicional na biossíntese da salinomicina.

## Abstract

The use of natural products in the pharmaceutical industry has been intensified along the last years. Nowadays, studies to improve applicable targets in therapy could be done changing the structure of the molecule without modifying the structure/activity relation through bioengineering. Salinomycin is a polyketide ionophore biosynthesized by the genus *Streptomyces* and studies showed therapeutic activity both in veterinary and human medicine. Different researches showed that salinomycin biosynthesis is blocked when *salC* and *salBIII* genes, tailoring gene from salinomycin biosynthetic gene cluster are deleted. This research project illustrates SalC enzyme *in-silico* and *in-vitro* characterization that together displayed that this enzyme belong to the monooxygenase FAD-dependent family and performs epoxidation reactions in unsaturation sites and uses NADH oxidation capacity for FAD restoration. NADH consumption measurements were done in reaction systems composed by SalC, the enzyme obtained by heterologous recombination in *Escherichia coli* BL21(DE3), unsaturated substrates and the linear intermediate obtained from the *Streptomyces albus*  $\Delta salC$  strain and potential oxidized products of the starting material were observed. Structural studies using circular dichroism revealed a highly flexible enzyme and combined with tridimensional models proposed by the I-TASSER tool we proposed a tridimensional structure for SalC that was used to simulate the interactions with the different stages of the FADH<sub>2</sub> oxireduction cycle. Finally, the whole set of results were pooled together to clarify an additional step in the salinomycin biosynthesis.

## LISTA DE ABREVAÇÕES

**ACP** Proteína transportadora de acila

**AT** Domínio aciltransferase

**BFP** Perfil do Fator – B

**pb** Pares de base

**CD** Dicroísmo circular

**CG/EM** Cromatografia Gasosa acoplada a Espectrometria de Massas

**DH** Domínio desidratases

**ER** Domínio enoilredutase

**FAD** Dinucleotídeo de flavina e adenina

**FAD<sup>+</sup>** Dinucleotídeo de flavina e adenina oxidado

**FADH<sub>2</sub>** Dinucleotídeo de flavina e adenina reduzido

**FMN** Flavina mononucleotídeo

**IPTG** Isopropil-β-D-1-tiogalactopiranosídeo

**KR** Domínio ceto-redutase

**KS** Domínio ceto-sintases

**KSQ** Domínio malonil-tioester descarboxilase

**LC/EM** Cromatografia Líquida acoplada a Espectrometria de Massas

**m/z** Razão massa/carga

**MS/MS** Espectrometria de Massas em tandem

**NAD(P)H** Nicotinamida adenina (fosfato) dinucleotídeo fosfato

**NADH** Nicotinamida adenina dinucleotídeo

**NCBI** National Center for Biotechnology Information (Centro Nacional para Informação Biotecnológica)

**PDB Protein DataBase** (Base de Dados de Proteínas)

**PKS** Sintase de policetídeo

**RMSD** Simply Root-Mean-Square Deviation (Desvio da raiz da média quadrada simples)

**SMO** Estireno Monooxigenase

**TM** Template Modeling (Planilhas de modelos)

**TS** Estado de Transição

**WT** Linhagem Selvagem

## LISTA DE FIGURAS

Figura 1. Molécula de salinomicina <sup>5</sup> .....	20
Figura 2. Cluster de genes responsáveis pela biossíntese da salinomicina <sup>7</sup> . O cluster de genes <i>sal</i> inclui 9 genes PKS ( <i>saIAI</i> – <i>saAIX</i> ) e 18 genes modificadores.....	21
Figura 3. Rota biossintética proposta para a obtenção da salinomicina <sup>7</sup> . .....	22
Figura 4. Estruturas obtidas na salinomicina seguindo as regras de Baldwin (azul) e anti-Baldwin (rosa).....	27
Figura 5. Estruturas da salinomicina (acima) e do intermediário linear (abaixo). .....	33
Figura 6. Cromatogramas dos extratos obtidos: A1_2 refere-se ao cultivo de <i>Streptomyces albus</i> (wt) onde a sinal <i>m/z</i> 769 [ $M_s + NH_4^+$ ] corresponde à salinomicina; B19D refere-se ao cultivo da bactéria <i>Streptomyces albus-ΔsaIC</i> , onde a sinal de <i>m/z</i> 744 [ $M_l + Na^+$ ] corresponde ao intermediário linear sodiado. .....	34
Figura 7. Comparação dos cromatogramas obtidos das análises de massas das frações eluídas por coluna Superclean™ LC-18 SPE observando-se o íon <i>m/z</i> 744 referente ao intermediário linear pré-salinomicina.....	35
Figura 8. Gel de agarose (1%) para a prova da integridade do plasmídeo após transformação em bactérias de <i>E. coli</i> BL21(DE3) usando plasmídeos isolados das cepas de estoque (DH10B) e de produção (BL21(DE3)). Linha 1: Marcador de massa molecular; linha 2 e 3: plasmídeo isolado de <i>E. coli</i> DH10B; linha 4: plasmídeo isolado de <i>E. coli</i> DH10B após digestão com as enzimas de restrição EcoRI e NdeI; linha 5: plasmídeo isolado de <i>E. coli</i> BL21(DE3); linha 6: plasmídeo isolado de <i>E. coli</i> BL21(DE3) após digestão com as enzimas de restrição EcoRI e NdeI; linha 7: plasmídeo isolado de <i>E. coli</i> DH10B pré-transformação; linha 8: plasmídeo isolado de <i>E. coli</i> BL21(DE3) pré-transformação; linha 9: Plasmídeo isolado de <i>E. coli</i> BL21(DE3) após transformação e clivado pelas enzimas de restrição NdeI e EcoRI; linha 10: plasmídeo isolado de <i>E. coli</i> DH10B após transformação e clivado pelas enzimas de restrição NdeI e EcoRI. ....	36
Figura 9. a) Gel de poliacrilamida (7%) feito para a prova da expressão do gene <i>saIC</i> . Pode-se ver a presença de bandas de 50 kDa tanto na fração solúvel sem IPTG como na solúvel com IPTG; linha 1 e 6: marcador molecular linha 2: fração	

insolúvel sem IPTG; linha 3: fração solúvel sem IPTG; linha 4: fração insolúvel com IPTG; linha 5: fração solúvel com IPTG, b) Gel de poliacrilamida (12%) com diminuição da concentração de acrilamida. Observa-se a produção da proteína parcialmente purificada com massa cerca de 50 kDa (linha 5), linha 1: marcador molecular; linha 2: fração solúvel; linha 3: <i>flowtought</i> ; linha 4: Lavagem com imidazol 60 mmol L <sup>-1</sup> ; linha 5: lavagem com imidazol 600 mmol L <sup>-1</sup> .....	37
Figura 10. Gel de poliacrilamida (12%) onde se apresenta a presença da banda maior a 50 kDa após da troca de tampão de fosfato 20 mmol L <sup>-1</sup> para Tris-HCl 20 mmol L <sup>-1</sup> . Linha 1: Marcador molecular; linha 2: Enzima SalC com tampão fosfato 20 mmol L <sup>-1</sup> , NaCl 300 mmol L <sup>-1</sup> ; linha 3: Enzima SalC com tampão Tris-HCl 20 mmol L <sup>-1</sup> . .....	38
Figura 11. Espectro CD da enzima SalC adquirido em temperatura de 25 °C.	39
Figura 12. Espectro de dicroísmo circular da enzima SalC compreendendo as faixas de temperaturas de 25 °C a 80 °C. ....	40
Figura 13. Variação do consumo de NADH feito pela enzima SalC com os substratos 3-buten-1-ol (preto); 3-metil-2-buten-1-ol (vermelho), 2-metil-3-buten-2-ol (azul), 2-metil-2-propen-1-ol (magenta), L-limoneno (verde), 4-penten-1-ol (roxo). ....	41
Figura 14. Cromatograma obtido da análise da transformação enzimática do 3-buten-1-ol usando SalC com uma proporção substrato:NADH 1:1. O pico eluído em 4,323 s mostra um padrão de fragmentação que pode ser atribuído para o epóxido esperado. ....	42
Figura 15. Espectro de massas gerado para o 2-(oxiran-2-il)-etan-1-ol .....	43
Figura 16. Comparativo do consumo de NADH dos substratos. A diminuição do consumo de NADH do L-limoneno e 4-penten-1-ol é estatisticamente significativa ( $p < 0.05$ ).....	44
Figura 17. Comparação dos cromatogramas obtidos: i. fração B17d (100% metanol) onde está presente a sinal do intermediário linear ( $m/z$ 744 [ $M_1 + Na^+$ ]); ii. fração de extrato que contém o intermediário linear com a enzima SalC (B17d 100% + SalC sem FAD) onde se observa o sinal de $m/z$ 776 [ $M_0 + Na^+$ ]. ....	45
Figura 18. Rota biosintética da salinomicina .....	46
Figura 19. Sequência de aminoácidos da enzima SalC. ....	47
Figura 20. Alinhamento entre as sequencias das enzimas SalC, monensina epoxidase (MonC), e LasC (lasalocida epoxidase). ....	48

Figura 21. Valores de BFP calculados para a enzima SalC comparados com as estruturas secundárias preditas e os sítios ativos e de ligação com FAD.....	49
Figura 22. Modelos gerados pela ferramenta I-TASSER para a sequência de aminoácidos da enzima SalC com seus respectivos valores de C-score. As imagens dos modelos foram feitas usando a ferramenta PyMOL, amarelo foi atribuído para as folhas- $\beta$ , vermelho para as $\alpha$ -hélices e verde para as estruturas randômicas.....	50
Figura 23. Alinhamentos do modelo 1 da enzima SalC (verde) com a enzima AlpK (ciano), PgaE (castanho claro), CabB (branco), aklavinona (magenta) e rifampicina monooxigenase (lilás). As linhas amarelas mostram as distâncias entre aminoácidos na mesma posição. Os cofatores e moléculas de água foram apagados.....	54
Figura 24. Estrutura tridimensional da enzima SalC com os valores considerados para sua escolha.....	56
Figura 25. <b>Acima:</b> Enzima Sal C (ciano) com o <i>Grid</i> escolhido (verde) para o estudo de <i>docking</i> . <b>Abaixo:</b> Superfície da enzima Sal C (esquerda) onde está ressaltado o bolso de ligação proposto para o FADH <sub>2</sub> (direita).....	60
Figura 26. Complexo SalC – FAD.....	61
Figura 27. Complexo SalC - FADH <sub>2</sub> . As ligações de hidrogênio formadas são apresentadas em verde.....	62
Figura 28. Complexo SalC – peroxiflavina. As ligações de hidrogênio são mostradas em verde.....	63
Figura 29. Complexo SalC - Hidroxiflavina. As ligações de hidrogênio são mostradas em verde.....	63

## LISTA DE TABELAS

Tabela 1. Substratos usados no teste de atividade enzimática da enzima SalC .....	31
Tabela 2. Contribuições das diferentes estruturas secundárias presentes na enzima SalC obtidas com os métodos K2D e CDNN .....	39
Tabela 3. Enzimas estruturalmente análogas com o modelo 1 da enzima SalC .....	53
Tabela 4. Valores de TM calculados para os alinhamentos entre os modelos 1, 2 e 3 e as enzimas AlpK e aklavinona.....	55
Tabela 5. Análise das diferenças entre as contribuições estruturais calculadas a partir do método K2D e comparadas com as porcentagens das estruturas secundárias do modelo 1 obtido do software I-TASSER e calculados com o método de STICKS (M1: Modelo 1, M2: Modelo 2, M3: Modelo 3, M4: modelo 4, M5, modelo 5) .....	55
Tabela 6. Domínios conservados na sequência de aminoácidos da enzima SalC .....	59
Tabela 7. Sítios de ligação para a enzima SalC.....	59
Tabela 8. Propriedades energéticas dos complexos obtidos .....	61

## LISTA DE ESQUEMAS

Esquema 1. Representação do mecanismo de reação da enzima estireno monooxigenase.....	24
Esquema 2. Resumo das etapas finais envolvidas na rota biossintética da monensina na qual a enzima MonCI promove a epoxidação estereoespecífica do intermediário linear e as enzimas MonBI e MonBII atuam como hidrolases ciclases (adaptado de Bhatt et. al, 2005). .....	25
Esquema 3. Resumo das etapas finais envolvidas na rota biossintética da lasalocida. LasC promove a epoxidação estereoespecífica do intermediário linear e a enzima LasB atua como hidrolase ciclase. (Adaptado de Minami et. al, 2008) .....	26
Esquema 4. Ciclo da reação de epoxidação catalisada pela enzima SalC .....	41
Esquema 5. Produto esperado após epoxidação enzimática da molécula 3-buten-1-ol. ....	42
Esquema 6. Oxidação do intermediário linear pré-salinomicina catalisado pela enzima SalC. ....	44
Esquema 7. Ciclo de reações proposto para a enzima LasC (Adaptado de Atsushi e colaboradores <sup>9</sup> ) .....	57

## Sumário

1	Introdução .....	19
1.1	Salinomicina.....	20
1.2	Epoxidação enzimática .....	23
1.2.1	Estireno monooxigenase.....	23
1.2.2	Monooxigenase putativa FAD-Dependente na biossíntese da Monensina: Enzima MonCI .....	24
1.2.3	LasC: Epoxidase FAD-Dependente na biossíntese da lasalocida .....	25
2	JUSTIFICATIVA.....	27
3	OBJETIVOS.....	29
4	METODOLOGIA .....	30
4.1	Fermentação dos metabólitos, expressão heteróloga e purificação da enzima SalC .....	30
4.2	Ensaio estruturais da enzima SalC .....	30
4.3	Ensaio de atividade enzimática <i>in vitro</i> .....	30
4.3.1	Ensaio da transformação do intermediário pré-salinomicina.....	32
5	Resultados .....	33
5.1	Fermentação e extração do intermediário linear .....	33
5.1.1	Expressão heteróloga da enzima SalC .....	35
5.1.2	Ensaio de dicroísmo circular .....	38
5.1.3	Ensaio de atividade enzimática <i>in vitro</i> .....	40
5.2	Estrutura tridimensional da enzima SalC .....	47
5.2.1	Estudo da sequência de aminoácidos da enzima SalC.....	47
5.2.2	Previsões <i>in silico</i> .....	52
5.2.3	Caraterização da atividade enzimática.....	56
5.3	Estudos de iteração da molécula FADH <sub>2</sub> com a enzima SalC .....	58
6	Conclusões .....	64
7	Referências.....	65

8 Anexos.....	68
---------------	----

## 1 Introdução

Segundo estudos recentes, o desenvolvimento de novos fármacos de fontes naturais, seja uma molécula natural isolada, um derivado de um produto natural ou um produto biológico isolado de um organismo, está readquirindo importância nos últimos anos<sup>1,2</sup>. Na revisão de Newman e Cragg do ano 2019<sup>3</sup>, foi realizada uma classificação da procedência dos novos fármacos com dados do *National Cancer Institute* e da *Food and Drugs Administration* nos Estados Unidos, de onde foi possível mostrar que os produtos naturais aumentaram a participação em medicamentos aprovados nos últimos 18 anos, principalmente para o tratamento de agentes infecciosos, evidenciando 88% dos fármacos são produtos naturais ou derivados destes<sup>3</sup>.

A despeito do seu potencial, a atividade biológica dos produtos naturais pode ser melhorada com diferentes métodos. O principal método atual de melhoramento são as modificações químicas na estrutura do composto, mas devido à dificuldade envolvida em alterar um grupo específico em uma molécula complexa estima-se que apenas 500 novos compostos de 3000 – 10000 produtos com potencial terapêutico tenham acesso farmacêutico<sup>2</sup>. Outra forma de melhorar a atividade biológica dos produtos naturais é pela bioengenharia, o que permite a derivatização da estrutura sem alterar a relação estrutura-atividade. Para fazer um melhoramento do potencial terapêutico de produtos naturais como a salinomicina, que possui um grande apelo no setor farmacêutico, pode-se aplicar bioengenharia e nesse sentido faz-se necessário entender como a biossíntese do metabolito é realizada. Nesta pesquisa, buscamos entender os aspectos bioquímicos e estereoquímicos relacionados a função da proteína SalC que atua na biossíntese da salinomicina, em especial na oxidação estereoespecífica do intermediário linear do tipo polieno obtido a partir da deleção do gene *salC*.

## 1.1 Salinomicina

A pesquisa de novos produtos naturais com potencial farmacêutico permitiu o estudo de muitos compostos, entre eles a salinomicina. Esta molécula é produzida por actinobactérias do gênero *Streptomyces* e contém em sua estrutura dois sistemas espirocetálicos, além de anéis do tipo pirano e furano<sup>4</sup> (Figura 1).

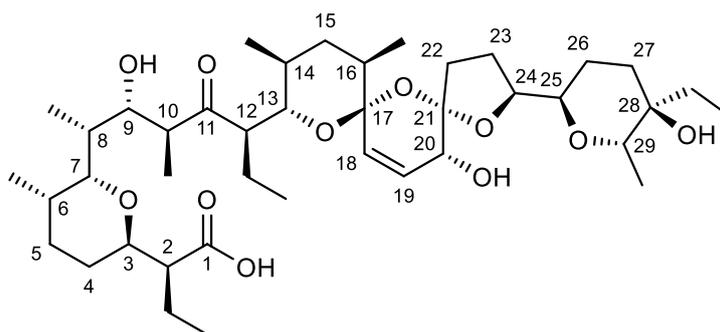


Figura 1. Molécula de salinomicina<sup>5</sup>

O núcleo composto por anéis pirano e furano é rico em oxigênio e permite a quelação de diferentes íons metálicos, como por exemplo o íon potássio. Além disso, devido a sua estrutura hidrofóbica, sua passagem é facilitada através da membrana celular. Esta passagem de íons pela membrana da célula gera uma despolarização e eventualmente a morte celular<sup>4,6</sup>.

A salinomicina é biossintetizada pelas enzimas codificadas no cluster de genes *sal* que incluem 9 genes PKS (*salA* – *salIX*) e 18 genes adicionais, onde pode-se encontrar enzimas com atividade regulatória, epoxidases FAD-dependentes, hidrolases, ciclases e oxidases dependentes de citocromo-P450, a maioria desses genes tem função putativa e sua atividade enzimática ainda está em estudo<sup>4,7</sup>. A Figura 2 mostra o *cluster* de genes biossintéticos *sal*.

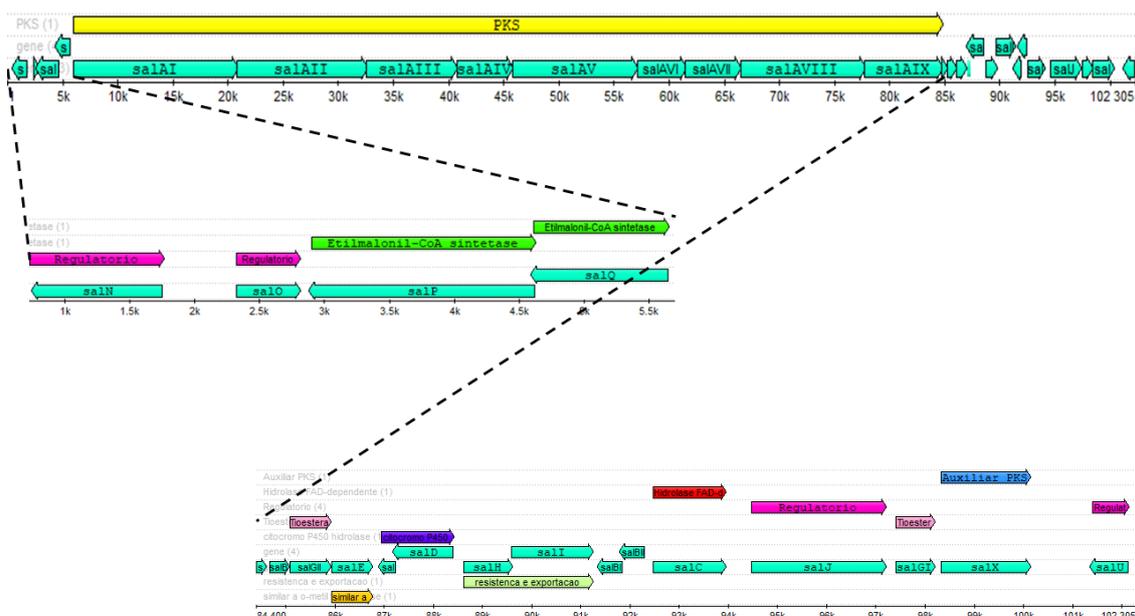


Figura 2. Cluster de genes responsáveis pela biossíntese da salinomicina<sup>7</sup>. O cluster de genes *sal* inclui 9 genes PKS (*salAI* – *salAIX*) e 18 genes modificadores.

Dois estudos propõem uma rota biossintética onde os genes PKS codificam para um módulo de carga compreendido por uma malonil-tioester descarboxilase (KSQ), aciltransferase específica para malonil-CoA (AT), proteína transportadora de acila (ACP) e 14 módulos de extensão. Após a biossíntese do policetídeo é proposto uma primeira ciclização nos carbonos C3 e C9 catalisada pela enzima SalBIII e então é feita a liberação do intermediário linear mediado pela proteína SalX, começando assim uma cascata de reações que envolvem a oxidação do policetídeo nos C24 e C28 pela enzima SalC, já que ela é uma epoxidase putativa, gerando um bis-epóxido. Na sequência ocorre a hidroxilação do carbono C20 feita pela enzima SalE para finalmente gerar o sistema de anéis furano e pirano feito pelas ciclases SalBI/II. Os estudos conseguem explicar uma sequência lógica dos passos da biossíntese mas tem a ausência da dupla ligação entre C18 e C19<sup>4,5,7</sup>. A Figura 3 ilustra a rota biossintética proposta para a salinomicina.

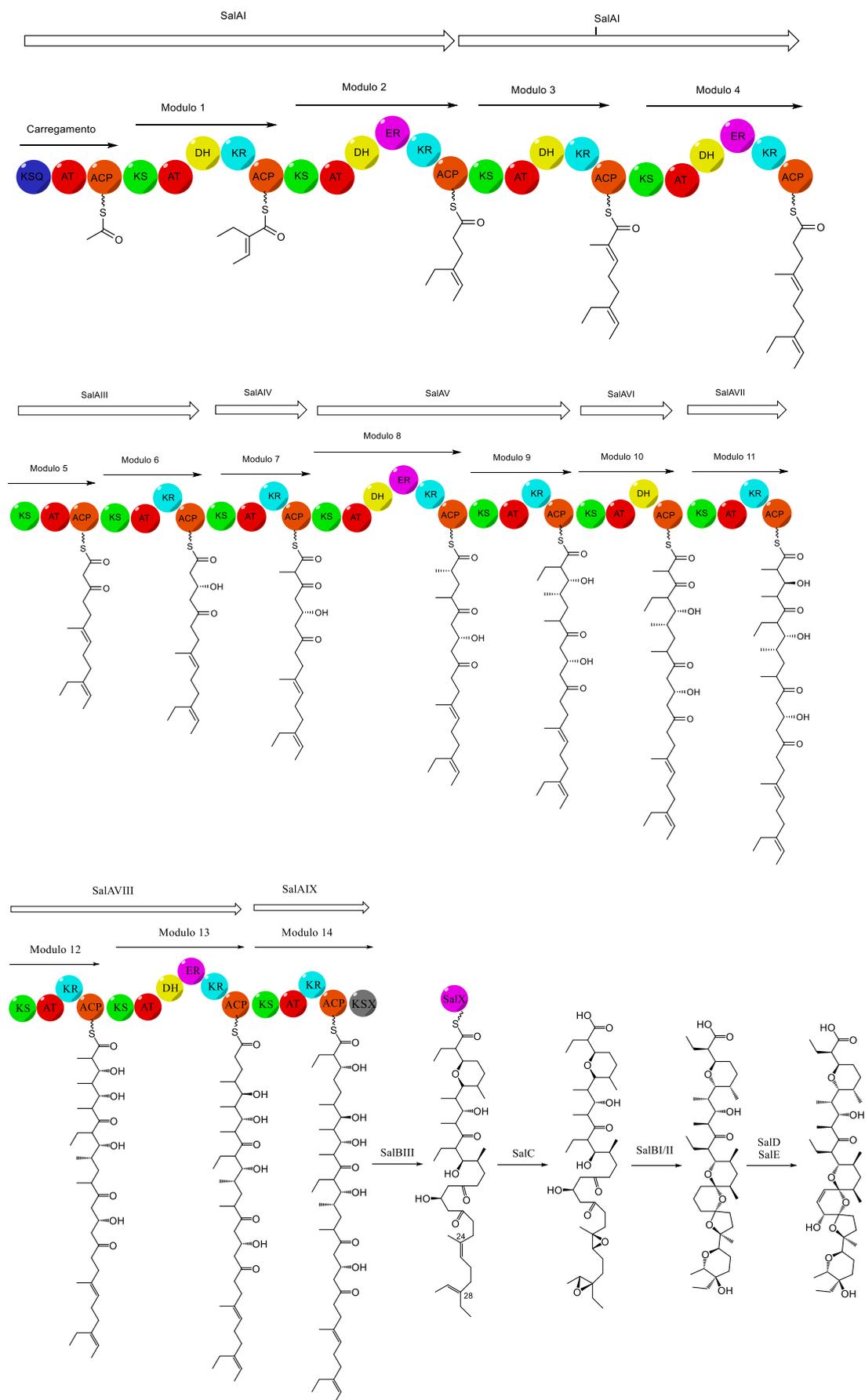


Figura 3. Rota biossintética proposta para a obtenção da salinomicina<sup>7</sup>.

## 1.2 Epoxidação enzimática

Diferentes organismos usam enzimas específicas com a capacidade de gerar grupos oxiranos em estruturas de certos metabólitos secundários, seja para adicioná-los na estrutura final ou para que os grupos epóxi sejam usados como precursores de menor energia na geração de estruturas mais complexas<sup>7-9</sup>. Tais enzimas são denominadas oxido-redutases que pertencem à família das monooxigenases, caracterizam-se pela necessidade de ter um cofator, seja um grupo FAD ou FMN nas flavino-dependentes ou um grupo heme nas heme-dependentes. Essas enzimas podem reagir também com substratos na presença ou na ausência de sítios de insaturação e necessitam de NAD(P)H, o qual participa como um regenerador do cofator. Esta pesquisa está focada no estudo de monooxigenases FAD-dependentes e a seguir são mostrados alguns exemplos de epoxidases FAD-dependentes.

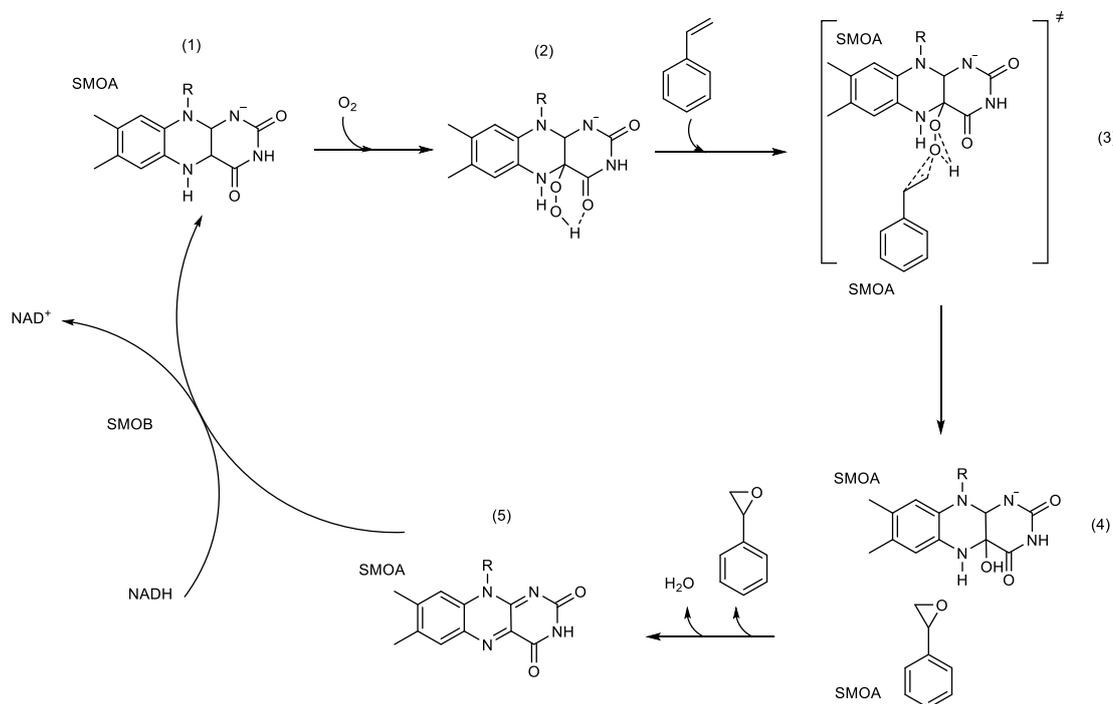
### 1.2.1 Estireno monooxigenase

Estireno monooxigenases compreendem um grupo de enzimas que convertem o estireno para óxido de estireno, encontram-se em diversos organismos e atualmente são usadas frequentemente em biossíntese *cell-free*<sup>8,10</sup>. As estireno monooxigenases (SMOs) têm na sua estrutura dois domínios conservados:

- O domínio A (SMOA) é a componente monooxigenase, que trabalha na formação do epóxido com alta enantiosseletividade com auxílio do cofator FAD<sup>+</sup>;
- O domínio B (SMOB) é a componente redutase dependente de NADH, que permite a regeneração do cofator<sup>10</sup>.

O sítio SMOA interage com a coenzima FADH<sub>2</sub> (1) que reage cataliticamente com o oxigênio molecular para gerar um grupo peróxido no cofator (2). Este grupo peróxido dispõe de um oxigênio eletrofílico que interage com os elétrons  $\pi$  da dupla ligação não aromática no estireno (3), gerando uma hidroxiflavina (4). A liberação do óxido de estireno e de água gera um FAD oxidado que interage fracamente com a SMOA (5), o FAD oxidado é reduzido pelo sítio SMOB NADH-dependente. No Esquema 1 é possível observar a

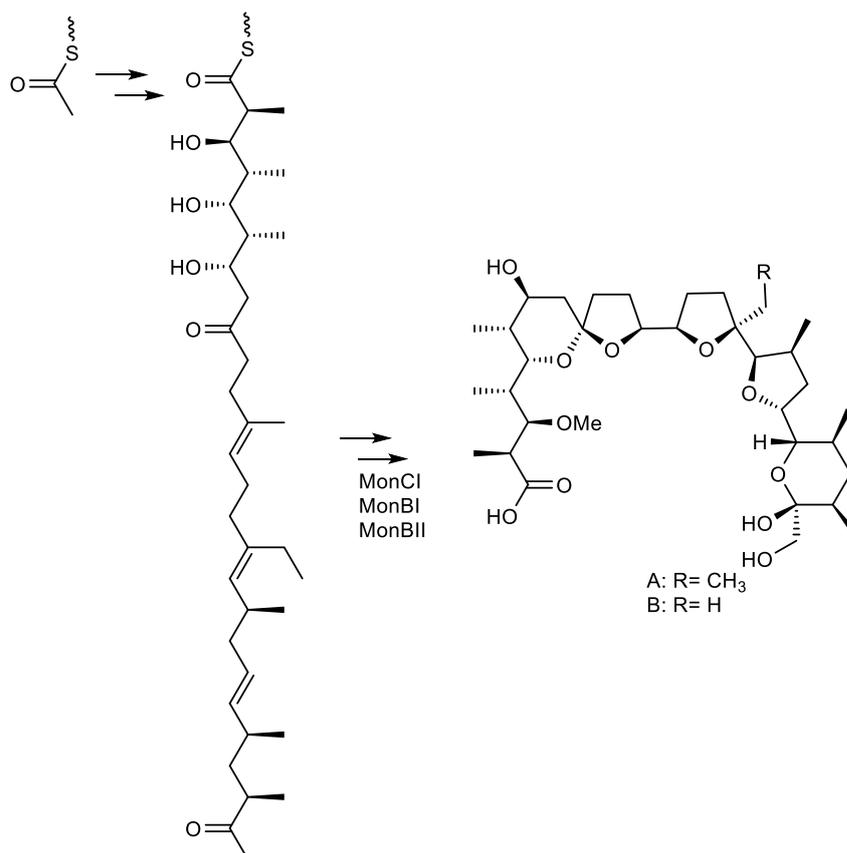
reação de epoxidação do estireno catalisado pela enzima estireno monooxigenase.



Esquema 1. Representação do mecanismo de reação da enzima estireno monooxigenase.

### 1.2.2 Monooxigenase putativa FAD-Dependente na biossíntese da Monensina: Enzima MonCI

A monensina é um ionóforo com atividade antimicrobiana obtida através da fermentação das bactérias *Streptomyces cinnamonensis*. Pertence à família de metabólitos secundários de policetídeos estruturalmente relacionados, sendo um dos mais bem estudados. A monensina é biossintetizada por um agrupamento de genes que inclui um complexo multienzimático PKS (MonAI – MonAVII) encarregados dos 12 ciclos de condensação necessários para a formação do esqueleto molecular policetídeo. Neste cluster de genes estão presentes também enzimas que fazem modificações estruturais para finalmente obter a monensina<sup>11</sup>.



Esquema 2. Resumo das etapas finais envolvidas na rota biossintética da monensina na qual a enzima MonCI promove a epoxidação estereoespecífica do intermediário linear e as enzimas MonBI e MonBII atuam como hidrolases ciclases (adaptado de Bhatt et. al, 2005).

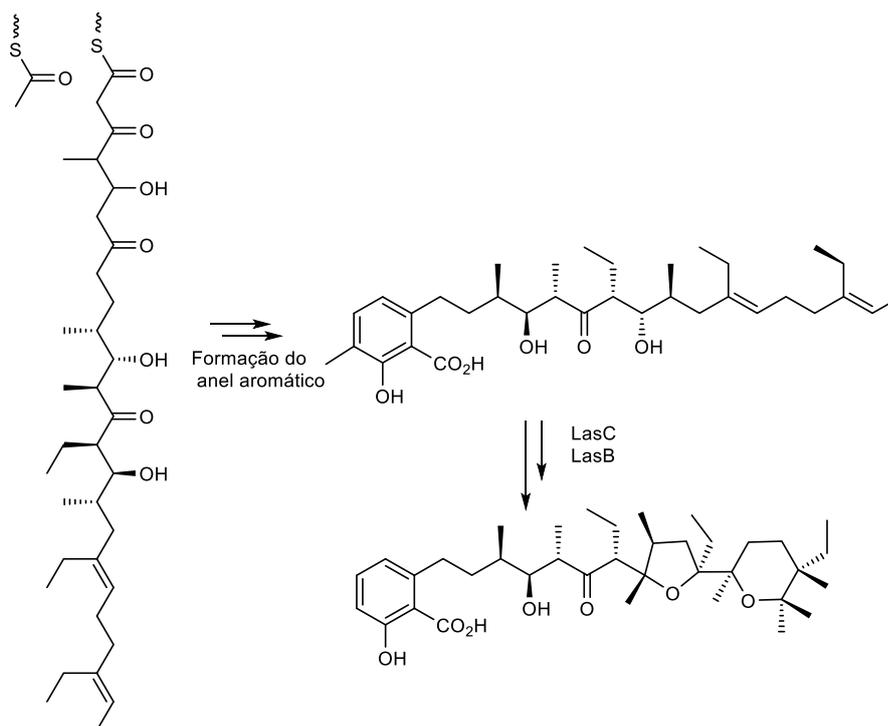
Estudos de biologia molecular feitos por Bhatt e colegas<sup>11,12</sup> mostraram que entre as enzimas encarregadas de fazer as mudanças estruturais no intermediário policetídeo, encontra-se a enzima MonCI, a qual promove a oxidação do intermediário linear para gerar 3 grupos epóxi que guiam a estereoquímica característica da monensina<sup>11,12</sup>. Experimentos de deleção do gene *monCI* mostrou o acúmulo do policetídeo não oxidado. O estudo do gene permitiu associá-lo com uma função putativa de epoxidase FAD-dependente<sup>12</sup>.

### 1.2.3 LasC: Epoxidase FAD-Dependente na biossíntese da lasalocida

A lasalocida é um agente antibacteriano obtido de cepas de bactérias *Streptomyces lasaliensis*. Assim como na biossíntese de metabólitos especializados como a monensina e salinomicina, a lasalocida contém um grupo de genes que codificam uma PKS tipo I (LasAI-LasAVII) que sintetizam um policetídeo como intermediário linear<sup>13,14</sup> que após sua liberação será

transformado enzimaticamente em lasalocida. Estudos demonstraram que as enzimas LasB e LasC participam na biossíntese da lasalocida com funções de ciclase e epoxidase respetivamente.<sup>13</sup>

No caso específico da enzima LasC, esta foi classificada como uma epoxidase putativa FAD-dependente<sup>13</sup>. De acordo com Minami e colaboradores<sup>9</sup>, a enzima LasC tem a capacidade de oxidar o intermediário linear, adicionando grupos epóxi em sítios de insaturação como foi predito. Para realizar a oxidação, a LasC utiliza uma flavina como cofator que atua como doadora de elétrons e que é restaurada pela oxidação de uma molécula de NAD(P)H. Também foi demonstrada a promiscuidade da enzima no momento de oxidar moléculas do tipo dieno truncadas<sup>9</sup>. O Esquema 3 apresenta o mecanismo de reação proposto para a epoxidação de alcenos catalisada pela enzima LasC.



Esquema 3. Resumo das etapas finais envolvidas na rota biosintética da lasalocida. LasC promove a epoxidação estereoespecífica do intermediário linear e a enzima LasB atua como hidrolase ciclase. (Adaptado de Minami et. al, 2008)

## 2 JUSTIFICATIVA

Como apresentado na introdução, as bactérias do gênero *Streptomyces* têm a capacidade de sintetizar ionóforos com propriedades antibióticas. Clusters de genes como *mon*, *las*, *tet*, e *sal* são responsáveis pela biossíntese dos poliéteres monensina, lasalocida, tetromicina e salinomicina respectivamente. O estudo desses clusters de genes mostrou que existem domínios conservados correspondentes para módulos PKS assim como enzimas pós-PKS que atuam como modificadoras do intermediário policetídeo linear liberado uma vez terminadas as etapas de alongação feita pelas enzimas PKS. Ditas enzimas tem a capacidade de epoxidar estereosseletivamente sítios de insaturação com geometria *trans* ou *E* e assim formar um poli-epóxido (enzimas da família *c*) que será usado por enzimas ciclases (família *b*) para gerar uma cascata de abertura e fechamento de anéis para formar os grupos pirano e furano característicos dos ionóforos<sup>4,7</sup>.

A biossíntese da salinomicina tem a atuação de enzimas modificadoras responsáveis pela formação de duas funcionalidades epóxido (SalC) e pela formação dos sistemas bis-espirocetálico e pirano (SalBI/SalBII). A formação dos anéis B e E não seguem as regras de ciclização de Baldwin, sendo favorecido o fechamento 6-*endo-tet* (ataque o C29') e não o esperado 6-*exo-tet* (ataque ao C28).

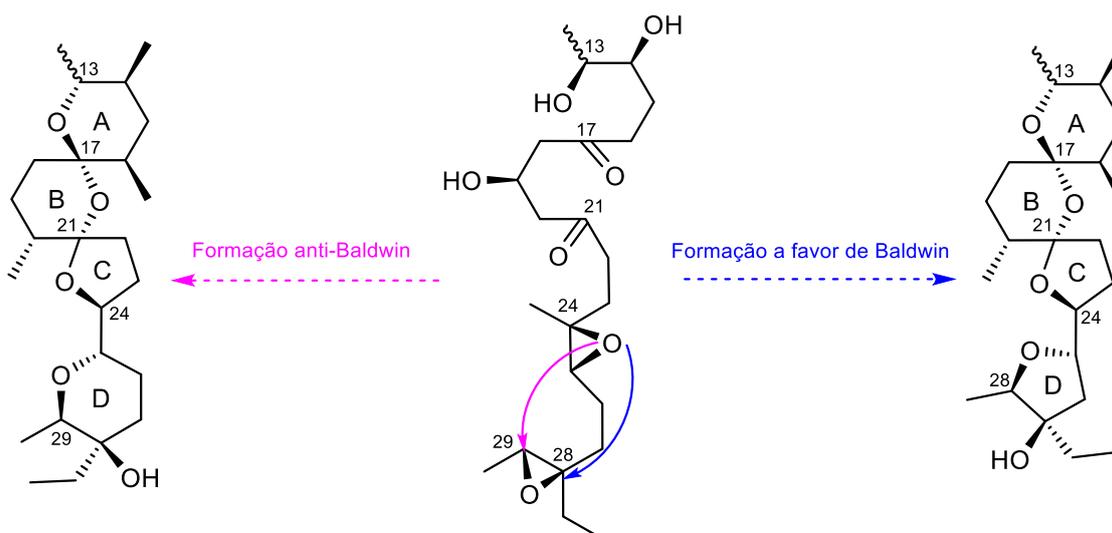


Figura 4. Estruturas obtidas na salinomicina seguindo as regras de Baldwin (azul) e anti-Baldwin (rosa).

Recentemente, diversos estudos mostraram que além de atividade antibiótica, a salinomicina também tem atividade contra células tronco de câncer, mas devido a alta toxicidade em células humanas seu uso está restringido na medicina veterinária. Pesquisas que obtiveram derivados da salinomicina adicionando diferentes substituintes nos carbonos C-1, C-19, C-20 e formando diferentes grupos ésteres derivados do grupo ácido carboxílico em C-1 permitiram concluir que a obtenção de derivados da salinomicina diminui a toxicidade em células humanas e aumenta a atividade anticâncer em células tronco e células de câncer pequenas<sup>6,15,16</sup>.

A possibilidade de obtenção de análogos da salinomicina por bioengenharia constitui a motivação para o entendimento da sua rota biossintética. Compreender a estereoespecificidade da reação de epoxidação originando o intermediário bis-epoxidado, que na sequência passa por etapas de ciclização anti-Baldwin promovida pelas enzimas B podem permitir a modificação desta rota para gerar estruturas análogas com menor toxicidade e aplicáveis na medicina humana.

### 3 OBJETIVOS

O objetivo principal do projeto é o estudo da enzima flavina monooxigenase SalC envolvida na oxidação estereosseletiva do intermediário linear insaturado pré-salinomicina. Esse intermediário foi isolado do mutante  $\Delta salC^5$ .

Os objetivos específicos compreendem:

- Expressão e caracterização funcional e bioquímica da enzima SalC;
- Análise estrutural *in silico* da enzima;
- Produção do intermediário pré-salinomicina;
- Estudos mecanísticos e estruturais enzima-substrato;
- Isolamento, purificação e caracterização do intermediário modificado enzimaticamente.

## 4 METODOLOGIA

### 4.1 Fermentação dos metabólitos, expressão heteróloga e purificação da enzima SalC

As metodologias usadas para a obtenção do intermediário do tipo policetídeo linear a partir de *Streptomyces albus* sem e com deleção do gene *salC*, a expressão e purificação da enzima SalC são metodologias padronizadas<sup>5</sup>.

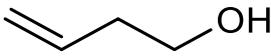
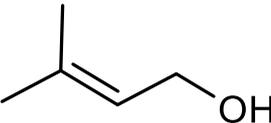
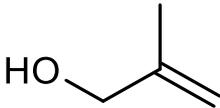
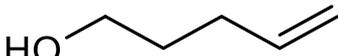
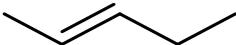
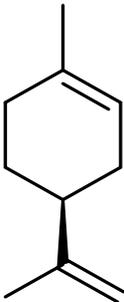
### 4.2 Ensaios estruturais da enzima SalC

A estrutura secundária da enzima SalC foi estudada usando a técnica de dicroísmo circular: 2 nmol de proteína foram dissolvidos em tampão Tris-HCl 20 mmol L<sup>-1</sup>, pH 8,0 e a solução irradiada com comprimentos de onda entre 196 nm e 260 nm usando um espectro polarímetro Jasco modelo J720. Fizeram-se provas de desnaturação e naturação térmica em rampas de temperatura de 25 °C até 85 °C. Os dados foram analisados usando a ferramenta online DICHROWEB<sup>17,18</sup>.

### 4.3 Ensaios de atividade enzimática *in vitro*

Para determinar o consumo de NADH pela proteína, foi usado um sistema de incubação contendo 3,7 µM de SalC, 240 µmol L<sup>-1</sup> de substrato e 6 mmol L<sup>-1</sup> de NADH e adicionado tampão de Tris-HCl 20 mmol L<sup>-1</sup> para completar um volume de 1 mL. A diminuição da concentração de NADH foi acompanhada usando um espectrofotômetro UV-Vis Biomate3 da Thermo-Electron no comprimento de onda de 340 nm durante 90 s. Para a análise de espectrometria de massas, aumentou-se o tempo e o volume de incubação para 4 horas e 25 mL respectivamente, o material obtido foi extraído usando acetato de etila e analisado usando CG/MS. Os substratos testados estão apresentados na Tabela 1.

Tabela 1. Substratos usados no teste de atividade enzimática da enzima SalC

Característica da molécula	Substrato utilizado
Alceno homoalílico sem substituinte na dupla ligação	 <b>3-buten-1-ol</b>
Alceno alílico com substituinte na dupla ligação	 <b>3-metil-2-buten-1-ol</b>
Alceno alílico com duplo substituinte	 <b>2-metil-3-buten-2-ol</b>
Alceno alílico com substituinte na dupla ligação	 <b>2-metil-2-propen-1-ol</b>
Alceno sem substituinte na dupla ligação	 <b>4-penten-1-ol</b>
Alceno sem substituinte na dupla ligação e ausência do grupo hidroxila	 <b>2-penteno</b>
Alceno cíclico	 <b>L-limoneno</b>

#### 4.3.1 Ensaio da transformação do intermediário pré-salinomicina

A oxidação do intermediário linear da biossíntese da salinomicina foi feita *in vitro* na presença da enzima SalC. Usaram-se 25 mL de meio de reação composto de Tris-HCl 20 mmol L<sup>-1</sup> pH 8,0, 2 mmol L<sup>-1</sup> NADH, 3,7 μmol L<sup>-1</sup> de enzima SalC e 1000 μL de fração de extrato obtido do cultivo da linhagem mutante de *Streptomyces albus-ΔsalC*. A reação foi feita *overnight* sob aquecimento com temperatura de 25 °C com agitação constante. O produto obtido foi extraído com acetato de etila em relação 1:1, o solvente foi removido e após evaporação foi diluído com metanol. Os produtos obtidos foram analisados usando um espectrômetro de massas acoplado a UPLC Waters, modelo TQD Quattro Micro API; fonte de ionização por electrospray (ESI); coluna de fase reversa InfinityLab Poroshell 120 C-18 2,7μm; 4.6x100mm. A gradiente de eluição seguiu as proporções de A e B a seguir: 0-10 min, 95% A/5% B; 15 – 20 min, 99% A/ 1% B; 20 – 40 min, 99% A/ 1%B; 40 – 45 min, 100%A, 0%B.

## 5 Resultados

### 5.1 Fermentação e extração do intermediário linear

Estudos anteriores revelam que a deleção do gene *salC* promove o acúmulo do intermediário linear pré-salinomicina, no qual a etapa de oxidação é inibida<sup>4</sup>. Esses estudos validam que SalC é a enzima responsável pela oxidação estereoespecífica do intermediário linear insaturado pré-saliomicina. Inicialmente foram testadas condições de cultivo para a produção de salinomicina e do intermediário linear utilizando a linhagem selvagem *Streptomyces albus* DSM 41398 (wt) e a cepa com a deleção do gene *salC* ( $\Delta salC$ ). A produção dos metabólitos foi acompanhada por espectrometria de massas, esperando-se a presença da sinal de salinomicina (750,49 g mol<sup>-1</sup>) e ausência do sinal do intermediário linear (721,03 g mol<sup>-1</sup>) no cultivo da linhagem wt, cujo acúmulo é então observado no cultivo da cepa em função da deleção do gene *salC* (Figura 4).

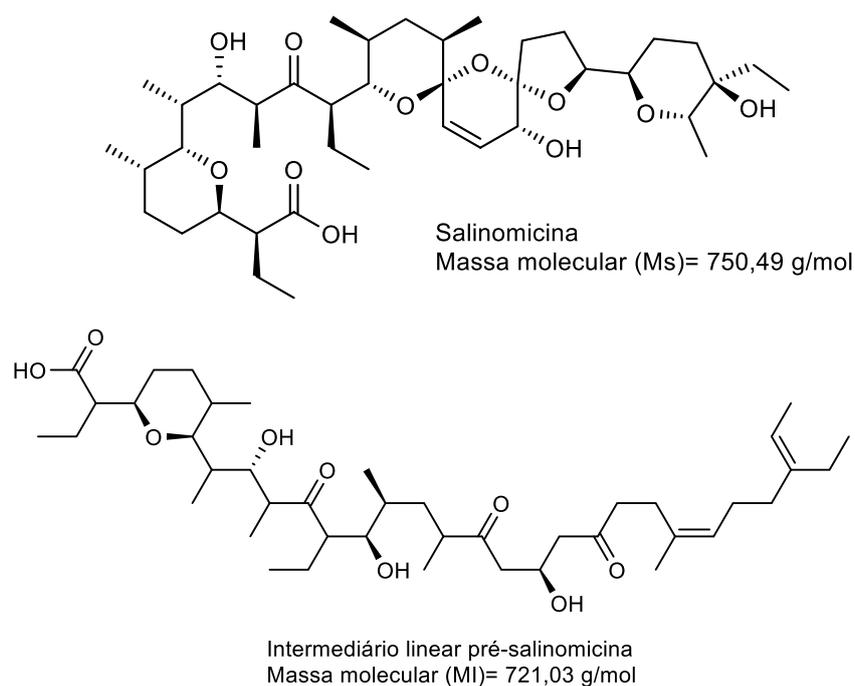


Figura 5. Estruturas da salinomicina (acima) e do intermediário linear (abaixo).

As cepas de bactérias  $\Delta salC$  e WT armazenadas em meio SFM sólido (Extrato de soja 20 g L<sup>-1</sup>, manitol 20 g L<sup>-1</sup>, agar 20 g L<sup>-1</sup>, MgCl<sub>2</sub> 10 mmol L<sup>-1</sup>) foram pre – inoculadas por 2 dias em meio SFM (Extrato de soja 20 g L<sup>-1</sup>, manitol 20 g L<sup>-1</sup>, MgCl<sub>2</sub> 10 mM) (seed) e depois transferidas para meio MP (Glicose 10 g L<sup>-1</sup>, amido solúvel 10 g L<sup>-1</sup>, peptona 2 g L<sup>-1</sup>, extrato de levedura 3 g L<sup>-1</sup>, soytone 3

g L<sup>-1</sup>, extrato de malta 2 g L<sup>-1</sup>, cloreto de sódio 5 g L<sup>-1</sup>, TES 25 mmol L<sup>-1</sup> pH 7,2) para fermentação por 14 dias sob agitação constante. Os metabólitos presentes no meio de cultivo foram extraídos após de centrifugação do meio de cultura (3000 rpm x 15 min) usando acetato de etila em relação 1:1 e depois da remoção do solvente, o extrato bruto foi diluído em metanol para análise por cromatografia líquida acoplado a espectrometria de massas (LC-EM) (Figura 5).

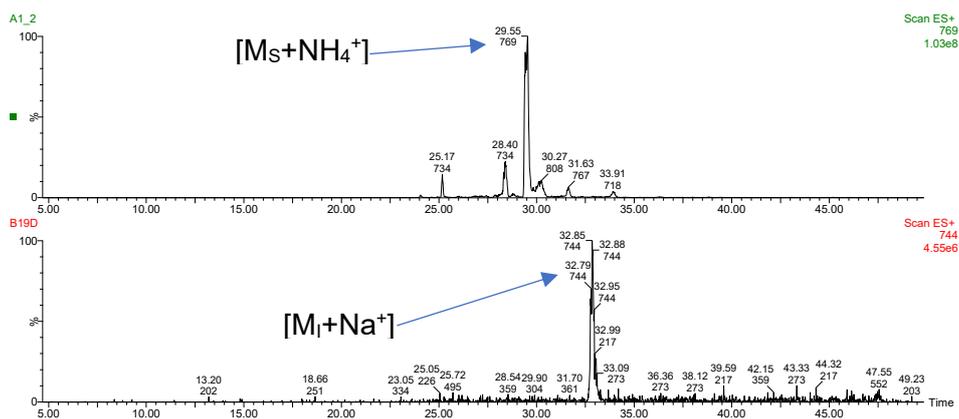


Figura 6. Cromatogramas dos extratos obtidos: A1\_2 refere-se ao cultivo de *Streptomyces albus* (wt) onde a sinal  $m/z$  769  $[M_S + NH_4^+]$  corresponde à salinomicina; B19D refere-se ao cultivo da bactéria *Streptomyces albus-Δsa/C*, onde a sinal de  $m/z$  744  $[M_I + Na^+]$  corresponde ao intermediário linear sodiado.

Nos cromatogramas mostrados há evidência da presença de salinomicina só no cultivo da cepa selvagem (sinal de  $m/z$  769  $[M_S + NH_4^+]$ ), enquanto a cepa mutante mostra a ausência do sinal de salinomicina e a presença de um sinal de  $m/z$  744 relacionado ao intermediário linear sodiado.

O extrato obtido do cultivo da cepa com deleção do gene *sa/C* foi semi-purificado passando-se por uma coluna Superclean™ LC-18 SPE de 3 mL, a coluna foi ativada com 3 mL de metanol e equilibrada com 6 mL de solução de metanol 5%, o extrato foi passado na coluna e a eluição foi guardada para análise posterior. Foram feitas 4 lavagens aumentando-se a concentração de metanol de 25% até 100% em incrementos de 25% de metanol. As frações eluídas foram analisadas usando LC-EM.

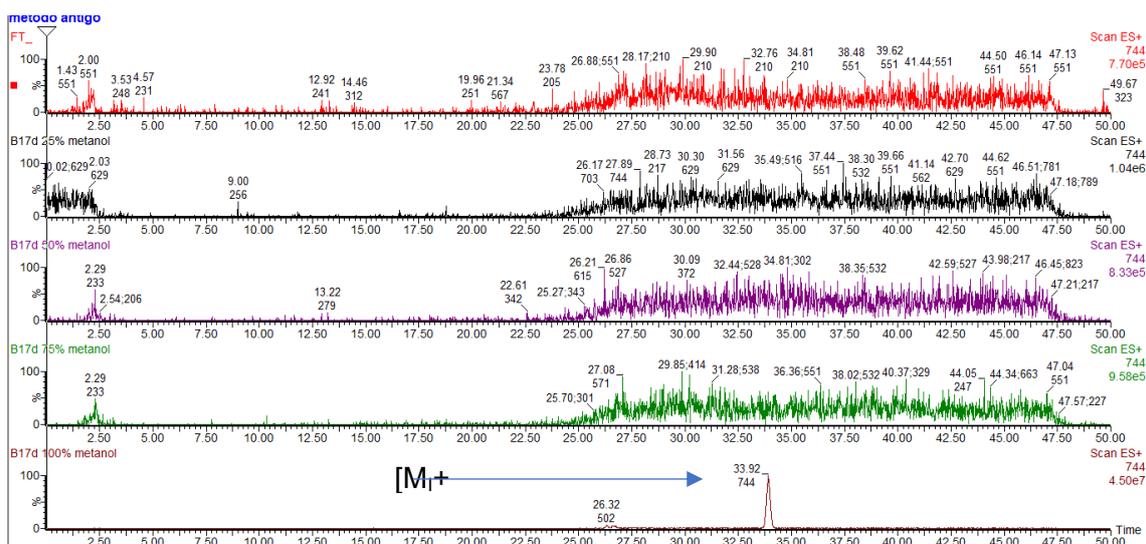


Figura 7. Comparação dos cromatogramas obtidos das análises de massas das frações eluídas por coluna Superclean™ LC-18 SPE observando-se o íon  $m/z$  744 referente ao intermediário linear pré-salinomicina.

Somente no cromatograma da fração 100% metanol (B17d 100% metanol) aparece o sinal de interesse ( $m/z$  744  $[M_1+Na^+]$ ). A obtenção de uma solução com quantidade reduzida de outros metabólitos especializados (fração 100%) permite que a incubação do intermediário com a enzima SalC seja mais fácil de ser estudada, já que a fração é a mais rica no composto de interesse.

### 5.1.1 Expressão heteróloga da enzima SalC

A expressão da enzima SalC foi feita usando como organismo heterólogo *Escherichia coli* BL21(DE3) contendo um plasmídeo no qual o gene *salC* foi inserido em um vetor pET29b(+), entre os sítios de NdeI e EcoRI, originando o plasmídeo pET29-salC, o qual foi mantido em *E. coli* DH10B. A integridade do plasmídeo foi verificada após análise de restrição utilizando-se as enzimas NdeI e EcoRI. A Figura 8 mostra o padrão das bandas obtidas após a digestão e corrida em um gel de agarose 1%, aonde é possível observar as bandas em ~6000 pb e ~1500 pb referentes ao plasmídeo e ao fragmento respectivamente. A construção correta do plasmídeo também foi avaliada por sequenciamento do fragmento utilizando-se os primers T7 e T7t.

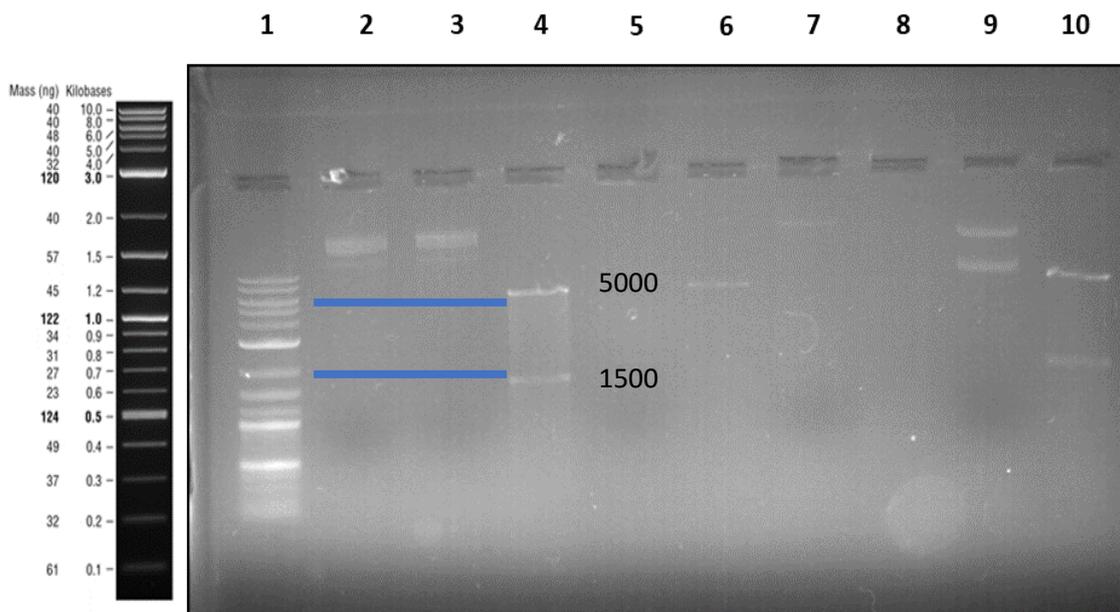


Figura 8. Gel de agarose (1%) para a prova da integridade do plasmídeo após transformação em bactérias de *E. coli* BL21(DE3) usando plasmídeos isolados das cepas de estoque (DH10B) e de produção (BL21(DE3)). Linha 1: Marcador de massa molecular; linha 2 e 3: plasmídeo isolado de *E. coli* DH10B; linha 4: plasmídeo isolado de *E. coli* DH10B após digestão com as enzimas de restrição EcoRI e NdeI; linha 5: plasmídeo isolado de *E. coli* BL21(DE3); linha 6: plasmídeo isolado de *E. coli* BL21(DE3) após digestão com as enzimas de restrição EcoRI e NdeI; linha 7: plasmídeo isolado de *E. coli* DH10B pré-transformação; linha 8: plasmídeo isolado de *E. coli* BL21(DE3) pré-transformação; linha 9: Plasmídeo isolado de *E. coli* BL21(DE3) após transformação e clivado pelas enzimas de restrição NdeI e EcoRI; linha 10: plasmídeo isolado de *E. coli* DH10B após transformação e clivado pelas enzimas de restrição NdeI e EcoRI.

Uma vez obtida uma transformação de *E. coli* BL21(DE3) com plasmídeo pET29-salC, foi realizado o experimento de expressão para obtenção de SalC. A Figura 9(a) mostra o acompanhamento do experimento de expressão em um gel de poliacrilamida no qual é comparada a expressão do gene feita em meio LB na presença de canamicina como marcador de seleção na presença e ausência de IPTG. Pode-se observar uma banda acima de 50 kDa mas a banda está presente tanto na fração solúvel sem IPTG como na fração solúvel com IPTG, para ter uma melhor resolução no gel e para esclarecer a produção da

proteína, diminuiu-se a concentração de poliacrilamida no gel de 12% para 7% e foi feita uma pré-purificação usando uma coluna His-Tag. O gel obtido depois da purificação é mostrado na Figura 9(b), onde pode-se observar a presença de uma banda de massa um pouco maior de 50 kDa isolada após a pré-purificação, corroborando a produção da proteína SalC suspendida em tampão fosfato 20 mmol L<sup>-1</sup> pH 8.0.

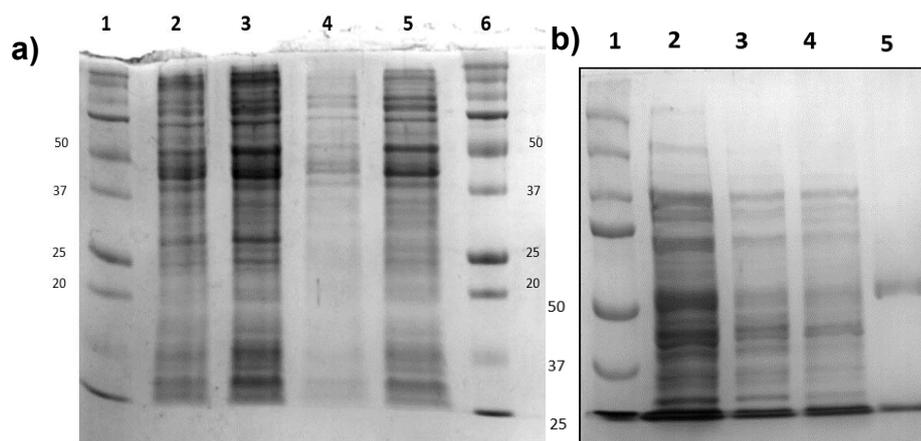


Figura 9. a) Gel de poliacrilamida (7%) feito para a prova da expressão do gene *sa/C*. Pode-se ver a presença de bandas de 50 kDa tanto na fração solúvel sem IPTG como na solúvel com IPTG; linha 1 e 6: marcador molecular linha 2: fração insolúvel sem IPTG; linha 3: fração solúvel sem IPTG; linha 4: fração insolúvel com IPTG; linha 5: fração solúvel com IPTG, b) Gel de poliacrilamida (12%) com diminuição da concentração de acrilamida. Observa-se a produção da proteína parcialmente purificada com massa cerca de 50 kDa (linha 5), linha 1: marcador molecular; linha 2: fração solúvel; linha 3: *flowtrought*; linha 4: Lavagem com imidazol 60 mmol L<sup>-1</sup>; linha 5: lavagem com imidazol 600 mmol L<sup>-1</sup>.

Com o protocolo estabelecido, a produção de enzima foi escalonada para realizar os ensaios de dicroísmo circular e de caracterização catalítica. Usaram-se 500 mL de meio LB com canamicina como antibiótico para a expressão do gene, obtivendo-se um rendimento da purificação de 0,08 mg mL<sup>-1</sup> de enzima no meio de cultura usado.

### 5.1.2 Ensaio de dicroísmo circular

Para o experimento de dicroísmo circular observou-se a necessidade de troca do tampão fosfato, uma vez que o mesmo absorve em valores de comprimento de onda baixos, promovendo a saturação do sinal no detector e dificultado medidas em valores entre 190 e 200 nm<sup>19</sup>. Para evitar a interferência, foi realizada a troca de tampão para Tris-HCl 20 mmol L<sup>-1</sup> e eliminou-se o NaCl da solução. A Figura 10 mostra o gel de poliácridamida (12%) feito para avaliar a presença de proteína depois do procedimento de troca de tampão usando uma coluna HiTrap™ Desalting, onde a banda perto de 50 kDa se mantém.

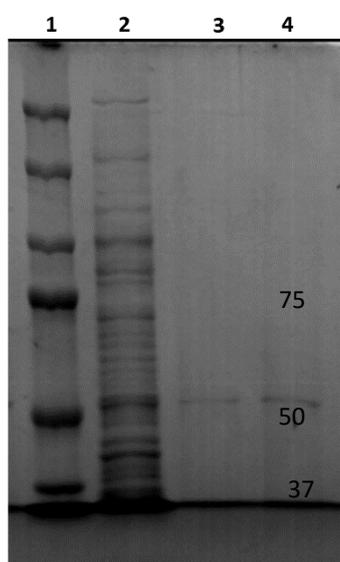


Figura 10. Gel de poliácridamida (12%) onde se apresenta a presença da banda maior a 50 kDa após da troca de tampão de fosfato 20 mmol L<sup>-1</sup> para Tris-HCl 20 mmol L<sup>-1</sup>. Linha 1: Marcador molecular; linha 2: Enzima SalC com tampão fosfato 20 mmol L<sup>-1</sup>, NaCl 300 mmol L<sup>-1</sup>; linha 3: Enzima SalC com tampão Tris-HCl 20 mmol L<sup>-1</sup>.

Fizeram-se ensaios de dicroísmo circular (CD) para deduzir a composição porcentual dos diferentes tipos de estruturas secundárias presentes na enzima. Para realizar a medição, foi necessário fazer diluições da enzima para evitar saturar o detector. A concentração ideal para o experimento foi de 0,1 mg mL<sup>-1</sup> em tampão Tris-HCl 20 mmol L<sup>-1</sup>. As faixas de comprimento de onda relevantes na análise estrutural da proteína compreendem 208 nm para  $\alpha$ -hélice, 215 nm para folhas- $\beta$  e 196 nm para estruturas randômicas<sup>19-21</sup>. Nessas condições foi obtido um espectro de CD em temperatura de 25 °C com 16 scans no total (Figura 11) onde pode-se apreciar que a enzima apresenta estruturas secundárias de  $\alpha$ -hélice, folhas- $\beta$  e arranjos randômicos.

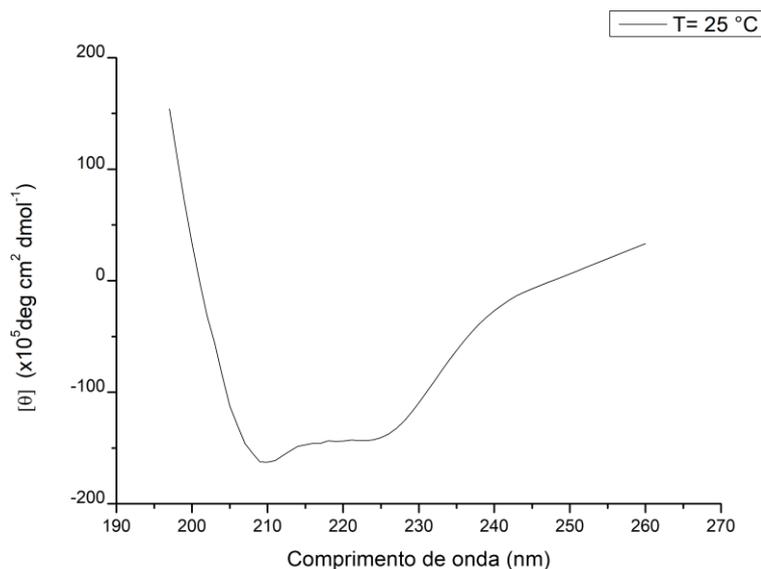


Figura 11. Espectro CD da enzima SalC adquirido em temperatura de 25 °C.

O espectro obtido a 25 °C foi analisado usando dois diferentes softwares, a ferramenta online DRICHROWEB, que usa o método K2D para calcular as porcentagens de  $\alpha$ -hélice, folhas- $\beta$  e estrutura randômica comparando o espectro CD com uma base de dados de espectros tanto teóricos, obtidos de estruturas armazenados no PDB como espectros experimentais de proteínas com estrutura já conhecida<sup>17,18,22,23</sup> e a ferramenta CNND, que usa uma rede neuronal com a capacidade de reconhecer os diferentes aportes da estrutura secundária, absorção da cisteína, cadeias laterais aromáticas, aporte do ruído de medição e erros não randômicos que modificam o espectro CD<sup>24</sup>. As porcentagens das conformações presentes na enzima obtidas a partir da deconvolução com os métodos já mencionados pode ser observada na Tabela 2.

Tabela 2. Contribuições das diferentes estruturas secundárias presentes na enzima SalC obtidas com os métodos K2D e CDNN

	<b>Enzima SalC (25°C)</b>	
	K2D	CDNN
<b><math>\alpha</math>-Hélice (%)</b>	23	16,2
<b>Folhas-<math>\beta</math> (%)</b>	24	33
<b>Randômico (%)</b>	53	48,6

Utilizou-se a análise de CD para avaliar a estabilidade térmica da enzima em uma faixa de temperatura compreendida entre os valores de 25 °C e 80 °C em intervalos de 5 °C. O gráfico obtido (Figura 12), compara os espectros obtidos, e evidencia que a enzima SalC tem uma faixa de estabilidade térmica compreendida entre as temperaturas de 25 °C e 50 °C, quando a enzima é aquecida até atingir os 55 °C começa o processo de desnaturação térmica, obtendo-se uma enzima totalmente desnaturada em temperaturas maiores que 70 °C.

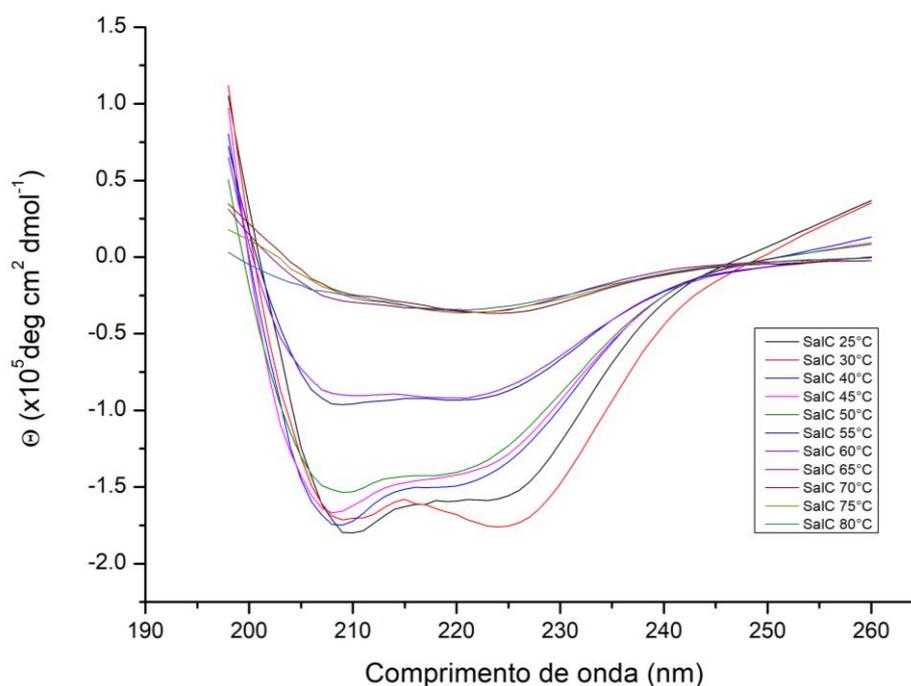
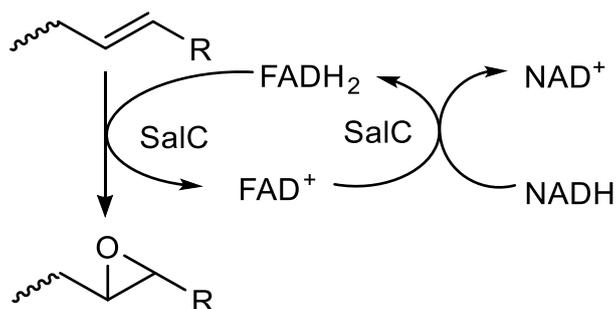


Figura 12. Espectro de dicroísmo circular da enzima SalC compreendendo as faixas de temperaturas de 25 °C a 80 °C.

### 5.1.3 Ensaios de atividade enzimática *in vitro*

A enzima SalC interage com substratos que tem insaturações na sua estrutura e como epoxidase, promovendo a adição de um átomo de oxigênio e originando um grupo epóxido<sup>4,7</sup>. Esta oxidação é feita pela redução do grupo FAD que doa elétrons para gerar as duas novas ligações. O FAD é reconstituído pela doação de prótons vindo da oxidação de duas moléculas de NADH, dita oxidação é feita também pela enzima SalC permitindo um novo ciclo de epoxidação<sup>9</sup> (Esquema 4).



Esquema 4. Ciclo da reação de epoxidação catalisada pela enzima SaIC

Experimentalmente foi possível observar que a enzima SaIC oxida NADH para  $\text{NAD}^+$  na presença de substratos que possui insaturações em sua estrutura (Figura 13). A observação da oxidação de NADH é uma prova indireta da conversão das moléculas testadas para seu respectivo epóxido (Esquema 5) promovida pela enzima SaIC, onde um grupo do tipo oxirano é gerado a partir da oxidação das insaturações presentes nos substratos precursores. O uso de substratos solúveis em água tem um consumo de NADH superior a 90%, mas quando o substrato tem uma natureza hidrofóbica, o consumo de NADH diminui.

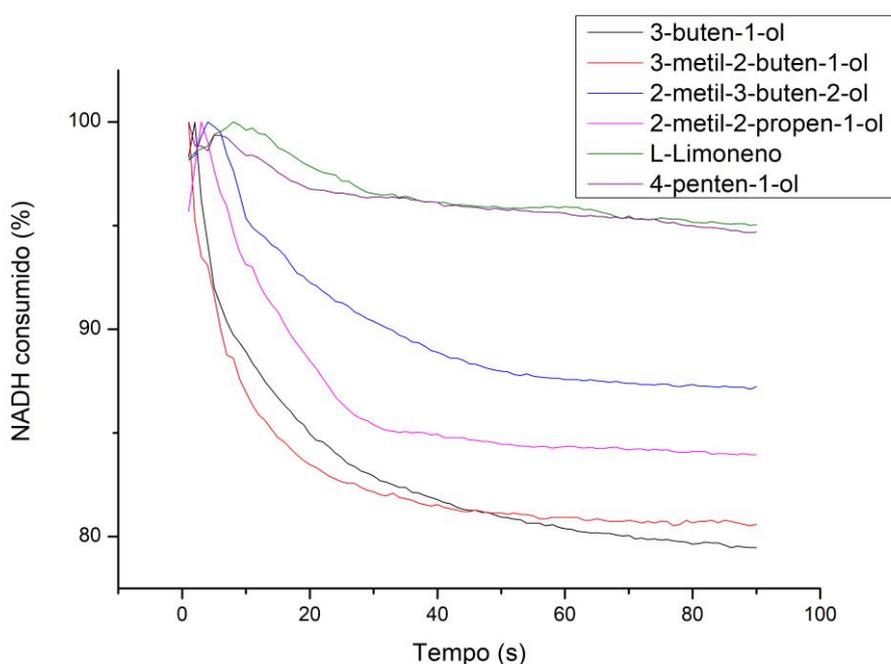
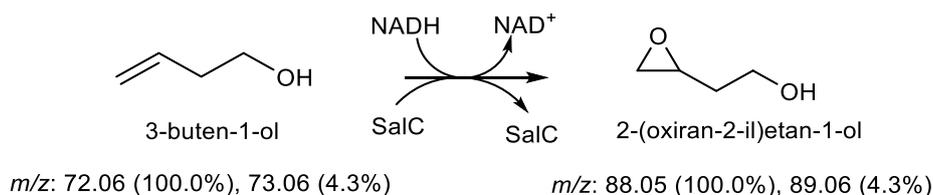


Figura 13. Variação do consumo de NADH feito pela enzima SaIC com os substratos 3-buten-1-ol (preto); 3-metil-2-buten-1-ol (vermelho), 2-metil-3-buten-2-ol (azul), 2-metil-2-propen-1-ol (magenta), L-limoneno (verde), 4-penten-1-ol (roxo).



Esquema 5. Produto esperado após epoxidação enzimática da molécula 3-buten-1-ol.

O cromatograma apresentado na Figura 14 mostra os produtos obtidos após incubação da molécula 3-buten-1-ol com a enzima SalC. A interação da enzima com o 3-buten-1-ol gera a molécula 2-(oxiran-2-yl)-etan-1-ol com massa de 88,05 g mol<sup>-1</sup>. A fragmentação do pico em 4,323 (baixa resolução) gera um padrão de fragmentos compatível com o esperado para 2-(oxiran-2-yl)-etan-1-ol (Figura 14).

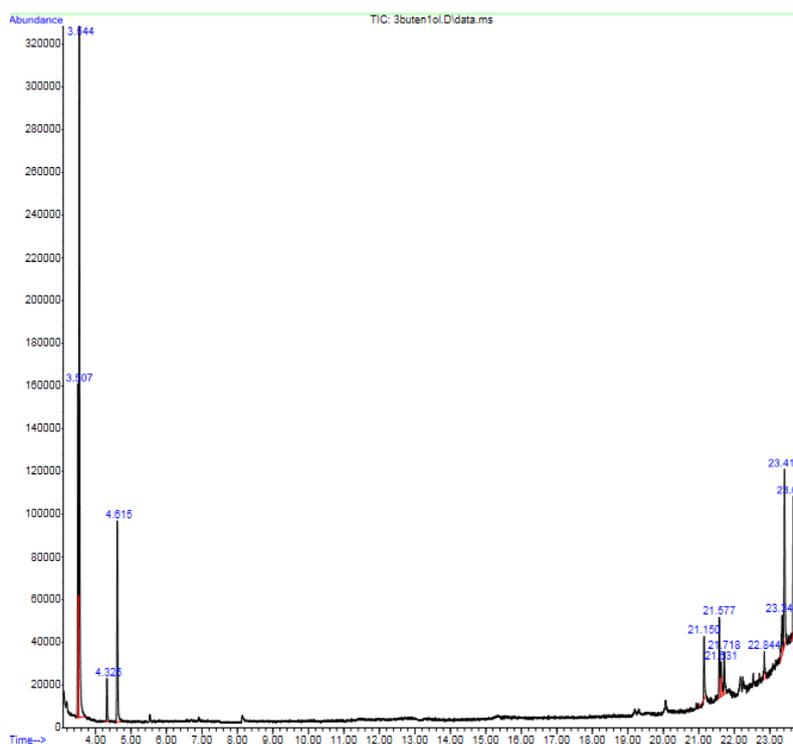


Figura 14. Cromatograma obtido da análise da transformação enzimática do 3-buten-1-ol usando SalC com uma proporção substrato:NADH 1:1. O pico eluído em 4,323 s mostra um padrão de fragmentação que pode ser atribuído para o epóxido esperado.

O sinal de  $m/z$  43,1 mostra o fragmento de óxido de etileno com uma diferença de  $m/z$  0,08, o sinal de  $m/z$  56,1 pode mostrar o fragmento resultante

da eliminação do grupo metanol da molécula (diferença de massa 0,93). O fragmento  $m/z$  73,0 representa a subtração de parte do grupo oxirano, com uma diferença de 1,04. A massa molecular é mostrada no sinal  $m/z$  de 87, com uma diferença de 1,05.

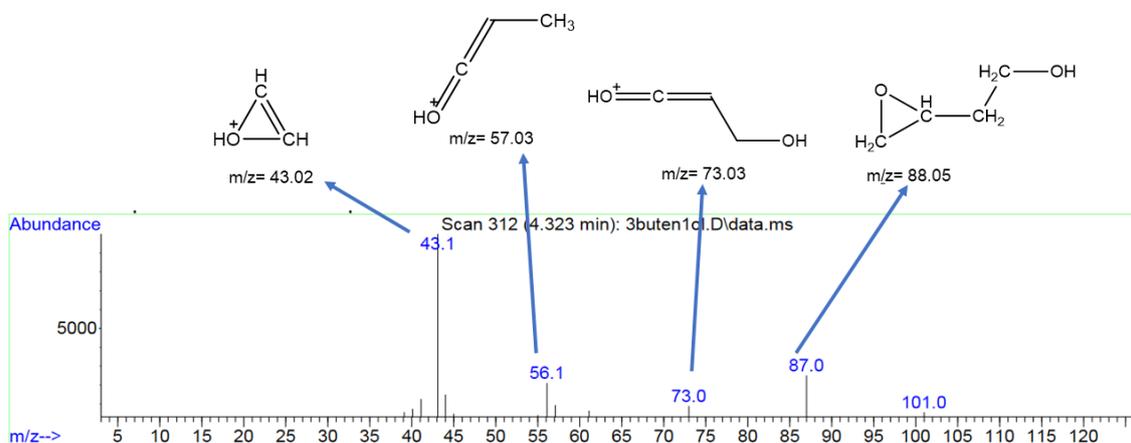


Figura 15. Espectro de massas gerado para o 2-(oxiran-2-il)-etan-1-ol

A importância da insaturação na reação foi avaliada usando a ferramenta analítica ANOVA do software ORIGINPro 8.5. A absorvância medida foi convertida para porcentagem de NADH consumido em 90 s de reação, compararam-se a porcentagem de NADH consumido usando os substratos 3-buten-1-ol, 3-metil-2-buten-1-ol, 2-metil-3-buten-2-ol, 2-metil-2-propen-1-ol, L-limoneno e 4-penten-1-ol. Para cada reação foram feitas 4 repetições e as médias das repetições foram analisadas usando o teste de Tukey<sup>36</sup> e para avaliar a igualdade das variâncias usou-se o teste de Levene<sup>36</sup>, juntado as duas informações pode-se concluir se essa variável, no caso o substrato, tem diferenças estatisticamente significativas. A Figura 16 mostra a comparação entre médias (mean) e desvios padrão das variáveis (factors) estudados. Os dados mostram que para os fatores estudado existe uma diferença estatisticamente significativa, especificamente para os substratos L-limoneno e 4-penten-1-ol ( $p < 0,05$ ). A ausência do grupo hidroxila nessas estruturas gera uma diminuição no momento dipolar da molécula, o que pode gerar uma interação menor com a enzima e assim evitando que haja interação com o sítio catalítico.

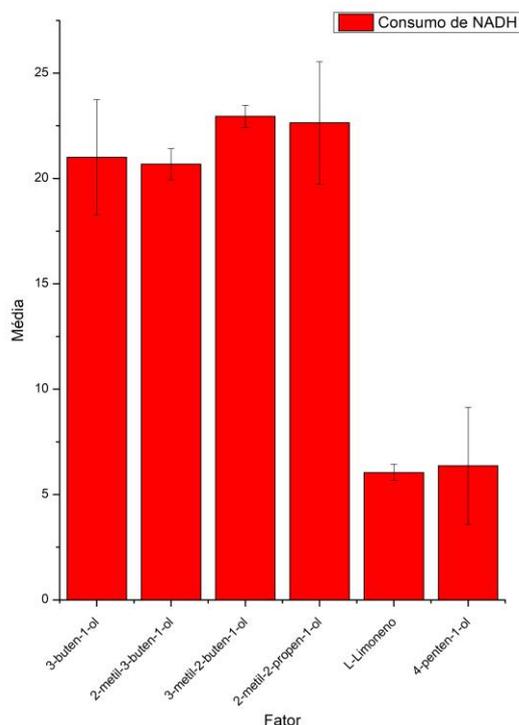
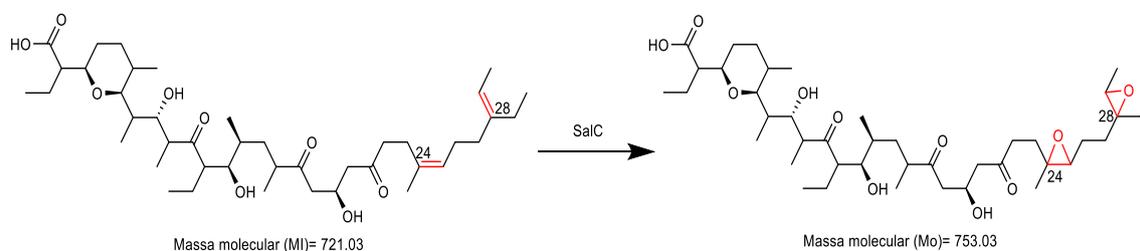


Figura 16. Comparativo do consumo de NADH dos substratos. A diminuição do consumo de NADH do L-limoneno e 4-penten-1-ol é estatisticamente significativa ( $p < 0.05$ )

A incubação do extrato obtido do cultivo da cepa com a deleção do gene *salC* foi estudada usando espectrometria de massas de baixa resolução. O esquema 6 mostra a possível mudança proposta pelo Yurkovich e colegas<sup>7</sup>, ditas mudanças são resultantes da presença de grupos do tipo oxirano resultantes da oxidação das insaturações localizadas nos carbonos 24 e 27 da molécula<sup>7</sup>.



Esquema 6. Oxidação do intermediário linear pré-salinomicina catalisado pela enzima SalC.

Finalmente, foi feito o estudo de oxidação enzimática do intermediário linear. Esperou-se nesse experimento observar a presença de epóxidos adicionais resultando nas massas dos intermediários sodiados de  $m/z$  744 [ $M_I+Na^+$ ] e  $m/z$  776 [ $M_O+Na^+$ ] no caso da oxidação em uma única etapa e considerando-se que o substrato não seja completamente convertido ou um único sinal de  $m/z$  776 [ $M_O+Na^+$ ]. Finalmente, ausência do sinal de  $m/z$  744 [ $M_I+Na^+$ ] mostra um consumo completo do substrato e a presença do sinal  $m/z$  776 [ $M_O+Na^+$ ] mostra uma reação completa com consumo dos substratos/intermediários ou uma reação em uma única etapa.

Com o objetivo de comprovar a hipótese de oxidação nas insaturações da molécula, a incubação foi feita *overnight* com excesso de enzima e na presença de NADH. Os espectros obtidos (Figura 17), mostram a ausência do sinal  $m/z$  740 [ $M_I+Na^+$ ] e a presença de sinais de  $m/z$  776 [ $M_O+Na^+$ ] no cromatograma. Os resultados mostram que o intermediário linear foi completamente convertido para o produto oxidado, mas deve-se fazer uma análise de fragmentação (MS/MS) desse pico para validar essa observação.

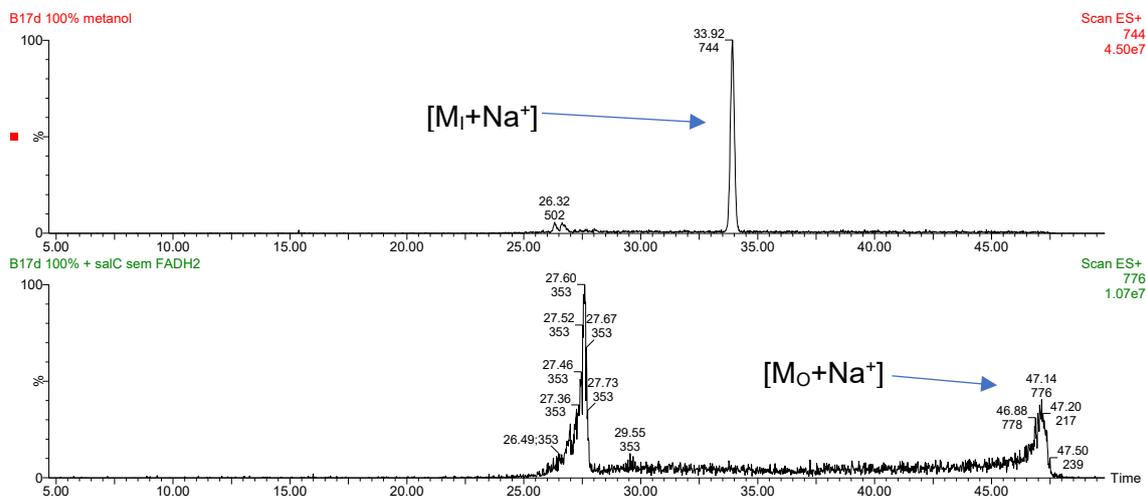


Figura 17. Comparação dos cromatogramas obtidos: i. fração B17d (100% metanol) onde está presente a sinal do intermediário linear ( $m/z$  744 [ $M_I+ Na^+$ ]); ii. fração de extrato que contém o intermediário linear com a enzima SalC (B17d 100% + SalC sem FAD) onde se observa o sinal de  $m/z$  776 [ $M_O+ Na^+$ ].

Com as informações obtidas e juntando-as com as contribuições feitas por Yurkovich e colaboradores<sup>7</sup> a rota biosintética da salinomicina pode ser explicada. A síntese começa com a extensão de um grupo acila através de 14

módulos que contêm domínios de ceto-redutase (KR), desidratases (DH), enoilredutase (ER), cetosintases (KS), descarboxilase (KSQ) e aciltransferase (AT) agrupados em 9 genes modulares (salAI – salAIX). Após a elongação da cadeia carbônica, a enzima SalX funciona como carregador do policetídeo linear liberado e o primeiro grupo pirano é formado pela enzima SalBIII e então, a enzima SalC oxida o intermediário linear, gerando um bis-epóxido ao adicionar dois grupos epóxi nas insaturações presentes nos carbonos C24 e C28. As enzimas SalD e SalE receberam o intermediário oxidado e adicionaram um grupo hidróxi- no C20. Finalmente, as enzimas SalBI e SalBII formam o núcleo rico em oxigênio constituído por dois piranos e um furano em um sistema bisespirocetálico, característico da salinomicina e um pirano com o qual a síntese da molécula é finalizada.

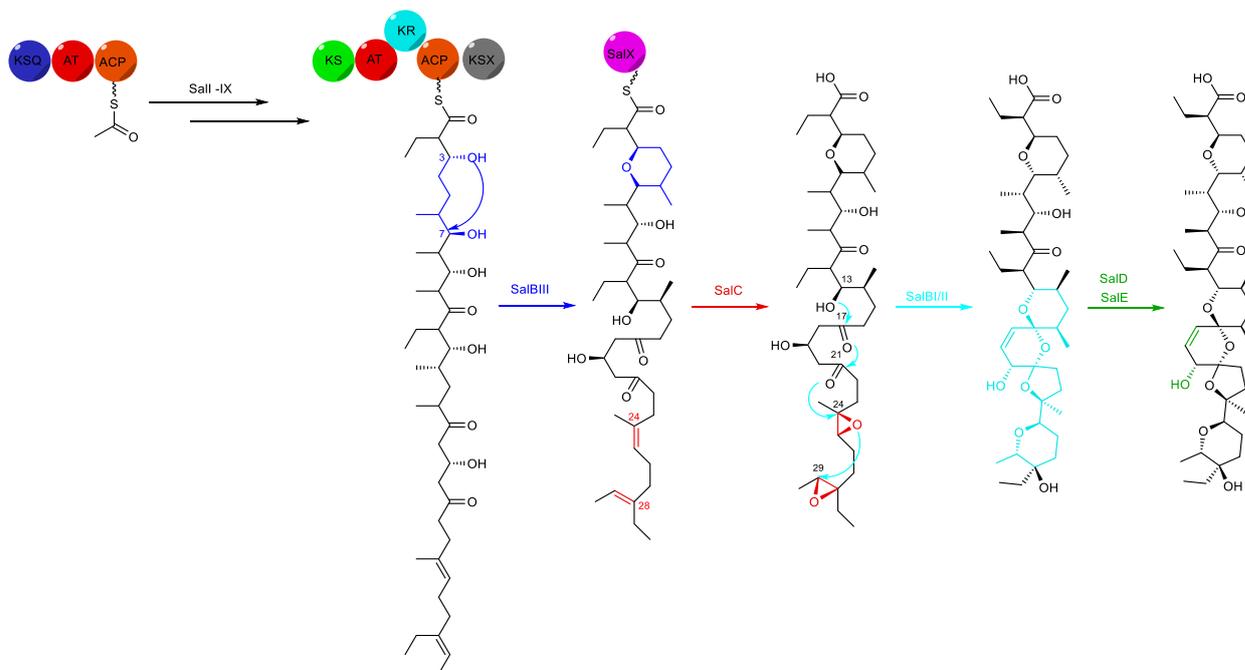


Figura 18. Rota biosintética da salinomicina

## 5.2 Estrutura tridimensional da enzima SalC

### 5.2.1 Estudo da sequência de aminoácidos da enzima SalC

A sequência da enzima SalC está armazenada no site do NCBI sob o código CCD31908.1.

```
>CCD31908.1 epoxidase [Streptomyces albus subsp. albus]
MPAADGRVTHLMTGETHAVVLGGGLTGMLTSAVRLARHLDKVTVIERDVLPEGPELRKGVQP
ARHAHLLWSGGARAIDSVLPGTVKQLIAEGAHRLLYLPDVVWLTPHGWQHRFSGSQFMVTC
SRALLDWVRRQALAEPKITVRQETEVLLGGLGGAGQVTGVRLRDRSGESTELAADLVVDAG
GRASALRRWLPELGLDQVEEDVVDGIAAYATRVFKAPAPVAQGFPMVNIAAAPGLGKPGQN
GALVPIEDGKWLVTLAGTRGGEPTDDDSFLDFARGLRHPVLADLLERAEPVLPVKGSRSTVN
RRLYYDRVANWPDGLLVLGDALAAFNPVYGHGMSCSALSALDAELGRSGLAPGMVQAV
LQKVAKVVDWPWLLATTQDICYPGTVTAQDPRIAPRGDQEQFADLLSTAALHDPVVSAAA
MQVTALAAPVSSLESPHLVAALRKGAHEPLTAPPFKDAELAVLDGDAAGAAASA
```

Figura 19. Sequência de aminoácidos da enzima SalC.

A enzima SalC possui 484 resíduos de aminoácidos com uma massa molecular de 51,04 kDa e um ponto isoelétrico teórico de 6,02. Analisando a sequência da enzima, observa-se uma porcentagem de resíduos de Tyr (1,2%), Phe (1,9%) e Trp (1,7%) que absorvem luz de 280 nm (total de 4,8%) que proporcionam a enzima um coeficiente de extinção molar de  $53065 \text{ M}^{-1}\text{cm}^{-1}$ .

O alinhamento da sequência com outras enzimas da família *c* que participam na epoxidação de intermediários lineares na biossíntese de diferentes ionóforos mostrou que a sequência tem uma cobertura que varia entre 87% - 95% comparadas às suas homólogas sendo que a identidade varia entre 53% - 74%

SalC/1-484	1	MPAADGRVTHLMTGETHAVVLGGGLTGMLTSAV	LARHLDKVTVIERDV	LPEGPELRKGV	59
Monensin/1-477	1	MTAETRAVVIGGGLAGMLASTV	LARHVDRVTVVERDVY	PEGPELRKGV	59
LasC/1-472	1	MTNTRSAVVLGGGMAGMLVSSML	ARHVGSVTVIDRDAF	PAGPDLRKGVP	59
SalC/1-484	60	PQARHAHLLWSGGARAI	DSVLPGTVKQLIAEG	AHRLYLPRD	VVWLT
Monensin/1-477	60	GGARAI	ESLLPGTLKALADEGA	AHRLYLPRDI	VWF
LasC/1-472	60	GGAR	IVEELLPGTTDRLLGAG	AHRIGIPD	GGVSYTAY
SalC/1-484	119	MVTC	SRALLDWVRRQALAE	PKITVRQ	ETEVLG
Monensin/1-477	119	VVRE	QALKNPRIEVRAQ	TGVTGLTGS	ADRVTG
LasC/1-472	119	TVRE	EETLREERIALVEK	TEVLALLGD	AGRV
SalC/1-484	178	L	VVDAGGRASALRRWL	PELGLDQVEED	VVDSG
Monensin/1-477	178	ML	KWLP	PELGLPEVEE	DVVD
LasC/1-472	178	PS	KRLLAELGL	PAPPEEF	VDSGM
SalC/1-484	237	G	L	GKPGQNGALVPI	EDGK
Monensin/1-477	237	P	I	EGGRWLVLAG	TRGA
LasC/1-472	237	L	M	PIEDGRWIVTV	SGTRG
SalC/1-484	296	L	G	PVKGSRSTVNRRL	YYDRVAN
Monensin/1-477	296	R	R	YYERLATWPDGL	LVLDALAA
LasC/1-472	296	T	V	NRRLLHYDRLAT	WPEGLV
SalC/1-484	355	G	R	SGLAPGMVQAVL	QKVAKVVD
Monensin/1-477	355	A	R	HAVAKVVD	PWLLATTQ
LasC/1-472	355	A	R	AQAIAVAVDD	AWVLATSHD
SalC/1-484	414	L	L	STAA	LHDPVVSAA
Monensin/1-477	414	N	A	AAMQV	TALAA
LasC/1-472	414	Q	V	NR	AAVALNTLS
SalC/1-484	473	L	D	G	AAGAAASA
Monensin/1-477	473	A	L	T	S
LasC/1-472		.	.	.	.

Figura 20. Alinhamento entre as sequencias das enzimas SalC, monensina epoxidase (MonC), e LasC (lasalocida epoxidase).

A proteína SalC é produzida pela expressão do gene localizado entre as posições dos nucleotídeos 741649 e 743268 do genoma do organismo *Streptomyces albus* sp. *albus* DSM 41398, a sequência de aminoácidos está disponível no banco de dados NCBI com o código CCD31908.1. O gene tem um tamanho de 1425 pb, codifica para uma proteína de 484 aminoácidos e possui uma massa de 51,8 kDa e um *pI* de 6,02<sup>25</sup>.

Utilizou-se a ferramenta online I-TASSER para analisar a estrutura tridimensional, flexibilidade, sítios conservados, sítios ativos, locais de ligação, proteínas semelhantes e classificação enzimática da enzima SalC<sup>26–28</sup>. A Figura 15 mostra o Fator B normalizado (B-Factor Profile, BFP) que indica a flexibilidade

de cada resíduo presente na enzima, onde valores menores de 0 indicam resíduos menos estáveis em condições experimentais<sup>29</sup>. Segundo a Figura 21, a enzima SalC tem uma alta flexibilidade, principalmente nos locais onde a estrutura de  $\alpha$ -hélice está presente. Valores de BFP positivos são encontrados principalmente nos extremos da enzima e alguns locais onde a estrutura randômica é predominante. Os sítios de ligação com FAD e os sítios ativos são mostrados como locais flexíveis possivelmente para acomodar-se mais facilmente ao substrato linear e extenso e para interagir mais facilmente com o cofator.

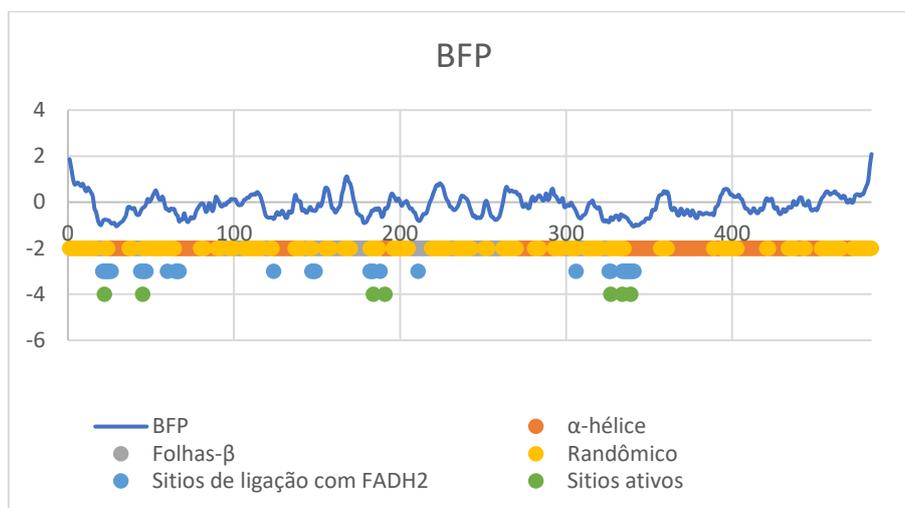


Figura 21. Valores de BFP calculados para a enzima SalC comparados com as estruturas secundárias previstas e os sítios ativos e de ligação com FAD.

A ferramenta I-TASSER também gerou 5 diferentes modelos usando o software SPICKER, os modelos gerados foram classificados com vários critérios<sup>26-28</sup>:

- **C-Score**: Valor entre  $[-5,2]$ , onde valores maiores mostram estruturas de maior confiabilidade, em outras palavras, mostra a qualidade do modelo gerado;
- **TM-score (Template Modeling score) e RMSD-score (Simply Root-Mean-Square Deviation)**: Medem a similaridade entre uma estrutura simulada e uma estrutura nativa conhecida, mas no caso de proteínas cuja estrutura é desconhecida, são comparados os valores de C com os modelos previstos. O valor de TM e RMSD mostram a qualidade do alinhamento e da topologia global da proteína. Valores de TM  $> 0,5$

indicam um modelo de topologia confiável e valores  $< 0,17$  mostram uma similaridade randômica;

- **Densidade do cluster:** Define-se como o número de réplicas de baixa temperatura em unidade de espaço, mostrando quanto muda o cluster obtido de SPICKES no momento de fazer pequenas deleções da estrutura proteica. Um valor alto da densidade de cluster mostra que a estrutura é simulada uma maior quantidade de vezes, o que representa uma melhor qualidade do modelo.

Os modelos gerados e os valores calculados e mencionados acima estão apresentados na Figura 22 e na Tabela 1A (Anexos).

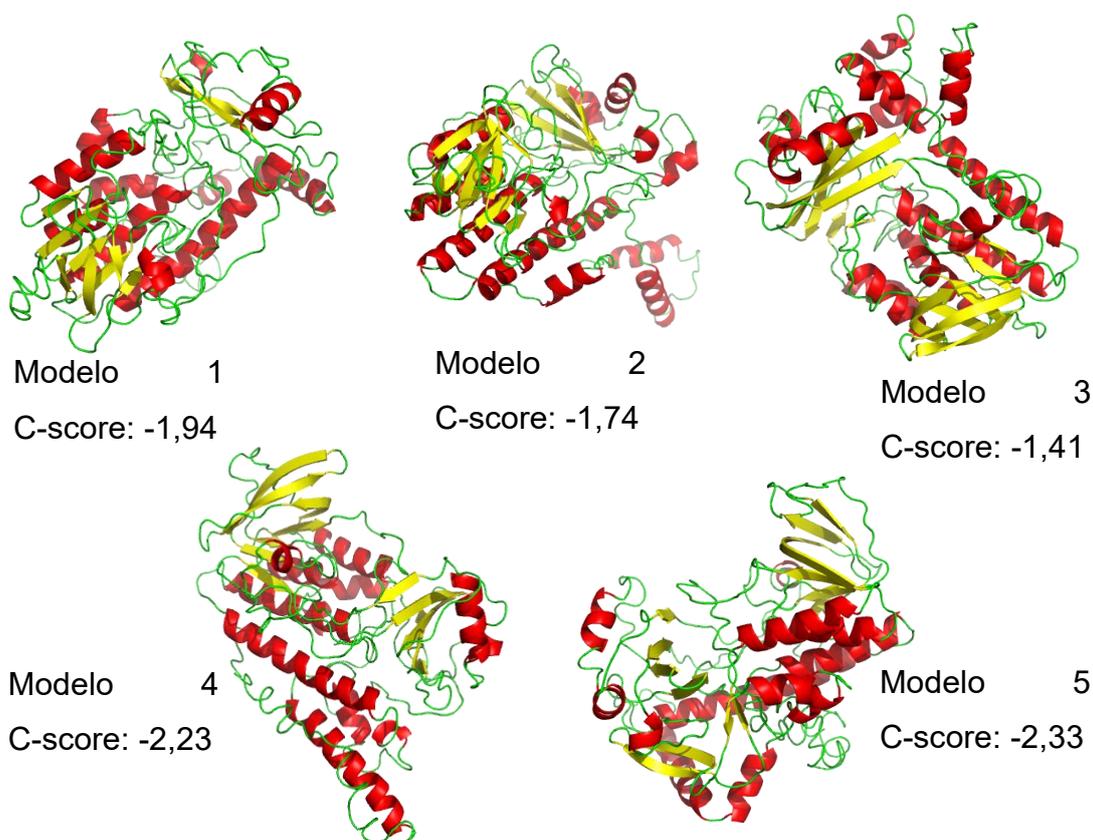


Figura 22. Modelos gerados pela ferramenta I-TASSER para a sequência de aminoácidos da enzima SalC com seus respectivos valores de C-score. As imagens dos modelos foram feitas usando a ferramenta PyMOL, amarelo foi atribuído para as folhas- $\beta$ , vermelho para as  $\alpha$ -hélices e verde para as estruturas randômicas.

É possível observar em todas as estruturas obtidas uma conformação majoritária de estrutura do tipo randômica, seguido da conformação de  $\alpha$ -hélice

em concordância com o obtido com os valores BFP, uma proteína de alta flexibilidade. Também se observa a presença de folhas- $\beta$  isoladas e apenas o modelo 5 mostra uma distribuição não uniforme dessas estruturas. Os outros modelos mostram que as folhas- $\beta$  estão presentes principalmente em duas localizações da enzima. Em geral, os modelos simulados concordam com os dados de BFP: enzimas com uma estrutura altamente flexível e pouca contribuição de folhas- $\beta$ .

Os valores de C-score (Tabela 1A Anexos) variam entre -1,41 até -2,23, sendo o modelo 3 quem tem o maior valor de C-score com -1,41. Entretanto, o modelo 3 não foi considerado como o principal já que o número de *decoys* é menor, sendo de 819 para o modelo 1 e 600 para o modelo 3, uma maior quantidade de *decoys* permite guiar mais eficientemente a construção da estrutura da proteína. Uma vez que os valores de C são negativos para todos os modelos e considerando-se que o valor de TM para o modelo 1 é menor que 0,5 e que os valores comparados entre os modelos 1, 2 e 3 não permitem concluir qual é o modelo que melhor se acomoda para a enzima SalC conclui-se que é preciso comparar os modelos gerados com os dados que foram medidos em experimentos *in vitro* para concluir qual é a estrutura que representa mais adequadamente SalC.

Calcularam-se a porcentagem de  $\alpha$ -hélice, folhas- $\beta$  e estrutura randômica dos modelos gerados pelo software I-TASSER usando duas estratégias diferentes: uma medição direta, contando a quantidade de resíduos em cada tipo de estrutura secundária e relacionando-o com o total de resíduos da proteína e usando o software online apresentados a seguir:

- DSSP: Ferramenta mais usada para calcular as contribuições das estruturas secundárias presentes em enzimas, baseia-se nas ligações de hidrogênio entre os diferentes resíduos<sup>30</sup>;
- STRIDE: Software que considera tanto as ligações de hidrogênio como os ângulos  $\phi$  e  $\psi$  dos resíduos presentes na proteína gerando uma quantificação mais sensível<sup>30</sup>;
- SECSTR: Assim como o software DSSP, baseia-se nas contribuições feitas pelas ligações de hidrogênio entre os resíduos de aminoácidos

presentes na enzima mas adiciona a quantificação das  $\pi$ -hélices presentes na estrutura<sup>30</sup>.

- XTL-SSTR: Calcula a contribuição de cada conformação estrutural baseado nas ligações de hidrogênio e interações amida-amida de cada resíduo. As distâncias e os ângulos são calculados a partir da geometria dos aminoácidos<sup>30</sup>;
- P-SEA: Baseia-se nos carbonos  $\alpha$  da estrutura protéica<sup>30</sup>;
- PALSSE: Usando os carbonos  $\alpha$ , baseia-se nos ângulos gerados entre os carbonos  $C_{\alpha i}$ ,  $C_{\alpha i+1}$ ,  $C_{\alpha i+2}$  e  $C_{\alpha i+3}$ <sup>31</sup>;
- STICKS: Baseia-se na localização dos carbonos  $\alpha$  e ligações de hidrogênio; pode detectar folhas- $\beta$  que outros softwares como DSSP considerariam como ausência de estrutura<sup>32</sup>;
- Simulação de dicroísmo circular (Simulação CD): A partir de um espectro de dicroísmo circular simulado usando a ferramenta PDBMD2CD, calculou-se a porcentagem de contribuição de cada tipo de estrutura secundária usando o fundamento teórico de DSSP<sup>33</sup>.

A Tabela 2A (Anexos) mostra as porcentagens de estrutura secundária obtidas a partir dos diferentes softwares. É possível observar que os diferentes métodos não diferem muito nos cálculos de estruturas do tipo  $\alpha$ -hélice e randômica, mas as porcentagens de folhas- $\beta$  variam consideravelmente (até 21%). Esses resultados podem ser em função da presença de folhas- $\beta$  isoladas, que são ignoradas por diferentes softwares como aqueles baseados em DSSP (DSSP, PALSSE, Simulação CD).

## 5.2.2 Previsões *in silico*

### 5.2.2.1 Enzimas estruturalmente análogas

Utilizando o programa de alinhamento TM-Align e a ferramenta I-TASSER estabeleceu relações de comparação entre estruturas de proteínas disponíveis no Protein Data Bank (PDB) e o modelo 1. A Tabela 3 mostra a identidade entre as duas estruturas, a porcentagem de cobertura da sequência e mostra as 5 proteínas análogas estruturalmente junto com os valores de TM<sup>34,35</sup>.

Tabela 3. Enzimas estruturalmente análogas com o modelo 1 da enzima SalC

<b>Classificação</b>	<b>PDB</b>	<b>Nome</b>	<b>Classificação enzimática</b>	<b>TM-score</b>	<b>Identidade (%)</b>	<b>Cobertura (%)</b>
<b>1</b>	6j0zC	AlpK	Óxido-redutase	0,1581	16	83,7
<b>2</b>	2qa1A	PgaE	Oxido-redutase	0,1569	16,5	83,7
<b>3</b>	2qa2A	CabE	Oxido-redutase	0,1562	15,5	82,6
<b>4</b>	3ihgA	Aklavinone-11 hidroxilase	Flavine Oxido-redutase	0,2280	13	86
<b>5</b>	5kowA	Rifampicina monooxigenase	Oxidoreductase	0,1522	15,1	83,1

As enzimas com analogia estrutural caracterizam-se pela mesma classificação como óxido-redutases. Entretanto é possível notar um valor de identidade baixo em relação ao modelo 1 da enzima SalC o que pode se explicar pelo fato de que os modelos têm um índice de confiança baixo dilatando erros no momento de identificar análogos estruturais. Adicionalmente os valores de identidade das enzimas com relação estrutural tem um valor baixo de identidade (máximo de 16,5% com a enzima PgaE) indicando que as estruturas têm pouca relação. Usou-se o software PyMOL para gerar uma imagem comparando cada enzima com o modelo 1 da proteína SalC.

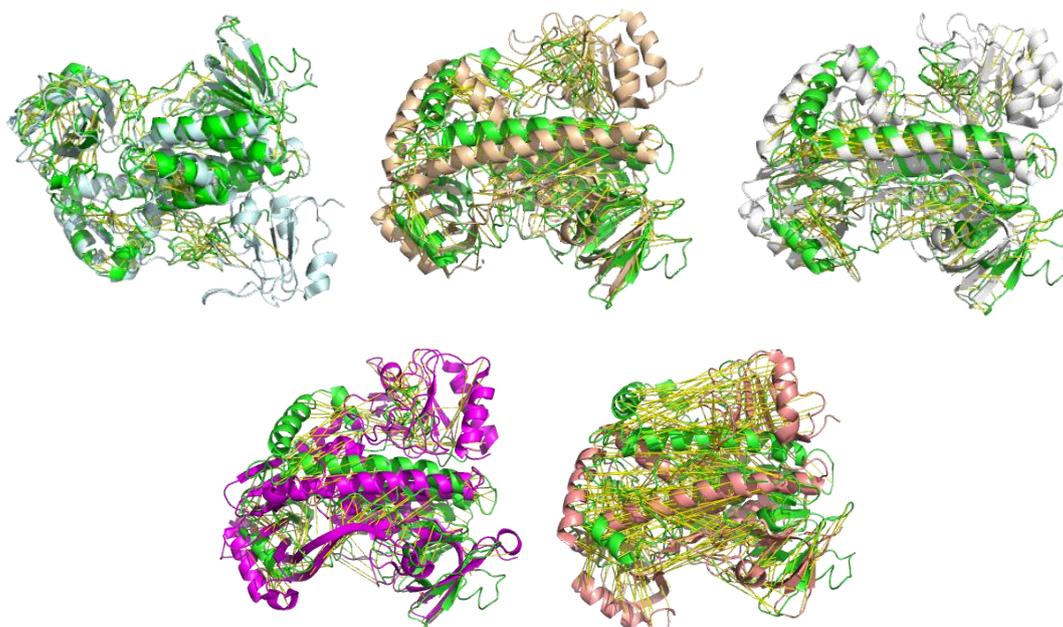


Figura 23. Alinhamentos do modelo 1 da enzima SalC (verde) com a enzima AlpK (ciano), PgaE (castanho claro), CabB (branco), aklavinona (magenta) e rifampicina monooxygenase (lilás). As linhas amarelas mostram as distâncias entre aminoácidos na mesma posição. Os cofatores e moléculas de água foram apagados.

Os valores calculados de TM mostram que os alinhamentos feitos não satisfazem os valores de referência para proteínas com alta similaridade estrutural, já que não superam 0,5, sendo então considerados de similaridade randômica. Como os modelos 2 e 3 mostram vantagens na densidade do cluster e maiores valores de C-score, fizeram-se os cálculos dos valores de TM com as enzimas de maior similaridade com o modelo 1, AlpK e aklavinona (Tabela 4). Observou-se que os valores de TM aumentaram para o caso da enzima AlpK nos modelos 2 e 3 enquanto que os valores para a enzima aklavinona-11 hidroxilase diminuíram 12,54% para o modelo 2 e 9,96% para o modelo 3, entretanto ainda não são aceitáveis podendo-se concluir que as enzimas apesar de ser estruturalmente semelhantes não tem a mesma função catalítica, o que explica a baixa porcentagem de identidade.

Tabela 4. Valores de TM calculados para os alinhamentos entre os modelos 1, 2 e 3 e as enzimas AlpK e aklavinona

Classificação	PDB	Nome	TM-score		
			Modelo 1	Modelo 2	Modelo 3
1	6j0zC	AlpK	0,1581	0,1622	0,1672
4	3ihgA	Aklavinona-11 hidroxilase	0,2280	0,1994	0,2053

A tabela 5 apresenta as porcentagens de variação das contribuições de  $\alpha$  – hélices, folhas  $\beta$  e estruturas randômicas calculadas com os dados de dicrósmo circular e os valores de estrutura secundária dos modelos propostos pelo I-TASSER usando o software STICKS com o objetivo de propor uma estrutura tridimensional pra a enzima SalC.

Tabela 5. Análise das diferenças entre as contribuições estruturais calculadas a partir do método K2D e comparadas com as porcentagens das estruturas secundárias do modelo 1 obtido do software I-TASSER e calculados com o método de STICKS (M1: Modelo 1, M2: Modelo 2, M3: Modelo 3, M4: modelo 4, M5, modelo 5)

	SalC	M 1	% Erro	M 2	% Erro	M 3	% Erro	M 4	% Erro	M 5	% Erro
<b><math>\alpha</math>-Hélice (%)</b>	23	34,7	0,34	30,6	0,25	34,9	0,34	30,2	0,24	32,2	0,29
<b>Folhas-<math>\beta</math> (%)</b>	24	18	0,33	27,5	0,13	25,6	0,06	26,9	0,11	21,7	0,11
<b>Randômica (%)</b>	53	47,3	0,12	41,9	0,26	39,5	0,34	43	0,23	46,1	0,15

Como discutiu-se na seção 5.3.1, os valores de C-score, a densidade do cluster e considerando que o valor de TM no modelo 1 não satisfaz o critério de topologia correta, guiando-se pelo valores de erro da comparação das porcentagens de estrutura secundária *in vitro* e *in silico*, pode-se concluir que o modelo 3, que tem o maior valor de densidade do cluster e um valor médio de C-score, acomoda-se melhor aos valores obtidos *in vitro*. A estrutura com os valores obtidos é apresentada na sequência:

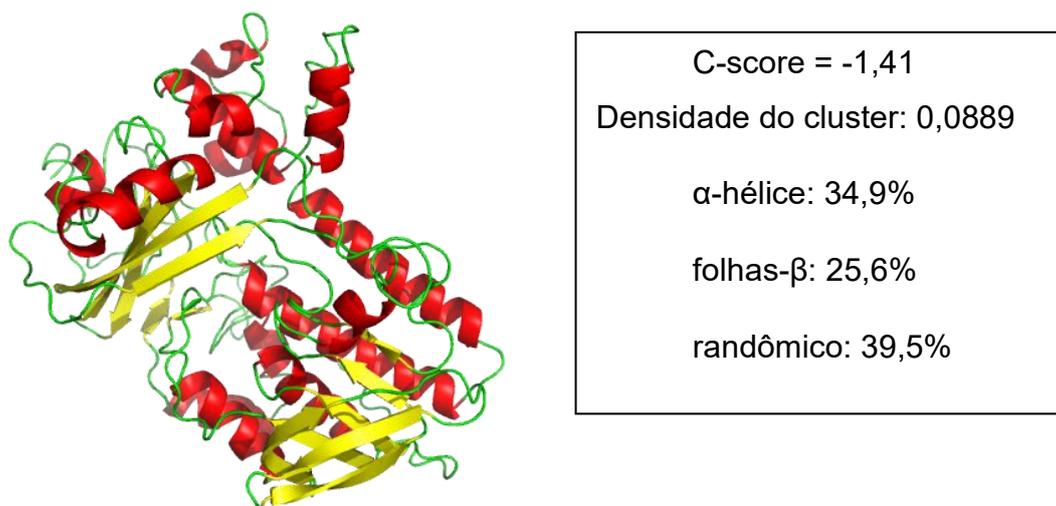


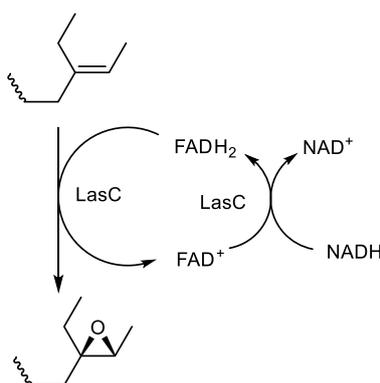
Figura 24. Estrutura tridimensional da enzima SalC com os valores considerados para sua escolha.

### 5.2.3 Caracterização da atividade enzimática

Usando a ferramenta BLAST, procuraram-se os domínios conservados da enzima SalC, onde um domínio de ligação de FAD compreendido entre as posições 16 – 374 da sequência da enzima, um sítio conservado nas enzimas que pertence à grande família de monooxygenases dependentes de FAD, o que encontra-se em acordo com a caracterização estrutural *in-silico* da enzima gerada pelo software I-TASSER, onde as proteínas propostas são dependentes de FAD. Tendo em conta o sítio ativo em comum, foi feito um BLAST comparando as sequências das proteínas com analogia estrutural e com a enzima LasC, uma enzima que exerce papel similar a SalC na biossíntese da lasalocida como descrito no trabalho de Atsushi e colaboradores<sup>9</sup>, na formação estereosseletiva de um bis-epóxido no intermediário linear pré-lasalocida, o qual posteriormente é ciclizado pela enzima LasB para gerar os grupos pirano e furano na molécula<sup>9</sup>. A Tabela 3A (Anexos) apresenta os valores de identidade, similaridade, cobertura da sequência e o valor de expectativa, E, que indica a probabilidade de encontrar no banco de dados um tamanho específico de sequência. Esses dados mostram que a enzima mais relacionada com a SalC é a monooxygenase dependente de FAD LasC, com valores de 75% de similaridade, que indica que um 75% dos aminoácidos presentes nas sequências das enzimas compartilham

a mesma natureza química. As sequências também mostram uma identidade de 61%, indicando que as sequências compartilham uma alta porcentagem de aminoácidos idênticos. Entretanto, 4 das 5 enzimas com analogia estrutural mostram uma baixa cobertura, variando entre 6% e 35%, valores que, junto com o valor de  $E > 10^{-4}$  indicam que as sequências não são semelhantes e os valores de identidade e similaridade não são confiáveis. A única enzima estruturalmente similar com alta cobertura é AlpK.

Considerando que as enzimas com valores de identidade superiores de 30% em geral exercem a mesma função, e valores de  $E < 10^{-4}$  mostram uma alta confiabilidade nos valores medidos, a enzima LasC pode descrever corretamente a função esperada de SalC. Entretanto, para o estudo dos locais de interação com o cofator será considerada a estrutura de AlpK, já que a enzima LasC não tem uma estrutura descrita na literatura. A enzima LasC como foi explicado anteriormente, é uma epoxidase que depende de FAD. Como epoxidase, classifica-se na família de óxido-redutases e o Esquema 7 mostra a reação de epoxidação proposta por Atsushi et al.<sup>9</sup>, onde a enzima consome NADH, que é o doador dos elétrons que são recebidos pelo FAD para gerar as ligações envolvidas na reação de epoxidação. Nesse sentido, espera-se que a enzima SalC funcione como uma óxido-redutase (epoxidase) dependente de FAD e consome NADH na reação.



Esquema 7. Ciclo de reações proposto para a enzima LasC (Adaptado de Atsushi e colaboradores<sup>9</sup>)

### 5.3 Estudos de interação da molécula FADH<sub>2</sub> com a enzima SaIC

Amparando-se na estrutura tridimensional que parece melhor representar SaIC e unindo essas informações aos dados de dicroísmo circular e I-TASSER discutido na seção 5.3, estudou-se a interação entre a enzima SaIC com a coenzima FADH<sub>2</sub> e suas diferentes conversões, hidroxiflavina, peroxiflavina, flavina oxidada e flavina reduzida. Como foi tratado na introdução, o ciclo da flavina já foi estudado na enzima estireno monooxigense e baseando-se nesse modelo, foram simuladas as interações usando o software Chimera 1.13.1e AutoDock Vina. O ciclo começa com uma interação forte do FADH<sub>2</sub> no sítio de ligação da flavina, após da oxidação nucleofílica, obtém-se uma peroxiflavina que tem a capacidade de doar um oxigênio na formação do grupo epóxido no substrato, quando essa doação é feita, forma-se uma hidroxiflavina a qual após reação de condensação forma uma flavina oxidada (FAD). O ciclo é repetido quando o NAD(P)H é oxidado pela enzima para reduzir o FAD para FADH<sub>2</sub>. Segundo a revisão feita por Montersino e colaboradores<sup>10</sup>, quando a flavina encontra-se oxidada, a interação com o sítio catalítico das enzimas FAD-dependentes diminui até atingir um equilíbrio entre enzima livre de FAD e enzima com FAD<sup>10</sup>. Baseando-se nos sítios de ligação de flavina obtidos pelos Softwares I-TASSER e BLAST (Tabela 6 e 7 e Figura 16) foi definido um *Grid* com centro em (62,86, 61,63, 65,94) e com dimensões (36,99, 27,41, 18,86), onde encontra-se o sítio catalítico para o FADH<sub>2</sub> (Figura 25).

Tabela 6. Domínios conservados na sequência de aminoácidos da enzima SalC

Nome	Código	Descrição	Intervalo	Valor de E
PRK06185	PRK06185	óxido-redutase FAD-dependente	132-339	6.91e-09

Pssm-ID: 235729 [Multi-domain] Cd Length: 407 Bit Score: 57.56 E-value: 6.91e-09

```

      10      20      30      40      50      60      70      80
.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|
Query 17208 132 RRQALAEPIIVRQETEVVLGGLGGAGQVTGVRLRDRSGEStELAADLVVDAGGRASALRRwlpELGLDQVEE----DVVD 207
Cdd:PRK06185 115 AEEASAYPNFTLRMGAETGLIEEGGRVTGVRARTPDGP-GEIRADLVVGADGRHSRVRA---LAGLEVREFFgapmDVL- 189
      90     100     110     120     130     140     150     160
Query 17208 208 sgiayatrVFKAPAPVAQGfpmvniAAAPGLGKPGQNGALVPIEDgKWLVTLAGTRGGEPTDdsfldfARGL--RHP 284
Cdd:PRK06185 190 -----WFRLPREPDDP-----ESLMGRFGPGQGLIMIDRGD-YWQCGYVIPKGGYAALR-----AAGLeafRER 247
      170     180     190     200     210
Query 17208 285 VLADLLERAEPVKGSRSTVNRRLYYDRVANW-PDGLLVLDALAAFNPFVYGHG 339
Cdd:PRK06185 248 VAELAPELADRVAELKSWDDVKLLDVRVDRLRWhRPGLLCIGDAAHAMSPVGGVG 303

```

FAD Binding	pfam01494	Domínio de ligação de FAD	15-340	1.34e-07
-------------	-----------	---------------------------	--------	----------

Pssm-ID: 366675 [Multi-domain] Cd Length: 345 Bit Score: 53.09 E-value: 1.34e-07

```

      10      20      30      40      50      60      70      80
.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|
Query 17208 15 ETHAVVLGGGLTGMLTSAVLAHLDKVTVIERdvlpegpeLRKGVPAQARhAHLWSSGGARAIDSVlpGTVKQLIAEGA-- 92
Cdd:pfam01494 1 ETDVLIVGAGPAGLMLALLARAGVRVVLVER-----HATTSVLP- AHGLNQRMTMELLRQA--GLEDRILAEGVph 69
      90     100     110     120     130     140     150     160
Query 17208 93 ---HRLYLPRDVVWLTTPH-GWQHRFSGSQFMVTcsRALLDWVVRQAaepkiTVRQETEVVLGGLGGAGQVTGVRLRDR 167
Cdd:pfam01494 70 egmgIAFYNP DARADLDLFLtSPPRVTVYYPQTELE--PILVEHAERGA-----QVRFGEVLSLEQDGDGVTAVVRDRR 141
      170     180     190     200     210     220     230     240
Query 17208 168 SGESTELAADLVVDAGGRASALRRwlpELGLDQVEEdvdsGIAYATRVFKAPAPVAQGFPMVNIAApgLGKPGQNGAL 247
Cdd:pfam01494 142 DGEYTVRAKYLVGCDGGRSPVRK---TLGIEFEGEE---GIFGSLDVLFDAPDLSDFVERAFVHYL--IYAPHSRGFV 212
      250     260     270     280     290     300     310     320
Query 17208 248 VPIEDGK---WLVTLAGTRGGEPTdDdsFLDFARGLRHFPVLADLLERAEPVKGSRSTVNRRLYYDRvanwpdgl 322
Cdd:pfam01494 213 GPWRSYGriryVQCPWDEEVEERPE---EFTDEELKQLRSLVIGTDLAPVEIDWkSIWCVRSRVATRYRKGVRV----- 282
      330
Query 17208 323 LVLGDALAAFNPFVYGHGM 340
Cdd:pfam01494 283 FLAGDAAHIHPTGGQGL 300

```

Tabela 7. Sítios de ligação para a enzima SalC

Rango	Valor de C	Tamanho do cluster	Nome do ligando	Resíduos do sítio de ligação
1	0,75	219	FAD	21, 22, 23, 24, 25, 26, 44, 45, 46, 47, 60, 65, 67, 124, 147, 148, 149, 182, 183, 184, 188, 211, 306, 326, 327, 334, 337, 338, 339, 340, 341

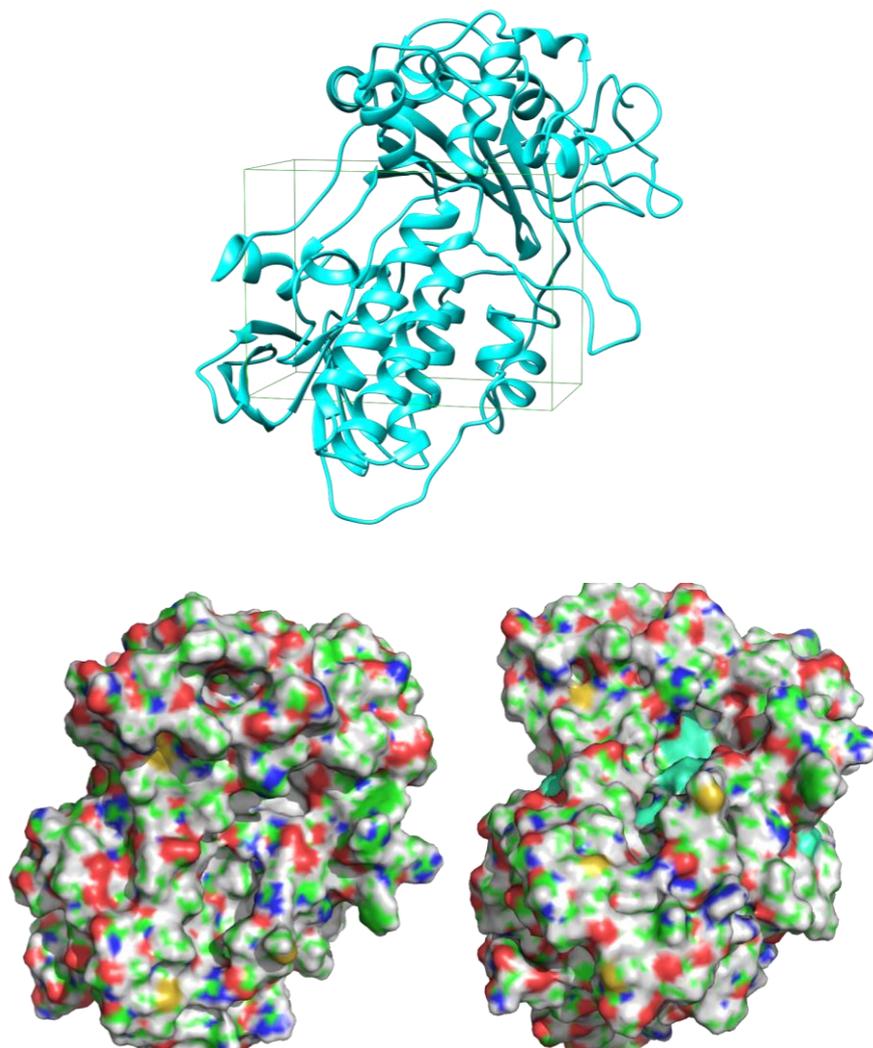


Figura 25. **Acima:** Enzima Sal C (ciano) com o *Grid* escolhido (verde) para o estudo de *docking*. **Abaixo:** Superfície da enzima Sal C (esquerda) onde está ressaltado o bolso de ligação proposto para o FADH<sub>2</sub> (direita).

Com as condições anteriormente descritas, foram obtidos diversos complexos SalC-Flavina que cumprem com as condições de energia,  $\Delta G$  e sítio de ligação. A Tabela 8 apresenta os valores dos complexos mencionados previamente. O único complexo que não apresenta ligações de hidrogênio para estabilizar a interação está formado entre a enzima e o cofator oxidado (SalC – FAD), resultado que concorda com o apresentado por Montersino<sup>10</sup>, ressaltando que as monooxigenases FAD-dependentes entram em equilíbrio com o cofator quando ele se encontra oxidado.

Tabela 8. Propriedades energéticas dos complexos obtidos

<b>Complexo Enzima – Coenzima</b>	<b>Energia [J]</b>	<b><math>\Delta G</math> [J]</b>
<b>SalC – FAD</b>	-0.407	-8.339
<b>SalC – FADH<sub>2</sub></b>	-1.764	- 10.658
<b>SalC – Peroxisflavina</b>	7.944	-9.902
<b>SalC – Hidroxisflavina</b>	-9.455	-9.884

O ciclo da flavina proposto começa com um equilíbrio SalC – FAD (Figura 26), já que a ausência de ligações de H diminui a estabilidade do complexo, essa observação condiz com o dado obtido pela energia do complexo, -0.407 J, indicando que a enzima recebe pouca estabilidade pela interação com a flavina oxidada.

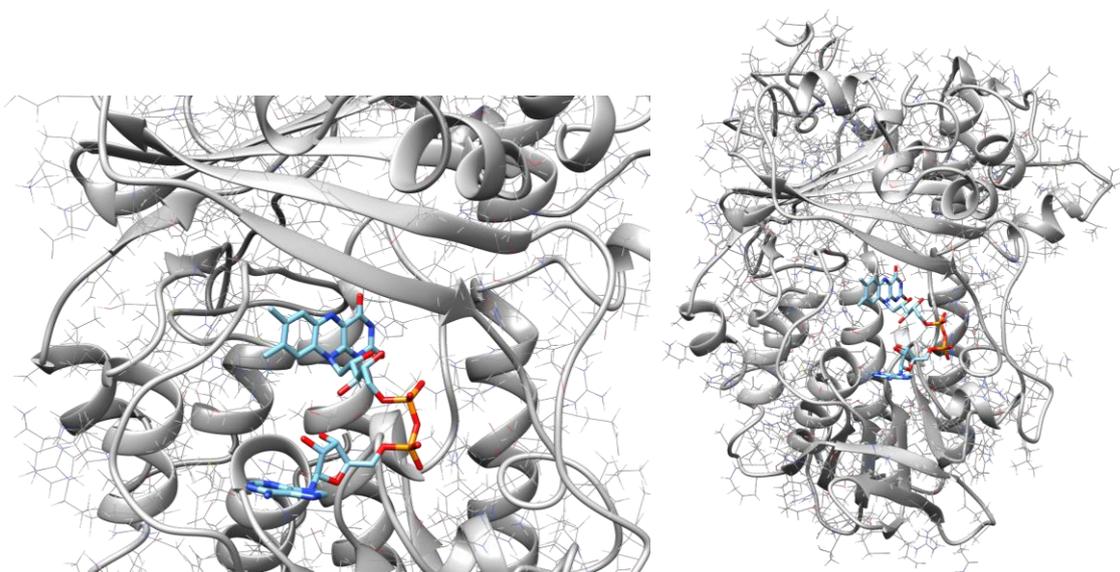


Figura 26. Complexo SalC – FAD.

Após a redução da flavina, o complexo obtém uma diminuição de energia de -1.764 J, indicando que a estrutura entra em um estado mais relaxado e mais espontâneo devido a sua energia livre. A energia da interação SalC – FADH<sub>2</sub>

diminui pelas ligações de hidrogênio que aparecem após da redução, 5 no total, que estabilizam o complexo e permitem a oxidação nucleofílica do cofator para continuar com a etapa seguinte.

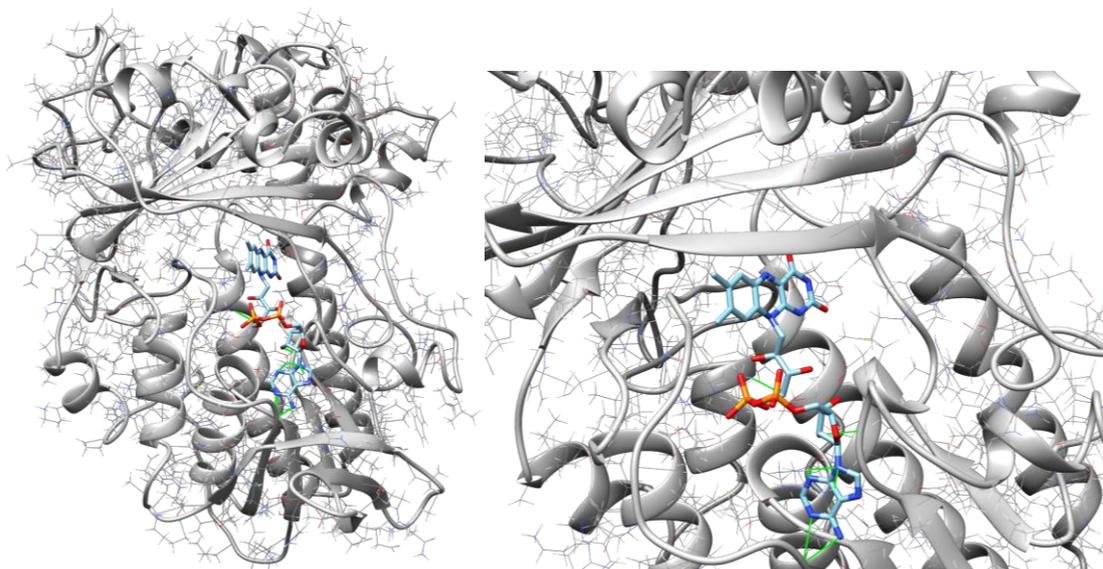


Figura 27. Complexo SalC - FADH<sub>2</sub>. As ligações de hidrogênio formadas são apresentadas em verde.

O complexo SalC – peroxiflavina formado após a oxidação nucleofílica aumenta a energia do complexo para 7.944 J mas a interação continua sendo espontânea, com uma energia livre de -9.902 J. Esse aumento de energia deve-se a mudança conformacional na adenina do cofator, que muda sua localização espacial para disponibilizar o grupo peróxi- estabilizando o complexo (Figura 28).

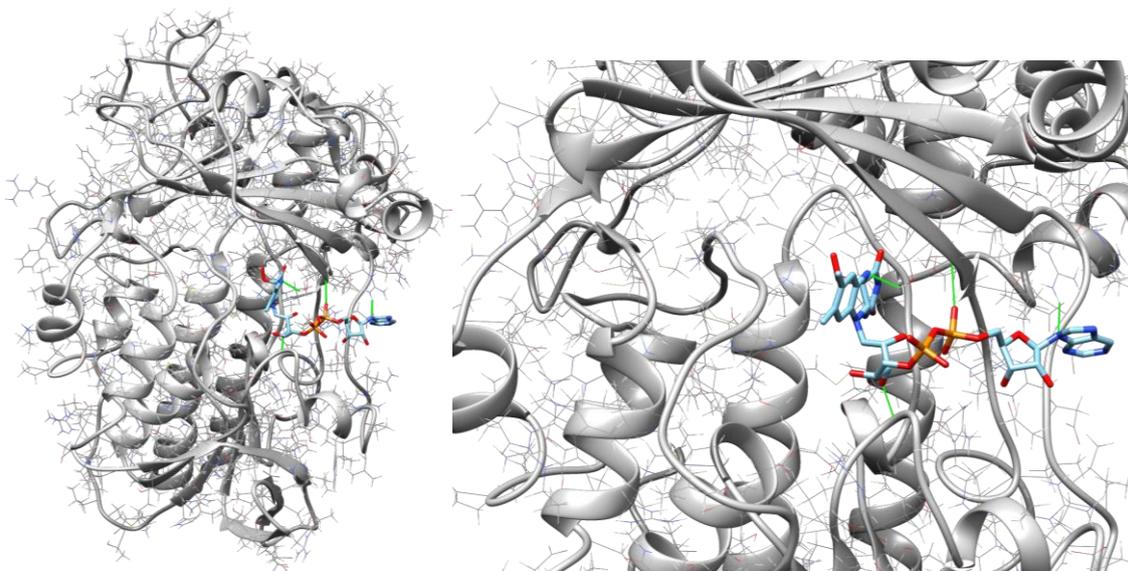


Figura 28. Complexo SalC – peroxiflavina. As ligações de hidrogênio são mostradas em verde.

Finalmente, a formação da hidroxiflavina diminui a energia do complexo para  $-9.455$  J, indicando que a transferência do oxigênio para o substrato favorece a formação do novo complexo.

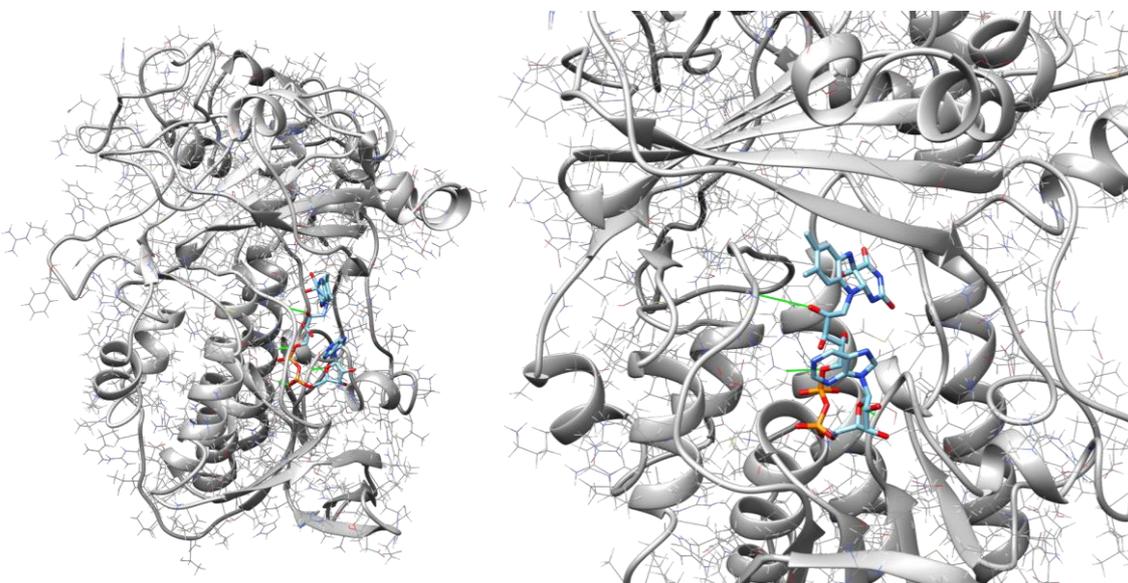


Figura 29. Complexo SalC - Hidroxiflavina. As ligações de hidrogênio são mostradas em verde.

## 6 Conclusões

- Obteve-se sucesso na produção da enzima SalC cataliticamente ativa e com ela promoveu-se a oxidação de diferentes substratos para estudar a promiscuidade da enzima e sua interação com o intermediário linear obtido da cultura da bactéria *S. albus*  $\Delta salC$ . Foi possível propor a obtenção do epóxido 2-(oxiran-2-il)-etan-1-ol e do bis-epóxido, intermediário da biossíntese da salinomicina, que foram caracterizados utilizando as técnicas de GC-MS e LC-EM de baixa resolução respectivamente.
- Foi possível, a partir de análises *in silico* (I-TASSER) e experimentos de dicroísmo circular, sugerir uma estrutura tridimensional para a enzima SalC. Com essa estrutura foi possível fazer um estudo *in silico* do ciclo catalítico do cofator FADH<sub>2</sub>, onde foi corroborado o ciclo catalítico característico das Baeyer-Villiger monooxigenases.
- Determinou-se a importância da molécula de NADH no meio reacional, mostrando que na ausência desse reagente não é possível observar ação catalítica por parte da enzima.
- Juntando as informações obtidas em estudos prévios com as observações feitas nesta pesquisa, foi esclarecido como atua a enzima SalC na epoxidação do intermediário linear na biossíntese do poliéter salinomicina. As contribuições com a elucidação biossintética facilitarão o estudo estruturas análogas com menor toxicidade.
- Os estudos enzimáticos demonstram a possibilidade de aplicação biotecnológica de SalC em sistemas de reações tipo *cell free*.

## 7 Referências

1. Vera B, Rodríguez AD, Romo D, et al. Simultaneous structure–activity studies and arming of natural products by C–H amination reveal cellular targets of eupalmerin acetate. *Nat Chem.* 2013; 5(6):510-517. doi:10.1038/nchem.1653
2. Harvey AL, Edrada-Ebel R, Quinn RJ. The re-emergence of natural products for drug discovery in the genomics era. *Nat Rev Drug Discov.* 2015; 14(2):111-129. doi:10.1038/nrd4510
3. Newman DJ, Cragg GM. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J Nat Prod.* 2020; 83(3):770-803. doi:10.1021/acs.jnatprod.9b01285
4. Luhavaya H, Williams SR, Hong H, Gonzaga De Oliveira L, Leadlay PF. Site-specific modification of the anticancer and antituberculosis polyether salinomycin by biosynthetic engineering. *ChemBioChem.* 2015; 16(17):2081-2085. doi:10.1002/cbic.201402300
5. Yurkovich ME. Characterisation of the Biosynthetic Gene Cluster for Salinomycin : An Ionophoric Polyether Inhibitor of Cancer Stem Cells Marie Eileen Yurkovich Department of Biochemistry University of Cambridge. (August 2011):1-126.
6. Antoszczak M, Popiel K, Stefańska J, et al. Synthesis, cytotoxicity and antibacterial activity of new esters of polyether antibiotic - Salinomycin. *Eur J Med Chem.* 2014; 76:435-444. doi:10.1016/j.ejmech.2014.02.031
7. Yurkovich ME, Tyrakis PA, Hong H, et al. A Late-Stage Intermediate in Salinomycin Biosynthesis Is Revealed by Specific Mutation in the Biosynthetic Gene Cluster. *ChemBioChem.* 2012; 13(1):66-71. doi:10.1002/cbic.201100590
8. Lin H, Liu JY, Wang HB, Ahmed AAQ, Wu ZL. Biocatalysis as an alternative for the production of chiral epoxides: A comparative review. *J Mol Catal B Enzym.* 2011; 72(3-4):77-89. doi:10.1016/j.molcatb.2011.07.012
9. Minami A, Shimaya M, Suzuki G, et al. Sequential enzymatic epoxidation involved in polyether lasalocid biosynthesis. *J Am Chem Soc.* 2012; 134(17):7246-7249. doi:10.1021/ja301386g
10. Montersino S, Tischler D, Gassner GT, Van Berkel WJH. Catalytic and structural features of flavoprotein hydroxylases and epoxidases. *Adv Synth Catal.* 2011; 353(13):2301-2319. doi:10.1002/adsc.201100384
11. Oliynyk M, Stark CBW, Bhatt A, et al. Analysis of the biosynthetic gene cluster for the polyether antibiotic monensin in *Streptomyces cinnamomensis* and evidence for the role of monB and monC genes in oxidative cyclization. *Mol Microbiol.* 2003; 49(5):1179-1190. doi:10.1046/j.1365-2958.2003.03571.x
12. Bhatt A, Stark CBW, Harvey BM, et al. Accumulation of anE,E,E-Triene by the Monensin-Producing Polyketide Synthase when Oxidative Cyclization is Blocked. *Angew Chemie.* 2005; 117(43):7237-7240.

doi:10.1002/ange.200501757

13. Smith L, Hong H, Spencer JB, Leadlay PF. Analysis of specific mutants in the lasalocid gene cluster: evidence for enzymatic catalysis of a disfavoured polyether ring closure. *ChemBiochem*. 2008; 9(18):2967-2975. doi:10.1002/cbic.200800585
14. Migita A, Watanabe M, Hirose Y, et al. Identification of a gene cluster of polyether antibiotic lasalocid from *Streptomyces lasaliensis*. *Biosci Biotechnol Biochem*. 2009; 73(1):169-176. doi:10.1271/bbb.80631
15. Mai TT, Hamaï A, Hienzsch A, et al. Salinomycin kills cancer stem cells by sequestering iron in lysosomes. *Nat Chem*. 2017; 9(10):1025-1033. doi:10.1038/nchem.2778
16. Huczynski A, Antoszczak M, Kleczewska N, et al. Synthesis and biological activity of salinomycin conjugates with floxuridine. *Eur J Med Chem*. 2015; 93:33-41. doi:10.1016/j.ejmech.2015.01.045
17. Whitmore L, Wallace BA. Protein Secondary Structure Analyses from Circular Dichroism Spectroscopy: Methods and Reference Databases. Biopolymers.
18. Whitmore L, Wallace BA. Dichroweb. 2004:W668-673.
19. Kelly SM, Jess TJ, Price NC. How to study proteins by circular dichroism. *Biochim Biophys Acta - Proteins Proteomics*. 2005; 1751(2):119-139. doi:10.1016/j.bbapap.2005.06.005
20. Lehninger, Albert L., Cox, Michael M. Nelson DL. *Lehninger Principles of Biochemistry*. 6a ed. New York: W.H. Freeman and Company; 2008.
21. VOET, Donald; VOET JG. *Bioquímica*. 4 ed. Artem; 2013.
22. Louis-Jeune C, Andrade-Navarro MA, Perez-Iratxeta C. Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins Struct Funct Bioinforma*. 2012; 80(2):374-381. doi:10.1002/prot.23188
23. Perez-Iratxeta C, Andrade-Navarro MA. K2D2: Estimation of protein secondary structure from circular dichroism spectra. *BMC Struct Biol*. 2008; 8:1-5. doi:10.1186/1472-6807-8-25
24. Böhm G, Muhr R, Jaenicke R. Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng Des Sel*. 1992; 5(3):191-195. doi:10.1093/protein/5.3.191
25. Gasteiger E, Hoogland C, Gattiker A, et al. Protein Identification and Analysis Tools on the ExPASy Server. In: *The Proteomics Protocols Handbook*.; 2005:571-607.
26. A Roy, A Kucukural YZ. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc*. 2010; 5:725-738.
27. J Yang, R Yan, A Roy, D Xu, J Poisson YZ. The I-TASSER Suite: Protein structure and function prediction. *Nat Methods*. 2015; 12:7-8.

28. J Yang YZ. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res.* 2015;W174-W181.
29. Yang J, Wang Y, Zhang Y. ResQ: An approach to unified estimation of B-factor and residue-specific error in protein structure prediction. *J Mol Biol.* 2016; 428(4):693–701. doi:10.1016/j.physbeh.2017.03.040
30. Martin J, Letellier G, Marin A, Taly JF, De Brevern AG, Gibrat JF. Protein secondary structure assignment revisited: A detailed analysis of different assignment methods. *BMC Struct Biol.* 2005; 5:1-17. doi:10.1186/1472-6807-5-17
31. Majumdar I, Krishna SS, Grishin N V. PALSSE: A program to delineate linear secondary structural elements from protein structures. *BMC Bioinformatics.* 2005; 6:1-24. doi:10.1186/1471-2105-6-202
32. Taylor WR, Aszódi A. *Protein Geometry, Classification, Topology and Symmetry.* 1 ed. New York: Taylor & Francis Group; 2005.
33. Drew ED, Janes RW. PDBMD2CD : providing predicted protein circular dichroism spectra from multiple molecular dynamics-generated protein structures. 2020; 175(7):1-8. doi:10.1093/nar/gkaa296
34. J. Xu YZ. How significant is a protein structure similarity with TM-score=0.5? *Bioinformatics.* 2010; 26:889-895.
35. Y. Zhang JS. Scoring function for automated assessment of protein structure template quality. *Proteins.* 2004; 57:702-710.
36. MILLER, JAMES N.; MILLER JC. *ESTADISTICA Y QUIMIOMETRIA PARA QUIMICA ANALITICA.* Madrid: PEARSON EDUCACION: PRENTICE-HALL; 2002.

## 8 Anexos

Tabela 1A. Valores de C, TM, RMSD e densidade do cluster calculado pela ferramenta I-TASSER para os modelos gerados a partir da sequência de aminoácidos da enzima SalC. Os valores de TM e RMSD para os modelos 2 e 3 não foram calculados

Nome	C-Score	Exp. TM-Score	Exp. RMSD	Decoys	Densidade do cluster
Modelo 1	-1,94	0,48 +- 0,15	11,9 +- 4,4	819	0,0523
Modelo 2	-1,74	-	-	600	0,0637
Modelo 3	-1,41	-	-	600	0,0889
Modelo 4	-2,23	-	-	466	0,0392
Modelo 5	-2,33	-	-	514	0,0355

Tabela 2A. Porcentagens de contribuição geradas pelos diferentes métodos de quantificação de estrutura secundária para os diferentes modelos gerados pela ferramenta I-TASSER para a enzima SalC

<b>Modelo 1</b>								
<b>Método</b>	DSSP	P-SEA	PALSSE	STICKS	STRIDE	XTLSSTR	PyMOL	Simulação CD
<b><math>\alpha</math>-Hélice (%)</b>	27,3	26,4	43,6	34,7	28,9	24,8	33,26	27,27
<b>Folhas-<math>\beta</math> (%)</b>	6,6	13	27,5	18	8,3	9,7	10,54	6,62
<b>Randômico (%)</b>	66,1	60,5	28,9	47,3	62,8	65,5	56,2	66,11
<b>Modelo 2</b>								
<b>Método</b>	DSSP	P-SEA	PALSSE	STICKS	STRIDE	XTLSSTR	PyMOL	Simulação CD
<b><math>\alpha</math>-Hélice (%)</b>	31,6	31	47,3	30,6	33,9	30,4	28,72	30,17
<b>Folhas-<math>\beta</math> (%)</b>	12,4	25	30	27,5	12,6	11,8	6,4	15,71
<b>Randômico (%)</b>	56	44	22,7	41,9	53,5	57,9	64,88	57,44
<b>Modelo 3</b>								
<b>Método</b>	DSSP	P-SEA	PALSSE	STICKS	STRIDE	XTLSSTR	PyMOL	Simulação CD
<b><math>\alpha</math>-Hélice (%)</b>	30	31	45,7	34,9	31,4	24,6	31,82	28,92
<b>Folhas-<math>\beta</math> (%)</b>	18,4	23,3	29,5	25,6	18,4	13,6	15,5	18,38
<b>Randômico (%)</b>	51,7	45,7	24,8	39,5	50,2	61,8	52,69	52,69
<b>Modelo 4</b>								
<b>Método</b>	DSSP	P-SEA	PALSSE	STICKS	STRIDE	XTLSSTR	PyMOL	Simulação CD
<b><math>\alpha</math>-Hélice (%)</b>	30,8	27,3	49,4	30,2	28,5	23,8	31,82	30,79
<b>Folhas-<math>\beta</math> (%)</b>	14,3	21,5	26	26,9	14,9	14,9	13,84	14,25
<b>Randômico (%)</b>	55	51,2	24,6	43	56,6	61,4	54,34	54,96
<b>Modelo 5</b>								
<b>Método</b>	DSSP	P-SEA	PALSSE	STICKS	STRIDE	XTLSSTR	PyMOL	Simulação CD
<b><math>\alpha</math>-Hélice (%)</b>	30	26,2	42,4	32,2	30	25	31,4	30,79
<b>Folhas-<math>\beta</math> (%)</b>	15,3	20,5	36	21,7	17,4	14,5	13,22	14,25
<b>Randômico (%)</b>	54,8	53,3	21,7	46,1	52,7	60,5	55,37	54,96

Tabela 3A. Estudo da homologia das enzimas com analogia estrutural e da enzima LasC com a enzima SalC

<b>Nome</b>	<b>Cobertura</b>	<b>E</b>	<b>Identidade</b>	<b>Similaridade</b>
<b>AlpK</b>	53%	4E-79	48%	68%
<b>PgaE</b>	6%	0,004	44%	62%
<b>CabE</b>	19%	0,09	27%	53%
<b>Aklavinone-11 hidroxidase</b>	35%	0,51	45%	65%
<b>Rifampicina monooxigenase</b>	22%	1,7	35%	56%
<b>LasC</b>	43%	3 E-86	61%	75%